# A Fast and Robust Max- C Projection Fuzzy Autoassociative Memory with Application for Face Recognition

Alex Santana dos Santos and
Marcos Eduardo Valle*

Department of Applied Mathematics
Institute of Mathematics, Statistics, and
Scientific Computing
University of Campinas - Brazil

5 October 2017

# Fuzzy Autoassociative Memories

The human brain ability to recall information by association instigated researches on associative memory models.

In an associative memory, an stimulus is associated to a response in such a way that the presentation of the stimulus gives rise to the response.

We speak of an autoassociative memory if the stimulus and the response coincide.

An autoassociative memory can be used, for instance, to recognize a face occluded by sunglasses or scarf.

A fuzzy associative memory (FAM) is a memory designed for the storage and recall of fuzzy sets!

# Outline of this Talk

# Fuzzy Logic

The symbols "$\vee$" and "$\wedge$" represent the supremum (maximum) and infimum (minimum) operations.

## Definition (Fuzzy conjunction)

An increasing mapping $C : [0,1] \times [0,1] \longrightarrow [0,1]$ is a fuzzy conjunction if $C(0,0) = C(0,1) = C(1,0) = 0$ and $C(1,1) = 1$.

## Example

- $C_M(x,y) = x \wedge y,$ (minimum)
- $C_G(x,y) = \begin{cases} 0, & x = 0, \\ y, & \text{otherwise.} \end{cases}$ (Gaines' conjunction)

# Fuzzy Implication

## Definition (Fuzzy Implication)

A mapping $I : [0,1] \times [0,1] \longrightarrow [0,1]$ decreasing in the first argument and increasing in the second argument is a fuzzy implication if $I(0,0) = I(0,1) = I(1,1) = 1$ and $I(1,0) = 0$.

## Example

- $I_M(x,y) = \begin{cases} 1, & x \leq y, \\ y, & x > y. \end{cases}$ (Gödel implication)
- $I_G(x,y) = \begin{cases} 1, & x \leq y, \\ 0, & x > y, \end{cases}$ (Gaines implication)

# Adjunction

Fuzzy conjunctions and implications are not inverse operations but they can be linked by an adjunction.

## Definition (Adjunction)

A fuzzy implication $I$ and a fuzzy conjunction $C$ form an adjunction if

$$C(x, y) \leq z \quad \Longleftrightarrow \quad x \leq I(y, z), \quad \forall x, y, z \in [0, 1].$$

## Example

The pairs $(I_M, C_M)$ and $(I_G, C_G)$ are adjunctions.

# Fuzzy Sets

A fuzzy set $\mathbf{x}$ on $U = \{u_1, u_2, \ldots, u_n\}$ can be identified with a vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \in [0,1]^n$, where $x_j = \mathbf{x}(u_j)$ denotes the degree of membership of $u_j$ in the fuzzy set $\mathbf{x}$.

Like linear combinations, $\mathbf{z} \in [0,1]^n$ is a max-$C$ combinations of $\mathcal{A} = \left\{\mathbf{a}^1, \ldots, \mathbf{a}^k\right\} \subseteq [0,1]^n$ if

$$\mathbf{z} = \bigvee_{\xi=1}^{k} C(\lambda_\xi, \mathbf{a}^\xi) \iff z_i = \bigvee_{\xi=1}^{k} C(\lambda_\xi, a_i^\xi), \ \forall i = 1, \ldots, n,$$

where $\lambda_\xi \in [0,1]$ for all $\xi = 1, \ldots, k$.

The family of all max-$C$ combinations of $\mathcal{A} = \left\{\mathbf{a}^1, \ldots, \mathbf{a}^k\right\}$ is

$$\mathcal{C}(\mathcal{A}) = \left\{ \mathbf{z} = \bigvee_{\xi=1}^{k} C(\lambda_\xi, \mathbf{a}^\xi) : \lambda_\xi \in [0,1] \right\}.$$

# Fuzzy Inclusion Measures

A fuzzy inclusion measure $Inc_{\mathcal{F}}(\mathbf{a}, \mathbf{b})$ yields the degree of inclusion of a fuzzy set **a** in a fuzzy set **b**.

