

A New Method to Build an Associative Memory Model

Arturo Gamino Carranza, and Juan Luis Díaz de León

Abstract—An associative memory is an artificial neural network model designed to store and recall input-output patterns pairs by association. A new method for obtaining two associative memories ${}^+M$ and ${}^-M$ is presented in this article, which uses a framework focused on two new binary operations \boxplus and \boxminus , and a unary operation called projection \downarrow . The conditions to obtaining perfect recovery and the noise boundaries that the memories can tolerate are studied. Memories ${}^+M$ and ${}^-M$ are robust to additive and subtractive noise, respectively, both types converge in one-step and operate in heteroassociative and autoassociative modes. The performance of the proposed memories is tested against other memory models under identical conditions using the gamma binary distance for measuring similarity between binary patterns. The computer simulation results based on the average of 500 trials with the binary set of 26 lowercase letters of 7x7 pixel size, showed that the proposed memories in autoassociative mode exhibited a gamma binary distance of above .8, .75 and .7 by distorting up to 10 pixels by additive, subtractive, and mixed noise, respectively, which implied that the recovered images had a high similarity. In absence of noise the performance was excellent, i.e., the 100% of the recovered images were identical.

Index Terms—Artificial neural networks, associative memory, binary operation, morphological associative memory, perfect recall.

I. INTRODUCCIÓN

Una memoria asociativa es un sistema que guarda la información, la mantiene de forma segura y la devuelve de forma eficiente [1], es por ello que las memorias son modelos que almacenan las asociaciones entre los patrones de entrada y salida para qué, posteriormente recuperen patrones a partir de entradas sin alterar o de versiones distorsionadas. A través del tiempo se han propuestos diversos modelos de memorias, en 1956 se introdujo un modelo que aprendía a asociar patrones a través de representaciones repetitivas de pares de patrones [2]. En 1961 se presentó la *learnmatrix* como un arreglo de interruptores receptores y actuadores para el proceso de aprendizaje [3]. En 1972 se desarrolló un modelo basado en una matriz de correlaciones, el cual almacena la información de los productos entre cada vector llave y su vector asociado [4], y en ese mismo año se describió el *associatron*, como un modelo que almacena los patrones mediante la suma de matrices de auto-correlación de sus vectores [5].

En el año de 1982 se definió uno de los modelos autoasociativo probablemente más estudiado, la red de Hopfield, la cual es una memoria donde la salida es calculada recursivamente

hasta que el sistema alcanza un punto de estabilidad [6], el número de patrones almacenados en una red de este tipo que se pueden recuperar sin error, no puede ser asintóticamente mayor que $n/(4 \log n)$ a medida que n tiende a infinito, siendo n el número de neuronas de la red [7].

Otro enfoque, son las memorias asociativas difusas (FAM), las cuales se fundamentan en el marco de la teoría de conjuntos difusos, el primer modelo de este tipo utilizó el operador mínimo para el almacenamiento de los patrones y para su recuperación [8]. En 1994 se presentó la memoria difusa relacional (FRM) como una sola matriz que empleó la composición max- t , esta memoria recordaba todas las reglas difusas sólo si los patrones de entrada constituían un conjunto de vectores ortonormales [9]. En 1997 en [10], se describió un modelo que utilizó un algoritmo de codificación max-min para el aprendizaje de las memorias difusas de multiplicación max (Codificación max-min para la FMM).

En el año de 1998 se presentaron las memorias asociativas morfológicas (MAM), este modelo basado en un álgebra minimax, en su fase de aprendizaje emplea el máximo (o mínimo) de restas entre los patrones de entrada y salida para construir la matriz de la memoria, mientras que en la fase de recuperación usa el operador mínimo (o máximo) de sumas entre la matriz y el patrón presentado a la memoria para recuperar un patrón [11]. Inspirados en las MAM se han propuesto modelos similares, en el 2009 se desarrolló un nuevo método para su obtención (NMMAM), el cual propuso intercambiar las operaciones de las fases de aprendizaje y recuperación empleadas por las MAM [12]. En el 2015 en [13] se presentaron memorias que almacenan contenido-asociación de patrones a través de un factor de control de proporción η , este factor se introduce como una multiplicación con el vector de entrada en la fase de aprendizaje y recuperación de las MAM. En 2016 en [14] se describió un modelo que reemplazan las operaciones utilizadas por las MAM (NR^2FMAM), a saber, se sustituyó la operación resta por la operación multiplicación en la fase de aprendizaje y la operación suma por la operación división en la fase de recuperación. Otro modelo relacionado con las MAM se desarrolló en la teoría de conjuntos difusos [15], el cual utiliza el aprendizaje implicativo de R -implicaciones de Lukasiewicz, Gödel y Goguen, mientras que en la recuperación se emplean las operaciones duales de las R -implicaciones.

Otra red más compleja que funcionan en modo autoasociativo es el autoencoder binario neuronal [16], esta red está construida por tres capas (entrada-oculta-salida), la asociación entre los atributos de entrada y salida se realiza con funciones hash mediante la capa oculta, es decir, la capa oculta genera la codificación binaria hash a L bits, donde L es la longitud del código binario y en la práctica se le asignan valores de 8,

Arturo Gamino Carranza, National Technological Institute of Mexico, Mexico City, Mexico, e-mails: d_docencia01@tecnm.mx, arturogamino@hotmail.com.

Juan Luis Díaz de León, Center for Computing Research, National Polytechnic Institute, Mexico City, Mexico, e-mail: jdiaz@cic.ipn.mx.

Corresponding author: Arturo Gamino Carranza.

12, 24, 32 o 48 bits.

Motivados en las ventajas que exhiben las MAM y modelo derivados [11], [17], [13], [18], [19], [20], como son la capacidad de almacenamiento ilimitada, convergencia en un solo paso, buen rendimiento de la memoria y tolerancia al ruido, en este artículo, se propone un método para obtener memorias asociativas binarias con estructura similar al de las MAM, estas memorias sustentan su marco de trabajo en la definición de dos nuevas operaciones binarias \boxplus , \boxminus y una unaria denominada proyección \downarrow .

Las operaciones binarias satisfacen las propiedades de distributividad respecto al operador máximo (o mínimo) y creciente por la derecha (o izquierda), la operación unaria asegura que los resultados sean binarios, y la concatenación de ambas (\boxplus, \downarrow o \boxminus, \downarrow) proporciona operaciones inversas. Estas propiedades son suficientes para caracterizar las soluciones de interés de las memorias propuestas. Las operaciones implican un enmascaramiento XOR, por lo que su implementación a nivel hardware es de bajo coste computacional y aplicable en problemas de gran escala.

La matriz de aprendizaje ${}^+M$ (o ${}^-M$) se construye calculando el máximo de los resultados de la operación \boxplus (o \boxminus) entre los patrones de salida y entrada, mientras que la recuperación de patrones se obtiene encontrando el mínimo de los resultados de la operación \boxminus (o \boxplus) entre la matriz y el patrón presentado a la misma. La matriz y el vector resultante, se les aplica la operación \downarrow . Las memorias ${}^+M$ son robustas al ruido aditivo y las ${}^-M$ al ruido sustractivo, ambas operan en modo autoasociativo o heteroasociativo, no tienen problemas de convergencia y su capacidad de almacenamiento es infinita en modo autoasociativo.

El artículo se encuentra organizado de la siguiente manera: en la Sección II se introducen los conceptos de memorias asociativas; en la Sección III se definen los fundamentos matemáticos de las memorias propuestas, además, se describe el método para el aprendizaje y recuperación de las memorias en los modos autoasociativo y heteroasociativo; en la Sección IV se exhiben los resultados experimentales en comparación con otras memorias, mientras que las conclusiones se presentan en la Sección V; finalmente, en el Apéndice A se muestran las demostraciones de los teoremas del artículo.

II. MEMORIAS ASOCIATIVAS

Las operaciones numéricas básicas de las memorias asociativas que emplean el máximo o mínimo de dos números, se representan por la notación $x \vee y$ y $x \wedge y$, respectivamente. Sean $A = \{0, 1\}$, \cdot una operación binaria, \mathbf{C} y \mathbf{D} dos matrices de dimensión $m \times p$, \mathbf{E} una matriz de dimensión $p \times n$, entonces en un sistema algebraico (A, \cdot, \vee, \wedge) se tiene lo siguiente:

- 1) El (i, j) -ésimo elemento de la matriz \mathbf{C} , se denota por c_{ij} .
- 2) El *máximo* y *mínimo* de las matrices \mathbf{C}, \mathbf{D} esta dado por $(\mathbf{C} \vee \mathbf{D})_{ij} = c_{ij} \vee d_{ij}$ y $(\mathbf{C} \wedge \mathbf{D})_{ij} = c_{ij} \wedge d_{ij}$, respectivamente.
- 3) La matriz \mathbf{C} se dice que es *menor o igual* que \mathbf{D} , representado por $\mathbf{C} \leq \mathbf{D}$, si y sólo si $c_{ij} \leq d_{ij}$, y la matriz \mathbf{C} se dice que es *estrictamente menor* que \mathbf{D} , representado por $\mathbf{C} < \mathbf{D}$, si y sólo si $c_{ij} < d_{ij}$.

- 4) El *producto máximo matricial* $\mathbf{D} \boxplus \mathbf{E}$ y el *producto mínimo matricial* $\mathbf{D} \boxminus \mathbf{E}$, para todo $i = 1, \dots, m$ y $j = 1, \dots, n$, están dados por

$$(\mathbf{D} \boxplus \mathbf{E})_{ij} = \bigvee_{k=1}^p (d_{ik} \cdot e_{kj}) \quad \text{y} \quad (\mathbf{D} \boxminus \mathbf{E})_{ij} = \bigwedge_{k=1}^p (d_{ik} \cdot e_{kj}) \quad (1)$$

Sean $I = \{1, \dots, n\}$ y $\mathbf{x} \in A^n$ un vector, entonces con base a la relación de orden establecida en el punto 3 de la lista anterior, se introducen tres tipo de ruido que alteran al vector \mathbf{x} [21]:

- 1) *Aditivo*: versión dilatada $\tilde{\mathbf{x}}$ del vector \mathbf{x} siempre que se cumpla que $\tilde{\mathbf{x}} \geq \mathbf{x}$, es decir, $\tilde{x}_i \geq x_i \forall i \in I$.
- 2) *Sustractivo*: versión erosionada $\tilde{\mathbf{x}}$ del vector \mathbf{x} siempre que se cumpla que $\tilde{\mathbf{x}} \leq \mathbf{x}$, es decir, $\tilde{x}_i \leq x_i \forall i \in I$.
- 3) *Mezclado*: versión distorsionada $\tilde{\mathbf{x}}$ del vector \mathbf{x} que incluye ruido aditivo y sustractivo al mismo tiempo, es decir, $\tilde{x}_i > x_i \forall i \in G$ y $\tilde{x}_i < x_i \forall i \in L$, con $L, G \subset I$ dos conjuntos no vacíos de índices y disjuntos entre sí.

Los conceptos y ecuaciones previas, son herramientas necesarias para establecer el marco teórico de las memorias asociativas presentadas en este trabajo. De acuerdo con [1], una memoria asociativa \mathbf{M} se considera como un sistema $\mathbf{x} \rightarrow \mathbf{M} \rightarrow \mathbf{y}$ el cual recupera un patrón de salida (vector columna \mathbf{y}) a partir de un patrón de entrada (vector columna \mathbf{x}). Sean $A = \{0, 1\}$ y $\xi, k, n, m \in \mathbb{Z}^+$, entonces:

- 1) k es el número pares de vectores binarios, ξ es un índice, y n, m la dimensión de los patrones de entrada y salida, respectivamente.
- 2) \mathbf{M} es una matriz que almacena k pares de vectores binarios, es decir, $(\mathbf{X}, \mathbf{Y}) = (\mathbf{x}^\xi, \mathbf{y}^\xi) \mid \xi = 1, \dots, k$ con $\mathbf{x}^\xi \in A^n$ y $\mathbf{y}^\xi \in A^m$.
- 3) El j -ésimo elemento de los vectores columna \mathbf{x}^ξ y \mathbf{y}^ξ , se representa por x_j^ξ y y_j^ξ , respectivamente.
- 4) La (i, j) -ésima entrada de las matrices $\mathbf{X} = (\mathbf{x}^1, \dots, \mathbf{x}^\xi)$ de dimensión $n \times k$ y $\mathbf{Y} = (\mathbf{y}^1, \dots, \mathbf{y}^\xi)$ de dimensión $m \times k$, está dada por x_i^j y y_i^j , respectivamente.
- 5) Si $\mathbf{X} = \mathbf{Y}$, es decir, $\mathbf{x}^\xi = \mathbf{y}^\xi \quad \forall \xi = 1, \dots, k$, entonces \mathbf{M} es una *memoria autoasociativa*, de lo contrario es una *memoria heteroasociativa*.
- 6) Se dice que existe una *recuperación perfecta*, cuando \mathbf{x}^ξ o una versión distorsionada $\tilde{\mathbf{x}}^\xi$, donde $\xi \in \{1, \dots, k\}$, es presentada a la memoria \mathbf{M} y sucede que la salida es igual a \mathbf{y}^ξ .

Finalmente, el problema fundamental del diseño de una memoria asociativa se divide en dos etapas:

- 1) *Fase de aprendizaje*: conjunto de reglas y operaciones empleadas para almacenar el par de matrices (\mathbf{X}, \mathbf{Y}) en una matriz \mathbf{M} .
- 2) *Fase de recuperación*: conjunto de reglas y operaciones empleadas para recuperar el patrón \mathbf{y}^ξ a partir del patrón de entrada \mathbf{x}^ξ o de una versión distorsionada $\tilde{\mathbf{x}}^\xi$.

III. NUEVAS MEMORIAS ASOCIATIVAS

A partir de esta sección, $A = \{0, 1\}$ y $C = \{01, 10, 11\} \subset B = \{00, 01, 10, 11\} \subset \mathbb{N}$ dos subconjuntos no negativos representados en forma binaria que tienen un buen orden con respecto a \leq .

TABLA I
OPERADORES BINARIOS Y EL OPERADOR PROYECCIÓN

w	w^\downarrow	x	y	$x \boxplus y$	$x \boxminus y$
00	0	0	0	10	00
01	1	0	1	01	11
10	0	1	0	11	01
11	1	1	1	10	10

(a) Operador \downarrow (b) Operadores \boxplus y \boxminus

TABLA II
PROPIEDADES DE LOS OPERADORES \boxplus Y \boxminus

Propiedad	operador \boxplus	operador \boxminus
1. Creciente	$x \leq y \leftrightarrow x \boxplus z \leq y \boxplus z$	$x \leq y \leftrightarrow z \boxminus x \leq z \boxminus y$
2. Distributiva sobre \vee	$(x \vee y) \boxplus z \leftrightarrow (x \boxplus z) \vee (y \boxplus z)$	$z \boxminus (x \vee y) \leftrightarrow (z \boxminus x) \vee (z \boxminus y)$
3. Distributiva sobre \wedge	$(x \wedge y) \boxplus z \leftrightarrow (x \boxplus z) \wedge (y \boxplus z)$	$z \boxminus (x \wedge y) \leftrightarrow (z \boxminus x) \wedge (z \boxminus y)$
4. Autoasociativa	$[(x \boxplus y) \boxplus z]^\downarrow = [x \boxplus (y \boxplus z)]^\downarrow$	$[(x \boxminus y) \boxminus z]^\downarrow = [x \boxminus (y \boxminus z)]^\downarrow$
5. Asociativa	$[(x \boxplus y) \boxplus z]^\downarrow = [x \boxplus (y \boxplus z)]^\downarrow$	$[(x \boxminus y) \boxminus z]^\downarrow = [x \boxminus (y \boxminus z)]^\downarrow$
6. Conmutativa	$(x \boxplus y)^\downarrow = (y \boxplus x)^\downarrow$	$(x \boxminus y)^\downarrow = (y \boxminus x)^\downarrow$
7. Complemento	$(x \boxplus \bar{x})^\downarrow = 1$	$(x \boxminus \bar{x})^\downarrow = 1$
8. Inversa	$[(x \boxplus y) \boxplus y]^\downarrow = x$	$[(x \boxminus y) \boxminus y]^\downarrow = x$
	$[(x \boxplus y) \boxplus x]^\downarrow = y$	$[(x \boxminus y) \boxminus x]^\downarrow = y$
	$[(x \boxplus y) \boxplus y]^\downarrow = x$	$[(x \boxminus y) \boxminus y]^\downarrow = x$
9. Autoinversa	$[(x \boxplus y) \boxplus x]^\downarrow = y$	$[(x \boxminus y) \boxminus x]^\downarrow = y$

Definición 1: Sean $A = \{0, 1\}$, $B = \{00, 01, 10, 11\}$ y $w \in B$; entonces, el operador *proyección* $\downarrow : B \rightarrow A$ es definido en la Tabla I (a).

Definición 2: Sean $A = \{0, 1\}$, $B = \{00, 01, 10, 11\}$, $C = \{01, 10, 11\}$ y $x, y \in A$; entonces, los operadores binarios $\boxplus : A \times A \rightarrow C$ y $\boxminus : A \times A \rightarrow B$ son definidos en la Tabla I (b).

Sean $x, y, z \in A = \{0, 1\}$ y \vee, \wedge las operaciones máximo y mínimo, respectivamente, entonces las operaciones de las Definiciones 1 y 2, exhiben las propiedades de la Tabla II.

La idea de las operaciones \boxplus y \boxminus es compensar la relación entre los patrones originales y los patrones a almacenar en la memoria, mediante la adición de un bit más significativo. Estas operaciones implican el enmascaramiento XOR del patrón original en el bit menos significativo, por lo que el este patrón puede reconstruirse a partir del patrón enmascarado con otra operación de enmascaramiento XOR. La operación proyección asegura que los resultados obtenidos por \boxplus y \boxminus sean binarios, además combinadas con estas dos operaciones binarias, satisfacen las propiedades de creciente por la izquierda o derecha, distributividad con respecto al máximo o mínimo, e inversas entre sí. Estas propiedades son suficientes para demostrar la mayoría de los resultados que caracterizan a las nuevas memorias. El coste computacional de las nuevas operaciones es bajo, debido a que su implementación en hardware presenta ventajas sobre la operaciones comunes de resta y suma: menor cantidad de compuertas digitales, menor tiempo de propagación de las señales eléctricas en los circuitos lógicos, y menor cantidad de celdas almacenamiento para la matriz de la memoria.

A. Memorias Heteroasociativas

Sustituyendo en (1) la operación \cdot por \boxplus y \boxminus , se introduce la terminología \boxplus (o \boxminus) para referirse a la operación matricial

\boxplus (o \boxminus) empleando la operación \boxplus , y $\bar{\boxplus}$ (o $\bar{\boxminus}$) para referirse a la operación matricial \boxplus (o \boxminus) empleando la operación \boxminus . Sean $A = \{0, 1\}$ y $\mathbf{x} = (x_1, \dots, x_n)^T \in A^n$ un patrón binario de entrada de dimensión n y $\mathbf{y} = (y_1, \dots, y_m)^T \in A^m$ un patrón binario de salida de dimensión m , entonces se definen las memorias asociativas ${}^+M$ y ${}^-M$ que almacenan la asociación de vectores (\mathbf{x}, \mathbf{y}) como

$${}^+M = (\mathbf{y} \times \mathbf{x}^T)^\downarrow = \begin{pmatrix} (y_1 \boxplus x_1)^\downarrow & \cdots & (y_1 \boxplus x_n)^\downarrow \\ \vdots & \ddots & \vdots \\ (y_m \boxplus x_1)^\downarrow & \cdots & (y_m \boxplus x_n)^\downarrow \end{pmatrix} \quad (2)$$

$${}^-M = (\mathbf{y} \times \mathbf{x}^T)^\downarrow = \begin{pmatrix} (y_1 \boxminus x_1)^\downarrow & \cdots & (y_1 \boxminus x_n)^\downarrow \\ \vdots & \ddots & \vdots \\ (y_m \boxminus x_1)^\downarrow & \cdots & (y_m \boxminus x_n)^\downarrow \end{pmatrix}$$

donde $(\mathbf{y} \times \mathbf{x}^T)^\downarrow = (\mathbf{y} \boxplus \mathbf{x}^T)^\downarrow = (\mathbf{y} \boxminus \mathbf{x}^T)^\downarrow$. Las entradas de ambas memorias son de un solo bit, a diferencia de otros modelos binarios que utilizan dos bits en su almacenamiento, por ejemplo [11], [19]. Estas memorias satisfacen las condiciones presentadas en (3).