## Example (Zadeh fuzzy inclusion measure)

$$Inc_{\mathcal{Z}}(\mathbf{a}, \mathbf{b}) = \begin{cases} 1, & a_j \leq b_j, \forall j = 1, \ldots, n, \\ 0, & \text{otherwise.} \end{cases}$$

## Example (Inf-$I$ inclusion measure)

$$Inc_{\mathcal{F}}(\mathbf{a}, \mathbf{b}) = \bigwedge_{j=1}^{n} I(a_j, b_j).$$

Zadeh fuzzy inclusion measure is obtained if $I \equiv I_G$.

# Fuzzy Similarity Measure

A fuzzy similarity measure yields the degree of similarity between two fuzzy sets $\mathbf{a}, \mathbf{b}$.

## Example

$$S_H(\mathbf{a}, \mathbf{b}) = 1 - \frac{1}{n} \sum_{i=1}^{N} |a_i - b_i| \, .$$

# Max-$C$ Projection Autoassociative Fuzzy Memories

Given a set $\mathcal{A} = \{\mathbf{a}^1, \ldots, \mathbf{a}^k\} \subseteq [0,1]^n$, called fundamental memory set, an autoassociative fuzzy memory (AFM) $\mathcal{M}$ should satisfy

$$\mathcal{M}(\mathbf{a}^\xi) = \mathbf{a}^\xi, \quad \forall \xi \in \mathcal{K} = \{1, 2, \ldots, k\}.$$

The memory $\mathcal{M}$ is also expected to exhibit some noise tolerance:

$$\mathcal{M}(\mathbf{x}) = \mathbf{a}^\xi, \quad \text{for a noisy version } \mathbf{x} \text{ of } \mathbf{a}^\xi.$$

A max-$C$ projection autoassociative fuzzy memory (max-$C$ PAFM) projects the input $\mathbf{x}$ on the max-$C$ combinations of $\mathcal{A} = \{\mathbf{a}^1, \ldots, \mathbf{a}^k\}$:

$$\mathcal{V}(\mathbf{x}) = \bigvee \{\mathbf{z} \in \mathcal{C}(\mathcal{A}) : \mathbf{z} \leq \mathbf{x}\}.$$

The output $\mathcal{V}(\mathbf{x})$ is the largest max-$C$ combination of the fundamental memories less than or equal to the input $\mathbf{x}$.

# Noise Tolerance and Convergence

**Theorem**

*A max-C PAFM satisfies* $\mathcal{V}(\mathbf{x}) \leq \mathbf{x}$ *and* $\mathcal{V}(\mathcal{V}(\mathbf{x})) = \mathcal{V}(\mathbf{x})$ *for all* $\mathbf{x} \in [0,1]^n$.

**Consequences:**

- A max-$C$ PAFM converges in a single iteration if employed with feedback.
- A max-C PAFM exhibits tolerance only with respect to dilative noise:
  - a distorted version $\mathbf{x}$ of $\mathbf{a}^\xi$ has undergone a dilative change if $\mathbf{x} \geq \mathbf{a}^\xi$.
  - $\mathbf{x}$ has undergone an erosive change $\mathbf{x} \leq \mathbf{a}^\xi$.
- A max-$C$ PAFM is extremely sensitive to either erosive or mixed (dilative and erosive) noise; $\mathbf{a}^\xi$ cannot be retrieved if $\mathbf{x} \geq \mathbf{a}^\xi$!