$$({}^+M \bar{\boxplus} \mathbf{x})^\downarrow = \begin{pmatrix} \left[\bigwedge_{i=1}^n (y_1 \boxplus x_i)^\downarrow \boxminus x_i \right]^\downarrow \\ \vdots \\ \left[\bigwedge_{i=1}^n (y_m \boxplus x_i)^\downarrow \boxminus x_i \right]^\downarrow \end{pmatrix} = \mathbf{y} \quad (3)$$

$$({}^-M \bar{\boxminus} \mathbf{x})^\downarrow = \begin{pmatrix} \left[\bigwedge_{i=1}^n (y_1 \boxminus x_i)^\downarrow \boxplus x_i \right]^\downarrow \\ \vdots \\ \left[\bigwedge_{i=1}^n (y_m \boxminus x_i)^\downarrow \boxplus x_i \right]^\downarrow \end{pmatrix} = \mathbf{y}$$

A continuación se describen las fases de aprendizaje y recuperación de las memorias.

1) *Aprendizaje:* Sean (\mathbf{X}, \mathbf{Y}) las matrices que contienen los k pares de vectores binarios de entrada y salida, entonces las matrices ${}^+M$ y ${}^-M$ están dadas por

$${}^+M_{\mathbf{X}\mathbf{Y}} = \left[\bigvee_{\xi=1}^k \left(\mathbf{y}^\xi \boxplus (\mathbf{x}^\xi)^T \right) \right]^\downarrow \quad (4)$$

$${}^-M_{\mathbf{X}\mathbf{Y}} = \left[\bigvee_{\xi=1}^k \left(\mathbf{y}^\xi \boxminus (\mathbf{x}^\xi)^T \right) \right]^\downarrow$$

donde la (i, j) -ésima entrada está dada por

$$({}^+M_{\mathbf{X}\mathbf{Y}})_{ij} = {}^+ m_{ij} = \left[\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right]^\downarrow \quad (5)$$

$$({}^-M_{\mathbf{X}\mathbf{Y}})_{ij} = {}^- m_{ij} = \left[\bigvee_{\xi=1}^k (y_i^\xi \boxminus x_j^\xi) \right]^\downarrow$$

Ambas memorias se obtienen con el operador máximo, la diferencia entre ellas es la operación binaria que se emplea en la matriz, \boxplus para ${}^+M$ y \boxminus para ${}^-M$.

2) *Recuperación*: El resultado de salida de las memorias asociativas ${}^+M$ y ${}^-M$ está dado por

$$({}^+M_{XY} \bar{\boxminus} x^\gamma)^\downarrow \text{ y } ({}^-M_{XY} \bar{\boxplus} x^\gamma)^\downarrow \quad (6)$$

donde la i -ésima entrada del vector resultante es

$$\begin{aligned} ({}^+M_{XY} \bar{\boxminus} x^\gamma)^\downarrow_i &= \left[\bigwedge_{j=1}^n ({}^+m_{ij} \boxminus x_j^\gamma) \right]^\downarrow = \left[\bigwedge_{j=1}^n \left(\left(\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right) \boxminus x_j^\gamma \right) \right]^\downarrow \\ ({}^-M_{XY} \bar{\boxplus} x^\gamma)^\downarrow_i &= \left[\bigwedge_{j=1}^n ({}^-m_{ij} \boxplus x_j^\gamma) \right]^\downarrow = \left[\bigwedge_{j=1}^n \left(\left(\bigvee_{\xi=1}^k (y_i^\xi \boxminus x_j^\xi) \right) \boxplus x_j^\gamma \right) \right]^\downarrow \end{aligned} \quad (7)$$

El método para construir y operar las memorias propuestas es similar, la diferencia radica en invertir el orden de las operaciones \boxplus y \boxminus en las fases de aprendizaje y recuperación.

Definición 3: Una matriz C es considerada una $\bar{\boxminus}$ -memoria de recuperación perfecta para (X, Y) si y sólo si $(C \bar{\boxminus} X)^\downarrow = Y$, es decir, $(C \bar{\boxminus} x^\xi)^\downarrow = y^\xi \forall \xi = 1, \dots, k$. Similarmente, una matriz C es considerada una $\bar{\boxplus}$ -memoria de recuperación perfecta para (X, Y) si y sólo si $(C \bar{\boxplus} X)^\downarrow = Y$, es decir, $(C \bar{\boxplus} x^\xi)^\downarrow = y^\xi \forall \xi = 1, \dots, k$.

El desempeño y las condiciones de recuperación de las memorias esta caracterizado por los siguientes tres teoremas y el corolario.

Teorema 1: ${}^+M_{XY}$ es una $\bar{\boxminus}$ -memoria de recuperación perfecta para (X, Y) si y sólo si para cada índice fila $i = 1, \dots, m$ y cada $\xi \in \{1, \dots, k\}$ existe un índice columna $j_0 \in \{1, \dots, n\}$, el cual depende de ξ e i , tal que

$${}^+m_{ij_0} = (y_i^\xi \boxplus x_{j_0}^\xi)^\downarrow \forall \xi = 1, \dots, k \quad (8)$$

Similarmente, ${}^-M_{XY}$ es una $\bar{\boxplus}$ -memoria de recuperación perfecta para (X, Y) si y sólo si para cada índice fila $i = 1, \dots, m$ y cada $\xi \in \{1, \dots, k\}$ existe un índice columna $j_0 \in \{1, \dots, n\}$, el cual depende de ξ e i , tal que

$${}^-m_{ij_0} = (y_i^\xi \boxminus x_{j_0}^\xi)^\downarrow \forall \xi = 1, \dots, k \quad (9)$$

Teorema 2: ${}^+M_{XY}$ es una $\bar{\boxminus}$ -memoria de recuperación perfecta para (X, Y) si y sólo si para cada $\xi = 1, \dots, k$, cada fila de la matriz $[{}^+M_{XY}] - [y^\xi \bar{\boxplus} (x^\xi)^T]^\downarrow$ contiene un elemento cero. Similarmente, ${}^-M_{XY}$ es una $\bar{\boxplus}$ -memoria de recuperación perfecta para (X, Y) si y sólo si para cada $\xi = 1, \dots, k$, cada fila de la matriz $[y^\xi \bar{\boxminus} (x^\xi)^T]^\downarrow - [{}^-M_{XY}]$ contiene un elemento cero.

El Teorema 2 $({}^+M_{XY} \bar{\boxminus} x^\xi)^\downarrow = y^\xi$ o $({}^-M_{XY} \bar{\boxplus} x^\xi)^\downarrow = y^\xi \forall \xi = 1, \dots, k$ si y sólo si para cada índice ξ y cada índice fila i , existe un índice columna j (el cual esta en función de ξ e i) tal que ${}^+m_{ij} = (y_i^\xi \bar{\boxplus} (x^\xi)^T)^\downarrow_{ij}$ o ${}^-m_{ij} = (y_i^\xi \bar{\boxminus} (x^\xi)^T)^\downarrow_{ij}$.

Corolario 1: $({}^+M_{XY} \bar{\boxminus} X)^\downarrow = Y$ si y sólo si para cada índice fila $i = 1, \dots, m$ y cada $\gamma \in \{1, \dots, k\}$, existe un

índice columna $j \in \{1, \dots, n\}$ (el cual esta en función de i y γ) tal que

$$x_j^\gamma = \left[\left(\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right) \boxminus y_i^\gamma \right]^\downarrow \quad (10)$$

Similarmente, $({}^-M_{XY} \bar{\boxplus} X)^\downarrow = Y$ si y sólo si para cada índice fila $i = 1, \dots, m$ y cada $\gamma \in \{1, \dots, k\}$, existe un índice columna $j \in \{1, \dots, n\}$ (el cual esta en función de i y γ) tal que

$$x_j^\gamma = \left[\left(\bigvee_{\xi=1}^k (y_i^\xi \boxminus x_j^\xi) \right) \boxplus y_i^\gamma \right]^\downarrow \quad (11)$$

Las memorias ${}^+M_{XY}$ son robustas al ruido aditivo y las memorias ${}^-M_{XY}$ son robustas al ruido sustractivo. Lo anterior se establece en

Teorema 3: Sea \tilde{x}^γ una versión distorsionada de x^γ ; entonces, $({}^+M_{XY} \bar{\boxminus} \tilde{x}^\gamma)^\downarrow = y^\gamma$ si y sólo si

$$\tilde{x}_j^\gamma \geq x_j^\gamma \vee \bigvee_{i=1}^m \left[\bigvee_{\xi \neq \gamma} \left((y_i^\xi \boxplus x_j^\xi)^\downarrow \boxminus y_i^\gamma \right) \right]^\downarrow \forall j = 1, \dots, n \quad (12)$$

y para cada índice fila $i \in \{1, \dots, n\}$, existe un índice columna $j_i \in \{1, \dots, m\}$ tal que

$$\tilde{x}_{j_i}^\gamma = x_{j_i}^\gamma \vee \left[\bigvee_{\xi \neq \gamma} \left((y_i^\xi \boxplus x_{j_i}^\xi)^\downarrow \boxminus y_i^\gamma \right) \right]^\downarrow \quad (13)$$

Similarmente, sea \tilde{x}^γ una versión distorsionada de x^γ ; entonces, $({}^-M_{XY} \bar{\boxplus} \tilde{x}^\gamma)^\downarrow = y^\gamma$ si y sólo si

$$\tilde{x}_j^\gamma \leq x_j^\gamma \vee \bigvee_{i=1}^m \left[\bigvee_{\xi \neq \gamma} \left((y_i^\xi \boxminus x_j^\xi)^\downarrow \boxplus y_i^\gamma \right) \right]^\downarrow \forall j = 1, \dots, n \quad (14)$$

y para cada índice fila $i \in \{1, \dots, n\}$, existe un índice columna $j_i \in \{1, \dots, m\}$ tal que

$$\tilde{x}_{j_i}^\gamma = x_{j_i}^\gamma \vee \left[\bigvee_{\xi \neq \gamma} \left((y_i^\xi \boxminus x_{j_i}^\xi)^\downarrow \boxplus y_i^\gamma \right) \right]^\downarrow \quad (15)$$

B. Memorias Autoasociativas

Las memorias ${}^\pm M$ en ausencia de ruido y en modo autoasociativo, exhiben recuperación perfecta para cualquier cantidad de patrones, es decir, tienen capacidad de almacenamiento infinita.

Teorema 4: $({}^+M_{XX} \bar{\boxminus} X)^\downarrow = X$ y $({}^-M_{XX} \bar{\boxplus} X)^\downarrow = X$.

Las memorias no presentan problemas de convergencia, el resultado se obtiene en un solo paso.

Teorema 5: Si $({}^+M_{XX} \bar{\boxminus} z)^\downarrow = u$, entonces $({}^+M_{XX} \bar{\boxminus} u)^\downarrow = u$. Similarmente, si $({}^-M_{XX} \bar{\boxplus} z)^\downarrow = u$, entonces $({}^-M_{XX} \bar{\boxplus} u)^\downarrow = u$.

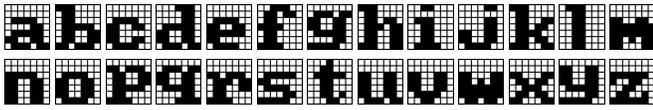


Fig. 1. Conjunto de imágenes de entrenamiento $\{p^\xi | \xi = 1, \dots, 26\}$. La numeración inicia de izquierda a derecha y de arriba abajo.

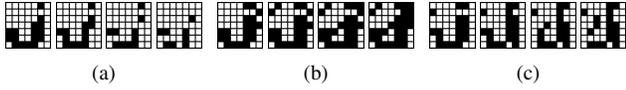


Fig. 2. Versiones de la letra "j" distorsionadas con ruido: (a) sustractivo, (b) aditivo y (c) mezclado. De izquierda a derecha en cada subfigura la cantidad de píxeles con ruido son: 2, 5, 8 y 10.

IV. RESULTADOS DE SIMULACIÓN POR COMPUTADORA

En las simulaciones se empleó el conjunto de 26 patrones binarios de 7×7 píxeles mostrado en la Fig. 1, donde los píxeles en color negro representan un valor 1 y los píxeles en blanco un valor 0. El código del programa fue realizado en lenguaje de programación C++ empleando una computadora con un procesador Intel® Core™ i5-7200U de arquitectura Kaby Lake (64 bits), frecuencia base de 2.5GHz y 8GB de memoria RAM instalada. Los patrones fueron distorsionados con ruido aditivo, sustractivo o mezclado, la Fig. 2 muestra un ejemplo de tres versiones distorsionadas de la letra "j".

El criterio utilizado para evaluar los resultados fue la distancia binaria gamma [22], la cual está dada por (16).

$$\gamma(\mathbf{x}, \tilde{\mathbf{x}}) = \left| 1 - \frac{2}{n} \sum_{i=1}^n |x_i - \tilde{x}_i| \right| \quad (16)$$

donde $\tilde{\mathbf{x}}$ es la versión distorsionada del vector \mathbf{x} . La distancia γ representa la similaridad entre imágenes y su categorización esta dada por: $0 \leq \gamma \leq 0.01$ imágenes distintas, $0.01 < \gamma \leq 0.2$ imágenes de mínima similaridad, $0.2 < \gamma \leq 0.4$ imágenes de baja similaridad, $0.4 < \gamma < 0.6$ imágenes de similaridad media, $0.6 \leq \gamma < 0.8$ imágenes de buena similaridad, $0.8 \leq \gamma < 1$ imágenes de alta similaridad, y $\gamma = 1$ imágenes idénticas. Si $\gamma = 1$, entonces la distancia indica una recuperación perfecta entre el vector obtenido y el vector deseado, es decir, $x_i = \tilde{x}_i$ para $i = 1, \dots, n$. Adicionalmente, se evaluó el rendimiento de las memorias con la *razón de recuperación perfecta*, la cual es una medida que muestra el número de patrones de entrenamiento perfectamente almacenados en la memoria asociativa.

Se realizaron tres experimentos para examinar el desempeño de las memorias asociativas propuestas con otros modelos, para ello se emplearon condiciones idénticas de versiones de entrada distorsionadas con un solo tipo de ruido, ya sea aditivo, sustractivo o mezclado. En los experimentos se utilizó el modo autoasociativo y los resultados son calculados por el promedio de 500 simulaciones. En los experimentos realizados con ruido aditivo y sustractivo, el desempeño de las memorias $\pm M$ fue comparado con las memorias de Hopfield, FRM, MAM, NMMAM, NR²FMAM y codificación max-min para la FMM, y en el caso del experimento efectuado con ruido mezclado la comparación fue con las MAM.

La Fig. 3, muestra los diagramas de flujo de los algoritmos de aprendizaje y recuperación utilizados para obtener las memorias $\pm M$. Con la modificación de algunos incisos en ambos algoritmos se obtuvieron las demás memorias, a saber, en el algoritmo de aprendizaje se modificaron las operaciones empleadas para el cálculo de las matrices parciales (inciso 2c) y las reglas para actualizar la matriz de la memoria (inciso 2d), además se eliminó el inciso 3, mientras que, en el algoritmo de recuperación se modificaron las operaciones y reglas para recuperar el vector de salida (inciso 2d). A continuación se describen los pasos de los procedimientos:

Procedimiento de aprendizaje (Fig. 3a):

- 1) Establecer la matriz inicial M con valores de entrada 0.
- 2) Para cada una de las 26 imágenes binarias de la Fig. 1, repetir lo siguiente:
 - a) Abrir la imagen $p^\xi(i, j)$.
 - b) Convertir la imagen $p^\xi(i, j)$ a un vector $\mathbf{x}^\xi = (x_1^\xi, \dots, x_{49}^\xi)$ definiendo $x_{7(i-1)+j}^\xi = 1$ si el pixel $p^\xi(i, j)$ es negro y $x_{7(i-1)+j}^\xi = 0$ si el pixel $p^\xi(i, j)$ es blanco.
 - c) Calcular la matriz M^ξ utilizando (4), es decir, $M^\xi = \mathbf{x}^\xi \bar{\boxplus} (\mathbf{x}^\xi)^T$ para obtener la memoria ${}^+M$ y $M^\xi = \mathbf{x}^\xi \bar{\boxminus} (\mathbf{x}^\xi)^T$ para la memoria ${}^-M$.
 - d) Calcular el máximo entre las matrices M y M^ξ .
- 3) Calcular la matriz de la memoria asociativa empleando la Definición 1, es decir, ${}^+M = M^\downarrow$ o ${}^-M = M^\downarrow$.
- 4) Guardar la matriz ${}^+M$ o ${}^-M$ en la memoria de la computadora.

Procedimiento de recuperación (Fig. 3b):

- 1) Cargar la matriz ${}^+M$ o ${}^-M$.
- 2) De 1 hasta 10 píxeles y para cada una de las 26 imágenes binarias de la Fig. 1, repetir lo siguiente:
 - a) Abrir la imagen $p^\xi(i, j)$.
 - b) Alterar la imagen $p^\xi(i, j)$ con un tipo de ruido. El número de píxeles distorsionados esta dado por la variable σ .
 - c) Convertir la imagen distorsionada $\tilde{p}^\xi(i, j)$ a un vector $\tilde{\mathbf{x}}^\xi = (\tilde{x}_1^\xi, \dots, \tilde{x}_{49}^\xi)$ definiendo $\tilde{x}_{7(i-1)+j}^\xi = 1$ si el pixel $\tilde{p}^\xi(i, j)$ es negro y $\tilde{x}_{7(i-1)+j}^\xi = 0$ si el pixel $\tilde{p}^\xi(i, j)$ es blanco.
 - d) Calcular el vector de salida \mathbf{y}^ξ empleando (6), es decir, $({}^+M \bar{\boxplus} \tilde{\mathbf{x}}^\xi)^\downarrow$ o $({}^+M \bar{\boxminus} \tilde{\mathbf{x}}^\xi)^\downarrow$.
 - e) Calcular la distancia binaria gamma γ entre los vectores \mathbf{y}^ξ y \mathbf{x}^ξ usando (16).
 - f) Convertir el vector encontrado $\mathbf{y}^\xi = (y_1^\xi, \dots, y_{49}^\xi)$ a una imagen de salida $q^\xi(i, j)$ de 7×7 píxeles, definiendo un color negro en $q^\xi(i, j)$ si $y_{7(i-1)+j}^\xi = 1$ y un color blanco en $q^\xi(i, j)$ si $y_{7(i-1)+j}^\xi = 0$.
 - g) Guardar la imagen obtenida q^ξ y su valor asociado γ . Estos resultados son para un valor de σ píxeles alterados con ruido.

La Fig. 4 muestra el desempeño de las memorias cuando los patrones de entrada son alterados con ruido aditivo y

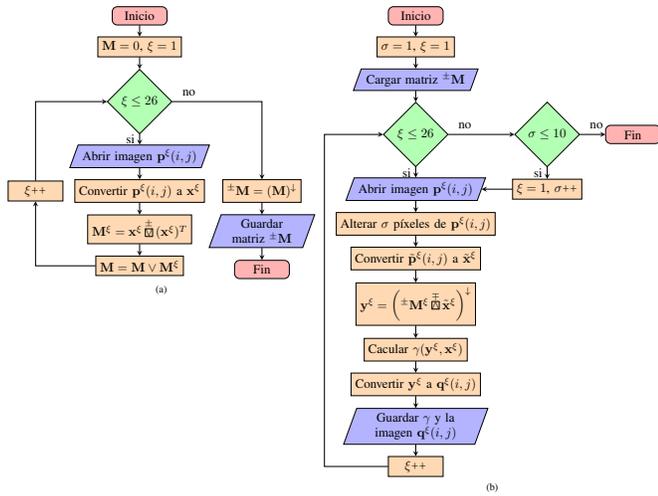


Fig. 3. Diagrama de flujo de la simulación: (a) fase aprendizaje y (b) fase de recuperación.