# Storage Capacity and Implementation

## Theorem (Optimal Absolute Storage Capacity)

*A max-C PAFMs satisfy* $\mathcal{V}(\mathbf{a}^\xi) = \mathbf{a}^\xi, \forall \xi \in \mathcal{K}$, *if the fuzzy conjunction C has a left identity.*

## Theorem (Implementation)

*Let* $(I, C)$ *be an adjunction. The output of the max-C PAFM satisfies*

$$\mathcal{V}(\mathbf{x}) = \bigvee_{\xi=1}^{k} C(\lambda_\xi, \mathbf{a}^\xi) \text{ where } \lambda_\xi = \bigwedge_{j=1}^{n} I(a_j^\xi, x_j), \forall \xi \in \mathcal{K}.$$

Note that $\lambda_\xi$ is the degree of inclusion of $\mathbf{a}^\xi$ in $\mathbf{x}$:

$$\lambda_\xi = Inc_{\mathcal{F}}(\mathbf{a}^\xi, \mathbf{x}), \quad \forall \xi \in \mathcal{K}.$$

## Example

Let $(I_M, C_M)$ and consider

$$\mathcal{A} = \left\{ \mathbf{a}^1 = \begin{bmatrix} 0.4 \\ 0.3 \\ 0.7 \\ 0.2 \end{bmatrix}, \mathbf{a}^2 = \begin{bmatrix} 0.1 \\ 0.7 \\ 0.5 \\ 0.8 \end{bmatrix}, \mathbf{a}^3 = \begin{bmatrix} 0.8 \\ 0.5 \\ 0.4 \\ 0.2 \end{bmatrix} \right\}.$$

Given the input

$$\mathbf{x}_d = \begin{bmatrix} 0.4 \\ 0.3 \\ 0.8 \\ 0.7 \end{bmatrix},$$

we obtain

$$\lambda_1 = 1.0, \quad \lambda_2 = 0.3, \quad \text{and} \quad \lambda_3 = 0.3.$$

## Example

The output of the max-$C_M$ PAFM is

$$\begin{aligned}
\mathcal{V}_{\mathcal{M}}(\mathbf{x}_d) &= C_M(\lambda_1, \mathbf{a}^1) \vee C_M(\lambda_2, \mathbf{a}^2) \vee C_M(\lambda_3, \mathbf{a}^3) \\
&= \left(1.0 \wedge \begin{bmatrix} 0.4 \\ 0.3 \\ 0.7 \\ 0.2 \end{bmatrix}\right) \vee \left(0.3 \wedge \begin{bmatrix} 0.1 \\ 0.7 \\ 0.5 \\ 0.8 \end{bmatrix}\right) \vee \left(0.3 \wedge \begin{bmatrix} 0.8 \\ 0.5 \\ 0.4 \\ 0.2 \end{bmatrix}\right) \\
&= \begin{bmatrix} 0.4 \\ 0.3 \\ 0.7 \\ 0.2 \end{bmatrix} \vee \begin{bmatrix} 0.1 \\ 0.3 \\ 0.3 \\ 0.3 \end{bmatrix} \vee \begin{bmatrix} 0.3 \\ 0.3 \\ 0.3 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 0.4 \\ 0.3 \\ 0.7 \\ 0.3 \end{bmatrix} \neq \mathbf{a}^1.
\end{aligned}$$

# Zadeh max-$C$ PAFM

We can improve the noise tolerance of a max-$C$ PAFM by defining

$$\lambda_\xi = Inc_{\mathcal{Z}}(\mathbf{a}^\xi, \mathbf{x}) = \begin{cases} 1, & a_j^\xi \leq x_j,\ \forall j = 1, \ldots, n, \\ 0, & \text{otherwise.} \end{cases}$$

and

$$\mathcal{V}_{\mathcal{Z}}(\mathbf{x}) = \bigvee_{\xi=1}^{k} C_G(\lambda_\xi, \mathbf{a}^\xi).$$

Equivalently, we have

$$\mathcal{V}_{\mathcal{Z}}(\mathbf{x}) = \bigvee_{\xi \in \mathcal{I}} \mathbf{a}^\xi, \qquad \text{where} \qquad \mathcal{I} = \{\xi : a_j^\xi \leq x_j, \forall j = 1, \ldots, n\}.$$

No arithmetic operation is performed! We only perform comparisons!