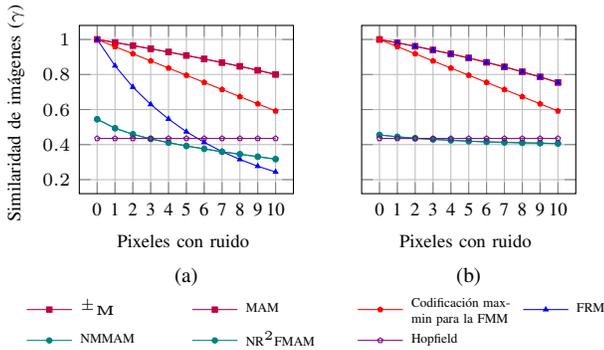


Fig. 4. Promedio de la distancia binaria gamma para los 26 patrones del alfabeto entrenados, cuando son distorsionados por: (a) ruido aditivo y (b) ruido sustractivo.

sustractivo (subfiguras a y b, respectivamente). Los experimentos exhibieron que en ausencia de ruido, las memorias $\pm M$, MAM, Codificación max-min para la FMM y la FRM, proporcionaron recuperación perfecta de los 26 patrones de la Fig. 1 ($\gamma = 1$). Cuando las 26 imágenes fueron distorsionadas con ruido aditivo o sustractivo, las memorias $\pm M$ y MAM, presentaron comportamientos similares y su tolerancia al ruido fue superior respecto a las otras memorias. En el caso de la memoria FRM, su tolerancia al ruido sustractivo es similar al de las memorias $\pm M$ y MAM min (Fig. 4a), sin embargo, ante la presencia de ruido aditivo, su desempeño es inferior (Fig. 4b). En general, las memorias $\pm M$, MAM y FRM (esta memoria sólo con ruido sustractivo), mostraron en promedio una recuperación de imágenes con alta similaridad cuando las 26 imágenes de entrada fueron distorsionadas con ruido aditivo o sustractivo, mientras que, en ausencia de ruido las imágenes fueron idénticas.

La Tabla III muestra la sensibilidad al ruido aditivo y sustractivo para las memorias asociativas. Cuando los píxeles de las 26 letras del alfabeto se alteraron con ruido aditivo, únicamente las memorias $\pm M$ y MAM max proporcionaron recuperación perfecta de patrones almacenados, en cambio

TABLA III
SENSIBILIDAD AL RUIDO CUANDO LAS 26 LETRAS DEL ALFABETO FUERON ALMACENADAS POR LAS MEMORIAS ASOCIATIVAS

Píxeles con ruido	Sensibilidad a ruido aditivo (Razón de recuperación perfecta [%])		Sensibilidad a ruido sustractivo (Razón de recuperación perfecta [%])	
	$\pm M$ y MAM max	Codificación max-min para la FMM y FRM	$\pm M$, MAM min y FRM	Codificación max-min para la FMM
0	100%	100%	100%	100%
1	57.8%	0%	53.5%	0%
2	32.9%	0%	33.2%	0%
3	18.8%	0%	23.5%	0%
4	10.6%	0%	17.0%	0%
5	5.8%	0%	12.9%	0%
6	3.1%	0%	9.9%	0%
7	1.6%	0%	7.8%	0%
8	0.9%	0%	6.2%	0%
9	0.5%	0%	4.8%	0%
10	0.3%	0%	3.8%	0%

cuando se alteraron con ruido sustractivo, las memorias $\pm M$, MAM min y FRM proporcionaron recuperación perfecta de patrones almacenados. La memoria Hopfield, NMMAM y NR²FMAM no se reportaron en los resultados de la tabla, debido a que no proporcionaron recuperación perfecta de algún patrón almacenado, es decir, su razón de recuperación perfecta fue de 0% para las 26 letras del alfabeto almacenadas.

En general, el desempeño y tolerancia al ruido aditivo o sustractivo que exhiben las memorias $\pm M$ y MAM son similares, sin embargo, en la implementación en hardware de las operaciones utilizadas por las memorias $\pm M$, presentan ventajas sobre las MAM:

- Número menor de compuertas digitales de 2 entradas empleadas en los circuitos lógicos que implementan las operaciones de las memorias, debido a que la operación \boxtimes usa 1 XOR, 1 OR y 1 NOT y la operación \boxminus usa 1 XOR, mientras que el medio restador emplea 1 XOR, 1 OR y 1 AND y el sumador completo 3 XOR, 2 AND y 1 OR.
- Menor tiempo de propagación de las señales eléctricas de los circuitos lógicos.
- Menor número de celdas de almacenamiento para el almacenamiento de la matriz de la memoria, debido a que cada entrada es de 1 bit, mientras que en las MAM es de 2 bits.

En el tercer experimento, se empleó la definición de *kernel* [11], es decir, se encontró una matriz $Z = (z^1, \dots, z^k)$ de dimensión $n \times k$ que satisface ${}^{-}M_{ZZ} \boxplus ({}^{+}M_{ZZ} \boxminus X) = {}^{-}M_{ZZ} \boxplus Z = X$. Los vectores kernels para las MAM se encontraron mediante el método convencional implementado por prueba y error, mientras que para las memorias $\pm M$ se modificó el algoritmo definido en [23]. Los vectores kernel de la Fig. 5, satisfacen las condiciones establecidas en [24] para el caso binario, esto es, $z^\gamma \ll x^\xi$, $z^\gamma \wedge x^\xi = \mathbf{0} \forall \gamma \neq \xi$, donde \ll (mucho menor que) hace énfasis al hecho de que z^ξ tiene más píxeles con valores cero y $\mathbf{0}$ es el vector con ceros. Se emplearon los procedimientos de la Fig. 3 para obtener las matrices M_{ZZ} y M_{ZX} .

La Fig. 6 muestra que las memorias $\pm M$ y MAM exhibieron en promedio una recuperación de imágenes de buena similaridad, sin embargo, el método propuesto muestra una curva con valores de distancia gamma mayores hasta .11 con respecto a

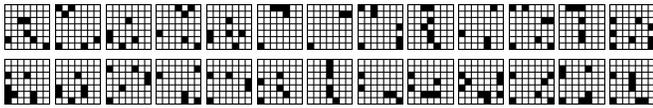
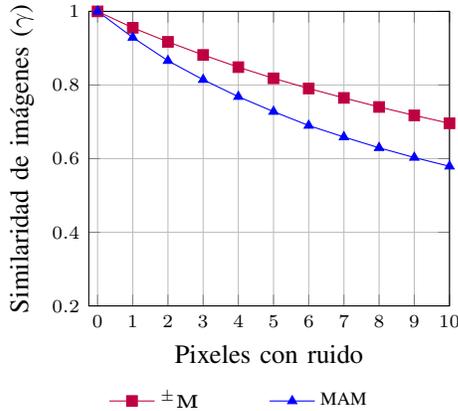
Fig. 5. Vectores kernels para las memorias $\pm M$.

Fig. 6. Promedio de la distancia binaria gamma para los 26 patrones del alfabeto entrenados, cuando son distorsionados por ruido mezclado

TABLA IV
SENSIBILIDAD AL RUIDO CUANDO LAS 26 LETRAS DEL ALFABETO
FUERON ALMACENADAS POR LAS MEMORIAS

Píxeles con ruido	$\pm M$	MAM
0	100%	100%
1	64.5%	38.4%
2	42.1%	16.6%
3	28.8%	7.6%
4	19.3%	3.9%
5	13.5%	2.2%
6	9.1%	1.0%
7	6.7%	0.7%
8	4.8%	0.4%
9	3.6%	0.3%
10	2.7%	0.2%

las MAM. La Tabla IV muestra que las memorias $\pm M$ exhiben mejor tolerancia al ruido mezclado que las MAM, debido a que sus porcentajes de recuperación perfecta asociados son mayores. En ausencia de ruido, ambas memorias, muestran un rendimiento del 100%.

Finalmente, los resultados sugieren que las memorias propuestas son una alternativa viable en la solución de diversos problemas, por ejemplo, en [25] se desarrolló un sistema de información geográfico que utiliza una red Hopfield para generar una propuesta de ruta que debe seguir un móvil dentro de una ciudad, por lo que la introducción de las memorias $\pm M$ en este tipo de sistemas, aportaría una reducción de tiempo para generar la ruta (procesamiento de los datos), lo cual implicaría un aumento en la velocidad de respuesta al móvil.

V. CONCLUSIONES

En este artículo se presentó un nuevo método para la obtención de memorias asociativas binarias, el cual se sustentó en dos nuevas operaciones binarias \boxplus y \boxminus , y la operación

unaria proyección \downarrow . La matriz de la memoria ${}^+M$ (o ${}^-M$) se calculó en la fase de aprendizaje encontrando el máximo de los resultados de la operación \boxplus (o \boxminus) entre los patrones de entrada y salida del conjunto de entrenamiento, mientras que la obtención de un patrón de salida se calculó en la fase de recuperación encontrando el mínimo de los resultados de la operación \boxminus (o \boxplus) entre el matriz de la memoria y el patrón presentado a la misma. Los resultados calculados en ambas fases, se les aplicó la operación proyección \downarrow para que las entradas de la matriz de la memoria y de los vectores salida fueran elementos binarios.

Se demostró que la memoria ${}^+M$ es robusta al ruido aditivo y la memoria ${}^-M$ es robusta al ruido sustractivo. El método presentado permite que ambas memorias funcionen en modo heteroasociativo o autoasociativo, en este último caso, se demostró que en ausencia de ruido, la capacidad de almacenamiento es infinita y la recuperación es perfecta para el conjunto de entrenamiento. Además, se mostró que los valores de entrada de la matriz $\pm M$ son de un bit, lo cual implica que las memorias propuestas utilizan un número menor de celdas binarias para su almacenamiento, a diferencia de las MAM donde los valores de entrada de la matriz son de dos bits.