## Example

Consider

$$\mathcal{A} = \left\{ \mathbf{a}^1 = \begin{bmatrix} 0.4 \\ 0.3 \\ 0.7 \\ 0.2 \end{bmatrix}, \mathbf{a}^2 = \begin{bmatrix} 0.1 \\ 0.7 \\ 0.5 \\ 0.8 \end{bmatrix}, \mathbf{a}^3 = \begin{bmatrix} 0.8 \\ 0.5 \\ 0.4 \\ 0.2 \end{bmatrix} \right\}.$$

Given the input

$$\mathbf{x}_d = \begin{bmatrix} 0.4 & 0.3 & 0.8 & 0.7 \end{bmatrix}^T,$$

we obtain

$$\mathcal{I} = \{\xi : a_j^\xi \leq x_j, \forall j = 1, \ldots, n\} = \{1\}.$$

The output of the Zadeh max-$C$ PAFM is

$$\mathcal{V}_{\mathcal{Z}}(\mathbf{x}_d) = \bigvee_{\xi \in \mathcal{I}} \mathbf{a}^\xi = \mathbf{a}^1.$$

# Noise Tolerance

**Theorem**

*The identity $\mathcal{V}_{\mathcal{Z}}(\mathbf{x}) = \mathbf{a}^{\eta}$ holds true if there exists an unique $\eta \in \mathcal{K}$ such that $\mathbf{a}^{\eta} \leq \mathbf{x}$.*

Consequence:

- The Zadeh max-$C$ PAFM is extremely robust to the dilative noise but it is extremely sensitive to erosive noise!

The noise tolerance of a max-$C$ PAFM can be significantly improved by masking the noise contained in a corrupted input.

# Noise Masking Strategy

In some sense, noise masking aims to remove the erosive noise from the input **x**.

1. Select an index $\eta$ such that

$$S(\mathbf{x}, \mathbf{a}^{\eta}) = \bigvee_{\xi=1}^{k} \left\{ S(\mathbf{x}, \mathbf{a}^{\xi}) \right\},$$

   where $S$ is a fuzzy similarity measure.
2. Define the novel memory

$$\mathcal{V}^{M}(\mathbf{x}) = \mathcal{V}(\mathbf{x} \vee \mathbf{a}^{\eta}).$$

If we use the fuzzy similarity measure $S_H$, $\mathcal{V}^{M}_{\mathcal{Z}}$ performs $\mathcal{O}(nk)$ operations during the retrieval phase.

# Computational Experiments

Face recognition problem using the AR database:

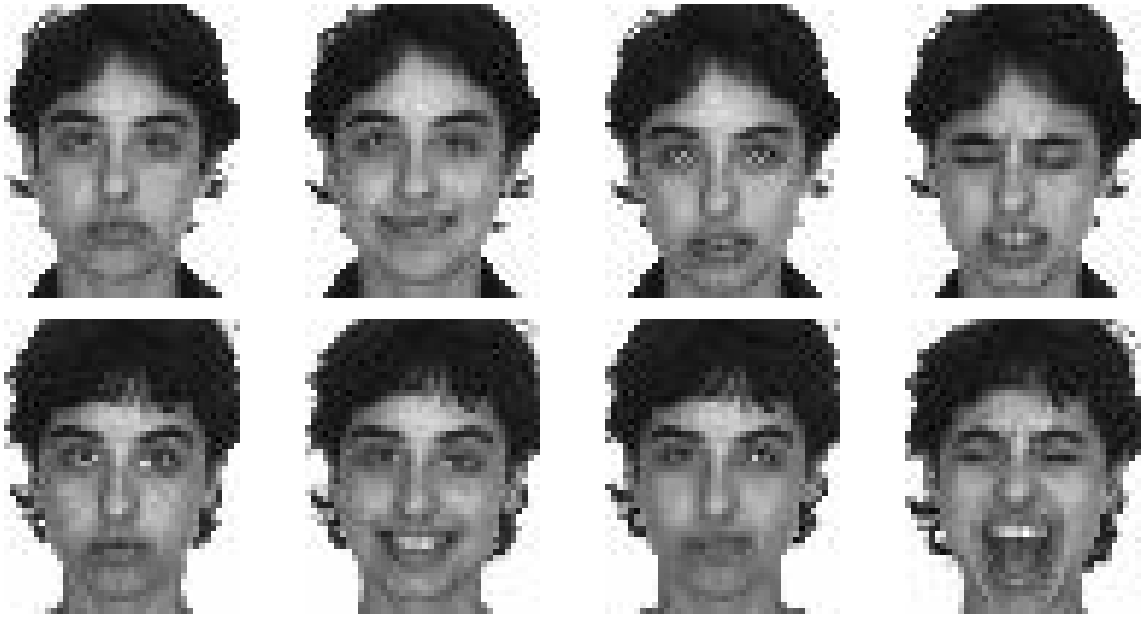
- Gray-scale face images of size $50 \times 40$ from 120 individuals.
- Eight frontal face images used for training.
- Two testing scenarios:
  a) Sunglasses plus illumination – 4 images.
  b) Scarf plus illumination – 4 images.

# Computational Experiments

Face recognition problem using the AR database:

- Gray-scale face images of size $50 \times 40$ from 120 individuals.
- Eight frontal face images used for training.
- Two testing scenarios:
  a) Sunglasses plus illumination – 4 images.
  b) Scarf plus illumination – 4 images.

**Training images from one individual:**

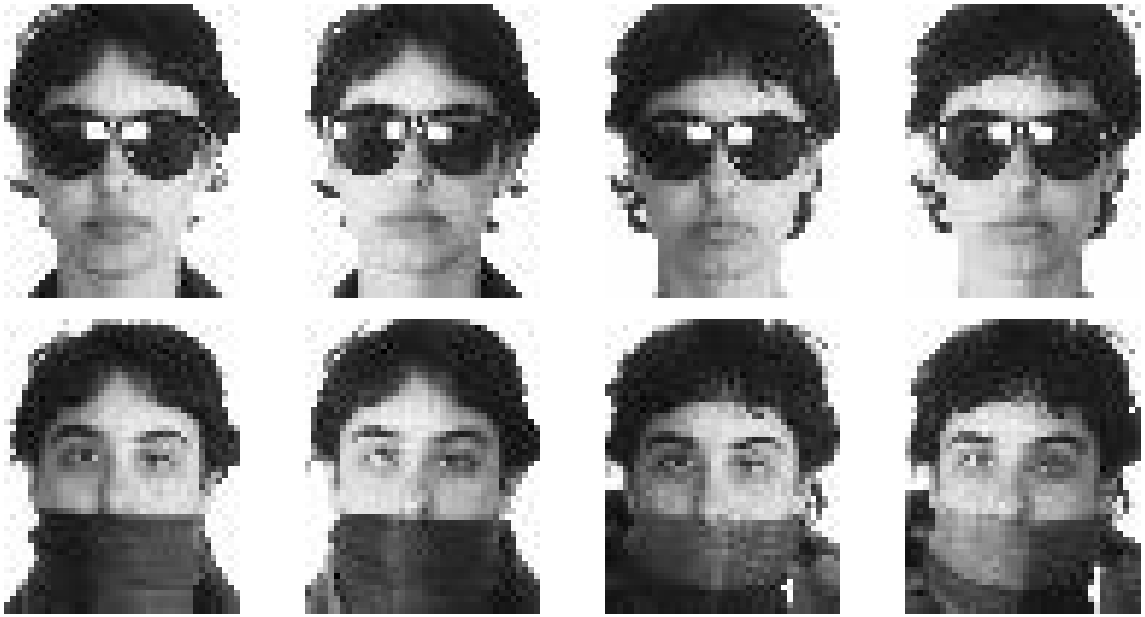


# Computational Experiments

Face recognition problem using the AR database:

- Gray-scale face images of size $50 \times 40$ from 120 individuals.
- Eight frontal face images used for training.
- Two testing scenarios:
  a) Sunglasses plus illumination – 4 images.
  b) Scarf plus illumination – 4 images.