Se verificó mediante los resultados obtenidos en los experimentos con ruido aditivo y sustractivo, que las memorias $\pm M$ y MAM exhibieron un desempeño similar y son superiores a los modelos de Hopfield, FRM, NMMAM, NR²FMAM y codificación max-min para FMM. En el experimento con ruido mezclado, se calculó la matriz Z de vectores kernel [11], [24], los vectores kernel de las memorias $\pm M$ se encontraron mediante el algoritmo propuesto por Hattori [23] y para las MAM a través del método convencional de prueba y error, los resultados exhibieron que el desempeño de las memorias $\pm M$ fue superior al de las MAM.

Debido a que la implementación en hardware de las operaciones definidas en este trabajo son de bajo coste computacional, en futuros trabajos se espera aplicar estas memorias en la solución de problemas del área de reconocimiento de patrones, particularmente en el diseño de sistemas embebidos de hardware especializado en tiempo real y algoritmos en paralelo que optimicen el tiempo de procesamiento consumido por el número de ciclos de reloj del CPU.

APÉNDICE A

Prueba del Teorema 1: Por la Definición 3, ${}^+M_{XY}$ es una $\bar{\boxminus}$ -memoria de recuperación perfecta para (X, Y) si y sólo si $({}^+M_{XY} \bar{\boxminus} x^\xi)_i^\downarrow = \left[\bigwedge_{j=1}^n ({}^+m_{ij} \boxminus x_j^\xi) \right]^\downarrow = y_i^\xi$ $\forall \xi = 1, \dots, k$ y $\forall i = 1, \dots, m$. Sean $i \in \{1, \dots, m\}$ y $\gamma \in \{1, \dots, k\}$ arbitrarios

$$({}^+M_{XY} \bar{\boxminus} x^\gamma)_i^\downarrow = \left[\bigwedge_{j=1}^n ({}^+m_{ij} \boxminus x_j^\gamma) \right]^\downarrow \quad (17)$$

si $j = j_0$, entonces se mantiene que

$$\left[\bigwedge_{j=1}^n (+m_{ij} \boxplus x_j^\gamma) \right]^\downarrow \leq (+m_{ij_0} \boxplus x_{j_0}^\gamma)^\downarrow \quad (18)$$

por hipótesis $+m_{ij_0} = (y_i^\gamma \boxplus x_{j_0}^\gamma)^\downarrow$, entonces

$$(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \mathbf{x}^\gamma)_i^\downarrow \leq ((y_i^\gamma \boxplus x_{j_0}^\gamma)^\downarrow \boxplus x_{j_0}^\gamma)^\downarrow = y_i^\gamma \quad (19)$$

por lo tanto $(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \mathbf{x}^\xi)^\downarrow \leq \mathbf{y}^\xi \forall \xi = 1, \dots, k$, debido a que i y γ se seleccionaron de manera arbitraria.

Demostrando lo contrario $(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \mathbf{x}^\xi)^\downarrow \geq \mathbf{y}^\xi$, sea $i \in \{1, \dots, m\}$ y $\gamma \in \{1, \dots, k\}$ arbitrarios, entonces

$$(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \mathbf{x}^\gamma)_i^\downarrow = \left[\bigwedge_{j=1}^n (+m_{ij} \boxplus x_j^\gamma) \right]^\downarrow \quad (20)$$

pero

$$+m_{ij} = \left[\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right]^\downarrow \quad (21)$$

lo cual implica

$$(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \mathbf{x}^\gamma)_i^\downarrow = \left[\bigwedge_{j=1}^n \left(\left[\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right]^\downarrow \boxplus x_j^\gamma \right) \right]^\downarrow \quad (22)$$

de (21) tenemos

$$\begin{aligned} & \left[\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right]^\downarrow \geq (y_i^\gamma \boxplus x_j^\gamma)^\downarrow \\ & \Leftrightarrow \left[\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right]^\downarrow \boxplus x_j^\gamma \geq (y_i^\gamma \boxplus x_j^\gamma)^\downarrow \boxplus x_j^\gamma \\ & \Leftrightarrow \bigwedge_{j=1}^n \left(\left[\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right]^\downarrow \boxplus x_j^\gamma \right) \geq \bigwedge_{j=1}^n ((y_i^\gamma \boxplus x_j^\gamma)^\downarrow \boxplus x_j^\gamma) \\ & \Leftrightarrow \left[\bigwedge_{j=1}^n \left(\left[\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right]^\downarrow \boxplus x_j^\gamma \right) \right]^\downarrow \geq \left[\bigwedge_{j=1}^n ((y_i^\gamma \boxplus x_j^\gamma)^\downarrow \boxplus x_j^\gamma) \right]^\downarrow = y_i^\gamma \end{aligned} \quad (23)$$

de esta última desigualdad, se mantiene que

$$(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \mathbf{x}^\gamma)_i^\downarrow \geq y_i^\gamma \quad (24)$$

por lo tanto $(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \mathbf{x}^\xi)^\downarrow \geq \mathbf{y}^\xi \forall \xi = 1, \dots, k$. De manera similar se demuestra para $-\mathbf{M}_{\mathbf{X}\mathbf{Y}}$. ■

Prueba del Teorema 2: Del Teorema 1, si $+\mathbf{M}_{\mathbf{X}\mathbf{Y}}$ es una $\bar{\boxminus}$ -memoria de recuperación perfecta para (\mathbf{X}, \mathbf{Y}) , entonces para cada índice columna $i = 1, \dots, m$ y cada $\gamma \in \{1, \dots, k\}$, existen índices columnas $j_0 \in \{1, \dots, n\}$ tal que $+m_{ij_0} = (y_i^\gamma \boxplus x_{j_0}^\gamma)^\downarrow$. De (2) y (5) se tiene que

$$\begin{aligned} (y_i^\gamma \boxplus x_{j_0}^\gamma)^\downarrow &= (\mathbf{y}^\gamma \times (\mathbf{x}^\gamma)^T)_{ij_0}^\downarrow \\ +m_{ij_0} &= (+\mathbf{M}_{\mathbf{X}\mathbf{Y}})_{ij_0} \end{aligned} \quad (25)$$

Entonces $+m_{ij_0} = (y_i^\gamma \boxplus x_{j_0}^\gamma)^\downarrow$ es equivalente a

$$\begin{aligned} (+\mathbf{M}_{\mathbf{X}\mathbf{Y}})_{ij_0} &= (\mathbf{y}^\gamma \times (\mathbf{x}^\gamma)^T)_{ij_0}^\downarrow \\ (+\mathbf{M}_{\mathbf{X}\mathbf{Y}})_{ij_0} - (\mathbf{y}^\gamma \times (\mathbf{x}^\gamma)^T)_{ij_0}^\downarrow &= 0 \\ (+\mathbf{M}_{\mathbf{X}\mathbf{Y}} - (\mathbf{y}^\gamma \times (\mathbf{x}^\gamma)^T)^\downarrow)_{ij_0} &= 0 \end{aligned} \quad (26)$$

Esto muestra que cada índice fila de la matriz $+\mathbf{M}_{\mathbf{X}\mathbf{Y}} - (\mathbf{y}^\gamma \times (\mathbf{x}^\gamma)^T)^\downarrow$ contiene una entrada con el elemento cero. De manera similar se demuestra para $-\mathbf{M}_{\mathbf{X}\mathbf{Y}}$. ■

Prueba del Corolario 1: Del Teorema 2, $(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \mathbf{x}^\xi)^\downarrow = \mathbf{y}^\xi \forall \xi = 1, \dots, k$ si y sólo si para cada $\gamma \in \{1, \dots, k\}$ y cada índice fila i , existe un índice columna $j \in \{1, \dots, n\}$ (el cual esta en función de γ y j) tal que

$$+m_{ij} = (\mathbf{y}^\gamma \times (\mathbf{x}^\gamma)^T)_{ij}^\downarrow = (y_i^\gamma \boxplus x_j^\gamma)^\downarrow \quad (27)$$

De (5) y (27) se tiene que

$$\begin{aligned} (y_i^\gamma \boxplus x_j^\gamma)^\downarrow &= \left(\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right)^\downarrow \\ \left[(y_i^\gamma \boxplus x_j^\gamma)^\downarrow \boxplus y_i^\gamma \right]^\downarrow &= \left[\left(\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right)^\downarrow \boxplus y_i^\gamma \right]^\downarrow \\ x_j^\gamma &= \left[\left(\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right)^\downarrow \boxplus y_i^\gamma \right]^\downarrow \end{aligned} \quad (28)$$

De manera similar se demuestra para $-\mathbf{M}_{\mathbf{X}\mathbf{Y}}$. ■

Prueba del Teorema 3: Suponga que $\tilde{\mathbf{x}}^\gamma$ es una versión distorsionada de \mathbf{x}^γ y que para $\gamma = 1, \dots, k$, $(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \tilde{\mathbf{x}}^\gamma)^\downarrow = \mathbf{y}^\gamma$. Entonces,

$$\begin{aligned} y_i^\gamma &= \left(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \tilde{\mathbf{x}}^\gamma \right)_i^\downarrow = \left[\bigwedge_{l=1}^n (+m_{il} \boxplus \tilde{x}_l^\gamma) \right]^\downarrow \quad \forall i = 1, \dots, m \\ &\leq (+m_{ij} \boxplus \tilde{x}_j^\gamma)^\downarrow \quad \forall i = 1, \dots, m \text{ y } \forall j = 1, \dots, n \end{aligned} \quad (29)$$

por lo que

$$\begin{aligned} y_i^\gamma &\leq (+m_{ij} \boxplus \tilde{x}_j^\gamma)^\downarrow \quad \forall i = 1, \dots, m \text{ y } \forall j = 1, \dots, n \\ (+m_{ij} \boxplus \tilde{x}_j^\gamma)^\downarrow &\leq (+m_{ij} \boxplus \tilde{x}_j^\gamma)^\downarrow \quad \forall i = 1, \dots, m \text{ y } \forall j = 1, \dots, n \\ [(+m_{ij} \boxplus \tilde{x}_j^\gamma)^\downarrow]^\downarrow &\leq [(+m_{ij} \boxplus \tilde{x}_j^\gamma)^\downarrow]^\downarrow \quad \forall i = 1, \dots, m \text{ y } \forall j = 1, \dots, n \\ \tilde{x}_j^\gamma &\geq (+m_{ij} \boxplus \tilde{x}_j^\gamma)^\downarrow \quad \forall i = 1, \dots, m \text{ y } \forall j = 1, \dots, n \\ \Leftrightarrow \tilde{x}_j^\gamma &\geq \left[\left(\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right)^\downarrow \boxplus y_i^\gamma \right]^\downarrow \quad \forall i = 1, \dots, m \text{ y } \forall j = 1, \dots, n \\ \Leftrightarrow \tilde{x}_j^\gamma &\geq \bigvee_{i=1}^m \left(\left[\left(\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right)^\downarrow \boxplus y_i^\gamma \right]^\downarrow \right) \quad \forall j = 1, \dots, n \\ \Leftrightarrow \tilde{x}_j^\gamma &\geq \bigvee_{i=1}^m \left(\left[\left(\bigvee_{\xi \neq \gamma}^k (y_i^\xi \boxplus x_j^\xi) \right)^\downarrow \boxplus y_i^\gamma \right]^\downarrow \vee [(y_i^\gamma \boxplus x_j^\gamma)^\downarrow \boxplus y_i^\gamma]^\downarrow \right) \quad \forall j = 1, \dots, n \\ \Leftrightarrow \tilde{x}_j^\gamma &\geq \bigvee_{i=1}^m \left(\left[\left(\bigvee_{\xi \neq \gamma}^k (y_i^\xi \boxplus x_j^\xi) \right)^\downarrow \boxplus y_i^\gamma \right]^\downarrow \vee x_j^\gamma \right) \quad \forall j = 1, \dots, n \\ \Leftrightarrow \tilde{x}_j^\gamma &\geq x_j^\gamma \vee \bigvee_{i=1}^m \left[\left(\bigvee_{\xi \neq \gamma}^k (y_i^\xi \boxplus x_j^\xi) \right)^\downarrow \boxplus y_i^\gamma \right]^\downarrow \quad \forall j = 1, \dots, n \end{aligned} \quad (30)$$

Esto demuestra a (12). Continuando con esta desigualdad

$$\tilde{x}_j^\gamma \geq x_j^\gamma \vee \left[\left(\bigvee_{\xi \neq \gamma}^k (y_i^\xi \boxplus x_j^\xi) \right)^\downarrow \boxplus y_i^\gamma \right]^\downarrow \quad \forall j = 1, \dots, n \text{ y } \forall i = 1, \dots, m \quad (31)$$

Suponga que (31) no contiene una desigualdad para $i = 1, \dots, m$; es decir, existe un índice fila $i \in \{1, \dots, m\}$ tal que

$$\tilde{x}_j^\gamma > x_j^\gamma \vee \left[\left(\bigvee_{\xi \neq \gamma} (y_i^\xi \boxplus x_j^\xi) \right) \boxminus y_i^\gamma \right] \quad \forall j = 1, \dots, n \quad (32)$$

entonces

$$\begin{aligned} (+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \tilde{\mathbf{x}}^\gamma)_i^\dagger &= \left[\bigwedge_{j=1}^n (+m_{ij} \boxminus \tilde{x}_j^\gamma) \right]^\dagger \\ &> \left[\bigwedge_{j=1}^n \left(+m_{ij} \boxminus \left(x_j^\gamma \vee \left[\left(\bigvee_{\xi \neq \gamma} (y_i^\xi \boxplus x_j^\xi) \right) \boxminus y_i^\gamma \right] \right) \right) \right]^\dagger \\ &= \left[\bigwedge_{j=1}^n \left(+m_{ij} \boxminus \left[\left(\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_j^\xi) \right) \boxminus y_i^\gamma \right] \right) \right]^\dagger \\ &= \left[\bigwedge_{j=1}^n (+m_{ij} \boxminus [+m_{ij} \boxminus y_i^\gamma]^\dagger) \right]^\dagger \\ &= \left[\bigwedge_{j=1}^n ([+m_{ij} \boxminus (+m_{ij})]^\dagger \boxminus y_i^\gamma) \right]^\dagger \\ &= \left[\bigwedge_{j=1}^n (0 \boxminus y_i^\gamma) \right]^\dagger = [0 \boxminus y_i^\gamma]^\dagger \\ &= y_i^\gamma \end{aligned} \quad (33)$$

Por lo tanto, $(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \tilde{\mathbf{x}}^\gamma)_i^\dagger > y_i^\gamma$, lo cual contradice la hipótesis $(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \tilde{\mathbf{x}}^\gamma)^\dagger = \mathbf{y}^\gamma$. Lo anterior implica que para cada índice fila i , debe existir un índice columna j_i que satisfaga (13). Suponga que

$$\tilde{x}_j^\gamma \geq x_j^\gamma \vee \bigvee_{i=1}^m \left[\left(\bigvee_{\xi \neq \gamma} (y_i^\xi \boxplus x_j^\xi) \right) \boxminus y_i^\gamma \right] \quad \forall j = 1, \dots, n \quad (34)$$

Por la primera parte de la demostración, esta desigualdad es verdadera si y sólo si

$$\tilde{x}_j^\gamma \geq (+m_{ij} \boxminus y_i^\gamma)^\dagger \quad \forall i = 1, \dots, m \text{ y } \forall j = 1, \dots, n \quad (35)$$

o equivalentemente, si y sólo si

$$\begin{aligned} (+m_{ij} \boxminus \tilde{x}_j^\gamma)^\dagger &\geq y_i^\gamma \quad \forall i = 1, \dots, m \text{ y } \forall j = 1, \dots, n \\ &\Leftrightarrow \left[\bigwedge_{j=1}^n (+m_{ij} \boxminus \tilde{x}_j^\gamma) \right]^\dagger \geq y_i^\gamma \quad \forall i = 1, \dots, m \quad (36) \\ &\Leftrightarrow (+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \tilde{\mathbf{x}}^\gamma)_i^\dagger \geq y_i^\gamma \quad \forall i = 1, \dots, m \end{aligned}$$

lo cual implica que $(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \tilde{\mathbf{x}}^\gamma)^\dagger \geq \mathbf{y}^\gamma \quad \forall \gamma = 1, \dots, k$. Ahora se demostrará que $(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \tilde{\mathbf{x}}^\gamma)^\dagger \leq \mathbf{y}^\gamma \quad \forall \gamma = 1, \dots, k$ para concluir que $(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \tilde{\mathbf{x}}^\gamma)^\dagger = \mathbf{y}^\gamma \quad \forall \gamma = 1, \dots, k$. Sean $\gamma \in \{1, \dots, k\}$ e $i \in \{1, \dots, m\}$ arbitrarios; entonces,

$$\begin{aligned} (+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \tilde{\mathbf{x}}^\gamma)_i^\dagger &= \left[\bigwedge_{j=1}^n (+m_{ij} \boxminus \tilde{x}_j^\gamma) \right]^\dagger \\ &\leq [+m_{ij} \boxminus \tilde{x}_{j_i}^\gamma]^\dagger \\ &= \left[+m_{ij} \boxminus \left(x_{j_i}^\gamma \vee \left[\left(\bigvee_{\xi \neq \gamma} (y_i^\xi \boxplus x_{j_i}^\xi) \right) \boxminus y_i^\gamma \right] \right) \right]^\dagger \\ &= \left[+m_{ij} \boxminus \left(\left(\bigvee_{\xi=1}^k (y_i^\xi \boxplus x_{j_i}^\xi) \right) \boxminus y_i^\gamma \right) \right]^\dagger \\ &= [+m_{ij} \boxminus (+m_{ij} \boxminus y_i^\gamma)^\dagger]^\dagger \\ &= y_i^\gamma \end{aligned} \quad (37)$$

Esto muestra que $(+\mathbf{M}_{\mathbf{X}\mathbf{Y}} \bar{\boxminus} \tilde{\mathbf{x}}^\gamma)^\dagger \leq \mathbf{y}^\gamma$. De manera similar se demuestra para $-\mathbf{M}_{\mathbf{X}\mathbf{Y}}$. ■

Prueba del Teorema 4: Debido a que $+m_{ii} = (\mathbf{x}^\xi \bar{\boxminus} (\mathbf{x}^\xi)^T)_{ii}^\dagger = (x_i^\xi \boxminus x_i^\xi)^\dagger = (10)^\dagger = 0$ para cada $i = 1, \dots, n$ y $\forall \xi = 1, \dots, k$, entonces cada fila de $(+\mathbf{M}_{\mathbf{X}\mathbf{X}} - (\mathbf{x}^\xi \bar{\boxminus} (\mathbf{x}^\xi)^T)^\dagger)^\dagger$ contiene un elemento cero; por lo tanto, del Teorema 2, $+\mathbf{M}_{\mathbf{X}\mathbf{X}}$ es una $\bar{\boxminus}$ -memoria de recuperación perfecta para (\mathbf{X}, \mathbf{X}) . De manera similar se demuestra para $-\mathbf{M}_{\mathbf{X}\mathbf{X}}$. ■

Prueba del Teorema 5: Sea $(+\mathbf{M}_{\mathbf{X}\mathbf{X}} \bar{\boxminus} \mathbf{z})^\dagger = \mathbf{u}$. Debido a que $+m_{ii} = 0$ para cada $i = 1, \dots, n$, entonces

$$\begin{aligned} (+\mathbf{M}_{\mathbf{X}\mathbf{X}} \bar{\boxminus} \mathbf{u})_i^\dagger &= \left[\bigwedge_{j=1}^n (+m_{ij} \boxminus u_j) \right]^\dagger \\ &\leq [+m_{ii} \boxminus u_j]^\dagger = [0 \boxminus u_j]^\dagger = u_j \end{aligned} \quad (38)$$

por lo que

$$\mathbf{u} \geq (+\mathbf{M}_{\mathbf{X}\mathbf{X}} \bar{\boxminus} \mathbf{u})^\dagger \quad (39)$$