**Testing images from one individual:**



# Max-$C$ PAFM Classifier

- We identify a face image of an individual $i$ with $\mathbf{a}^{\xi,i} \in [0,1]^{2000}$.
- We formed the fundamental memory sets $\mathcal{A}^i = \{\mathbf{a}^{1,i}, \ldots, \mathbf{a}^{8,i}\}$, for $i = 1, \ldots, c = 120$.

Independent max-$C$ PAFM model for each individual:

$$\mathcal{M}^i \quad (= \mathcal{V}^i_{\mathcal{M}} \text{ or } \mathcal{V}^i_{\mathcal{Z}}).$$

An unknown face image $\mathbf{x} \in [0,1]^n$ is assigned to $\eta$ such that

$$S_H\big(\mathbf{x}, \mathcal{M}^\eta(\mathbf{x})\big) \geq S_H\big(\mathbf{x}, \mathcal{M}^i(\mathbf{x})\big), \quad \forall i = 1, \ldots, c.$$

In words, $\mathbf{x}$ belongs to an individual such that the recalled vector is the most similar to the input!

| Classifier | Sunglasses | Scarf | Complexity |
|---|---|---|---|
| LRC | 90.2 | 30.4 | $\mathcal{O}(pqT(k+1)+(pq)^3)$ |
| SRC | 95.6 | 54.8 | $\mathcal{O}(T^2(T+pq)^{1.3})$ |
| NMR | 94.6 | 70.4 | $\mathcal{O}((pq)^{1.5}+pqT)$ |
| SN$L_2$R2 | 95.7 | 70.2 | $\mathcal{O}(pq^2+pqT+T^2)$ |
| SN$L_1$R1 | 96.1 | 71.2 | $\mathcal{O}(pq^2+pqT+T^2)$ |
| DN$L_2$R2 | 95.8 | 70.4 | $\mathcal{O}(pq^2+pqT+T^2)$ |
| DN$L_1$R1 | **96.7** | **72.3** | $\mathcal{O}(pq^2+pqT+T^2)$ |
| $\mathcal{V}^M_{\mathcal{M}}$ | 92.3 | 46.5 | $\mathcal{O}(pqT+T)$ |
| $\mathcal{V}^M_{\mathcal{Z}}$ | 92.7 | 47.3 | $\mathcal{O}(pqT+T)$ |
| $\mathcal{V}^M_{\mathcal{Z}}$ (partitioned) | 95.8 | **80.8** | $\mathcal{O}(pqT+\beta T)$ |

- $p$ and $q$ denote the height and width of the face images images,
- $c$ is the number of individuals,
- $k$ is the number of training images per individual,
- $T = ck$, and $\beta$ denotes the number of blocks in which the face images are partitioned.

# Concluding Remarks

We introduced the Zadeh max-$C$ projection autoassociative fuzzy memory (max-$C$ PFAM):

- Optimal absolute storage capacity.
- Excellent tolerance to dilative noise.
- Uses only comparisons (no operations)!

We also proposed the noise making strategy based on a fuzzy similarity measure to improve the tolerance with respect to noise.

Preliminary experiments suggested that the novel memory may be competitive with other classifiers by providing fairly recognition rates with low computational cost.

Thank you!

# Have an idea?

IEEE WCCI 2018
Rio de Janeiro, Brazil

IEEE WORLD CONGRESS ON COMPUTATIONAL INTELLIGENCE

8-13 July 2018, Rio de Janeiro, Brazil

Submit a paper to 2018 WCCI
(FUZZ-IEEE or IJCNN)!