Sean $i, j, l \in \{1, \dots, n\}$, entonces

$$+m_{il} = \left(\bigvee_{\xi=1}^k (x_i^\xi \boxplus x_l^\xi) \right)^\dagger \geq (x_i^\gamma \boxplus x_l^\gamma)^\dagger \quad \text{para } \gamma = 1, \dots, k \quad (40)$$

$$+m_{lj} = \left(\bigvee_{\xi=1}^k (x_l^\xi \boxplus x_j^\xi) \right)^\dagger \geq (x_l^\gamma \boxplus x_j^\gamma)^\dagger \quad \text{para } \gamma = 1, \dots, k \quad (41)$$

lo cual implica que

$$\begin{aligned} [+m_{il} \boxminus (+m_{lj})]^\dagger &\geq [(x_i^\gamma \boxplus x_l^\gamma)^\dagger \boxminus (x_l^\gamma \boxplus x_j^\gamma)^\dagger]^\dagger \quad \text{para } \gamma = 1, \dots, k \\ &= [x_i^\gamma \boxminus (x_l^\gamma \boxminus (x_l^\gamma \boxplus x_j^\gamma)^\dagger)^\dagger]^\dagger \quad \text{para } \gamma = 1, \dots, k \\ &= (x_i^\gamma \boxminus x_j^\gamma)^\dagger \quad \text{para } \gamma = 1, \dots, k \end{aligned} \quad (42)$$

por lo que

$$[+m_{il} \boxminus (+m_{lj})]^\dagger \geq \left(\bigvee_{\xi=1}^k (x_i^\xi \boxplus x_j^\xi) \right)^\dagger = +m_{ij} \quad (43)$$

De esta última desigualdad se tiene que para $i = 1, \dots, n$

$$u_i = \left[\bigwedge_{j=1}^n (+m_{ij} \boxplus z_j) \right]^\downarrow \quad (44)$$

$$\leq \left[\bigwedge_{j=1}^n \left([+m_{il} \boxplus (+m_{lj})]^\downarrow \boxplus z_j \right) \right]^\downarrow \quad \forall l = 1, \dots, n$$

por consiguiente

$$u_i \leq \left[\bigwedge_{j=1}^n \left((+m_{il} \boxplus (+m_{lj} \boxplus z_j))^\downarrow \right) \right]^\downarrow \quad \forall l = 1, \dots, n$$

$$= \left[+m_{il} \boxplus \left(\bigwedge_{j=1}^n (+m_{lj} \boxplus z_j) \right) \right]^\downarrow \quad \forall l = 1, \dots, n$$

$$= \left[\bigwedge_{l=1}^n \left(+m_{il} \boxplus \left(\bigwedge_{j=1}^n (+m_{lj} \boxplus z_j) \right) \right) \right]^\downarrow \quad (45)$$

$$= \left[\bigwedge_{l=1}^n (+m_{il} \boxplus u_l) \right]^\downarrow$$

$$= (+\mathbf{M}_{\mathbf{X}\mathbf{X}} \bar{\boxplus} \mathbf{u})^\downarrow_i$$

Lo cual demuestra que

$$\mathbf{u} \leq (+\mathbf{M}_{\mathbf{X}\mathbf{X}} \bar{\boxplus} \mathbf{u})^\downarrow \quad (46)$$

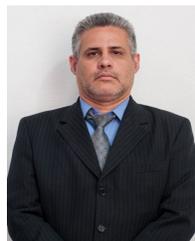
(39) y (46) implican que $\mathbf{u} = (\mathbf{M}_{\mathbf{X}\mathbf{X}} \boxplus \mathbf{u})^\downarrow$. De manera similar se demuestra para $\bar{\boxplus} \mathbf{M}_{\mathbf{X}\mathbf{X}}$. ■

REFERENCES

- [1] M. H. Hassoun, *Associative neural memories: theory and implementation*. New York, United States: Oxford University Press, Inc., 1993.
- [2] W. K. Taylor, "Electrical simulation of some nervous system functional activities," in *Proc. Information Theory: Third London Symposium*, (Butterworths, London), p. 314–328, Sept. 1955.
- [3] K. Steinbuch, "Die lernmatrix," *Kybernetik*, vol. 1, pp. 36–45, Jan. 1961.
- [4] T. Kohonen, "Correlation matrix memories," *IEEE Trans. Comput.*, vol. C-21, pp. 353–359, Apr. 1972.
- [5] K. Nakano, "Associatron—a model of associative memory," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-2, pp. 380–388, July 1972.
- [6] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 79, pp. 2554–2558, Apr. 1982.
- [7] R. J. McEliece, E. C. Posner, E. R. Rodemich, and S. S. Venkatesh, "The capacity of the hopfield associative memory," *IEEE Trans. Inf. Theory*, vol. 33, pp. 461–482, July 1987.
- [8] B. Kosko, "Fuzzy associative memories," in *Proc. of the 2nd Joint Technology Workshop on Neural Networks and Fuzzy Logic*, (California, United States), pp. 3–58, Feb. 1990.
- [9] F.-L. Chung and T. Lee, "Towards a high capacity fuzzy associative memory model," in *Proc. of 1994 IEEE International Conference on Neural Networks (ICNN'94)*, (Florida, United States), pp. 1595–1599, July 1994.
- [10] P. Xiao, F. Yang, and Y. Yu, "Max-min encoding learning algorithm for fuzzy max-multiplication associative memory networks," in *Proc. 1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, (Orlando, United States), pp. 3674–3679, Oct. 1997.
- [11] G. X. Ritter, P. Sussner, and J. L. Díaz de León, "Morphological associative memories," *IEEE Trans. Neural Netw.*, vol. 2, pp. 281–293, Mar. 1998.
- [12] N. Feng, X. Cao, S. Li, L. Ao, and S. Wang, "A new method of morphological associative memories," in *Proc. Emerging Intelligent Computing Technology and Applications. With Aspects of Artificial Intelligence. ICIC 2009*, (Ulsan, South Korea), pp. 407–416, Sept. 2009.
- [13] T. D. Bui, T. H. Nong, and T. K. Dang, "Improving learning rule for fuzzy associative memory with combination of content and association," *Neurocomputing*, vol. 149, pp. 59–64, Feb. 2015.
- [14] N. Feng and Y. Yao, "No rounding reverse fuzzy morphological associative memories," *Neural Network World*, vol. 26, no. 6, pp. 571–587, 2016.
- [15] P. Sussner and M. E. Valle, "Implicative fuzzy associative memories," *IEEE Trans. Fuzzy Syst.*, vol. 14, pp. 793–807, Nov. 2006.
- [16] A. Ferreyra-Ramírez, E. Rodríguez-Martínez, C. Avilés-Cruz, and F. López-Saca, "Image retrieval system based on a binary auto-encoder and a convolutional neural network," *IEEE Latin America Transactions*, vol. 100, pp. 1–8, Sept. 2020.
- [17] N. Feng, Y. Qiu, F. Wang, and Y. Sun, *A Unified Framework of Morphological Associative Memories*, vol. 344 of *Lecture Notes in Control and Information Sciences*, ch. Intelligent Control and Automation: International Conference on Intelligent Computing, ICIC 2006 Kunming, China, August 16–19, 2006, pp. 1–11. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006.
- [18] B. Sun and N. Feng, "Exploring the new application of morphological neural networks," in *Intelligent Computing Theories and Application*, (Liverpool, United Kingdom), pp. 113–120, Aug. 2017.
- [19] C. Yáñez-Márquez, I. López-Yáñez, M. Aldape-Pérez, O. Camacho-Nieto, A. J. Argüelles-Cruz, and Y. Villuendas-Rey, "Theoretical foundations for the alpha-beta associative memories: 10 years of derived extensions, models, and applications," *Neural Processing Letters*, vol. 48, p. 811–847, Oct. 2018.
- [20] N. Feng and B. Sun, "On simulating one-trial learning using morphological neural networks," *Cognitive Systems Research*, vol. 53, pp. 61–70, 2019.
- [21] G. Urcid and G. X. Ritter, *Noise Masking for Pattern Recall Using a Single Lattice Matrix Associative Memory*, pp. 81–100. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.
- [22] A. A. Mustafa, "Probabilistic binary similarity distance for quick binary image matching," *IET Image Processing*, vol. 12, no. 10, pp. 1844–1856, 2018.
- [23] M. Hattori, A. Fukui, and H. Ito, "A fast method of constructing kernel patterns for morphological associative memory," in *Proc. 9th International Conference on Neural Information Processing, 2002. ICONIP'02*, (Singapore, Singapore), pp. 1058–1063, Nov. 2002.
- [24] P. Sussner, "Observations on morphological associative memories and the kernel method," *Neurocomputing*, vol. 31, pp. 167–183, Mar. 2000.
- [25] J. Y. Montiel, R. Romero, and J. L. López, "Generation of trip routes on the basis of a geographic information system," *IEEE Latin America Transactions*, vol. 2, pp. 196–200, Sept. 2004.



Arturo Gamino Carranza received the B. S. Degree in Electronic Engineering from the Instituto Tecnológico de Veracruz México in 2002, the M. Sc. Degree in Automatic Control from the Centro de Investigación y Estudios Avanzados del IPN México (CINVESTAV-IPN) in 2005. Currently, he is a professor at the Tecnológico Nacional de México. His research interest includes associative memories, mathematical morphology and educational innovation.



Juan Luis Díaz de León received the B. S. Degree in Electronic Engineering from the Instituto Tecnológico de Veracruz México in 1991, the M. Sc. and the Ph. D. degrees both in Electrical Engineering from the Centro de Investigación y Estudios Avanzados del IPN México (CINVESTAV-IPN), in 1993 and 1996. He is currently a research professor at the Centro de Investigación en Computación de the IPN (CIC-IPN). His research interest includes associative memories, mathematical morphology, topology, image processing and control of electromechanical

systems using visual data.