ASSOCIATIVE MEMORIES NETWORK FOR FACE RECOGNITION AND OBJECT RECOGNITION

Roberto A. Vazquez ¹, Humberto Sossa ²

¹) Center for Computing Research (CIC-IPN), Av. Juan de Dios Batiz s/n Col. Nueva Industrial Vallejo, CP. 07738 Mexico City, Mexico, ravem@ipn.mx
²) Center for Computing Research (CIC-IPN), Av. Juan de Dios Batiz s/n Colonia Nueva Industrial Vallejo CP. 07738 Mexico City, Mexico, hsossa@cic.ipn.mx

Abstract: An associative memory (AM) is a special kind of neural network that allows associating an output pattern with an input pattern. Some problems require associating several output patterns with a unique pattern. Classical associative and neural models cannot solve this simple task and less if these patterns are complex images, for example faces. In this paper a network of AMs to recall a collection of patterns is proposed. The accuracy of the proposal is tested with two benchmarks. One is composed by 20 objects and the other is composed by 20 images of 15 different people faces. First the all, the benchmarks are split into several collections and then this collections are used to train the network of AMs. During training an image of a collection is associated with the rest of the images belonging to the same collection. Once trained the network we expected to recover a collection of images by using as an input pattern any image belonging to the collection.

Keywords: Associative memories, face recognition, object recognition.

1. INTRODUCTION

An associative memory AM is a special kind of neural network that allows recalling one output pattern given an input pattern as a key that might be altered by some kind of noise (additive, subtractive or mixed). Several models of AMs are described in [1], [2], [3], [4], [5], [6], [7] and [8]. Most common application of these models is as a filter. In general, these models have severe constraints which limit their applicability in complex problems such as object and face recognition. In particular, models described in [4], [5] and [6] cannot handle with mixed noise. Associative model presented in [7] and [8] is robust to mixed noise.

An association between input pattern \( x \) and output pattern \( y \) is denoted as \( (x^k, y^k) \), where \( k \) is the corresponding association. AM \( W \) is represented by a matrix whose component \( w_{ij} \) can be seen as the synapse of the neural network. Operator \( W \) is generated from a finite a priori set of know associations, known as the fundamental set of association and is represented as:\n
\[
\{ (x^k, y^k) \}_{k=1, \ldots, p}
\]

where \( p \) is the number of associations. If \( x^k = y^k \forall k = 1, \ldots, p \) then \( W \) is auto-associative, otherwise it is hetero-associative.

A distorted version of a pattern \( x \) to be restored will be denoted as \( \tilde{x} \). If an AM \( W \) is fed with a distorted version of \( x^k \) and the output obtained is exactly \( y^k \), we say that recalling is perfect.

These models only associate an input pattern with a unique output pattern which in some problems could be seen as a limitation. Suppose for example that instead of recall a face of a person associated to an input pattern you want to recall different faces of the person, probably took in different stages of his life, using the same input pattern, or any pattern related to the face. In order to solve these kind problems, we will need more than one associative memory. We will need a network of associative memories because an associative memory only recalls an output pattern and not a collection of patterns.

In this paper we present how a network of AMs can be used to recall nor just one pattern but several of them given an input pattern. Further more we shown how the adopted associative model is also useful in complex problems such as object and face recognition. In this proposal an association between input pattern \( x \) and a collection of output pattern \( Y \) is denoted as, \( \{ (x^k, Y^k) \}_{k=1, \ldots, p} \) where \( p \) is the number of association, \( Y^k = \{ y^l, \ldots, y^r \} \) is a
collection of output patterns and \( r \) is the number of patterns belonging to collection \( Y \).

The remaining of the paper is organized as follows. In sections 2 and 3 the associative model used in this research is described. In section 4 the proposed network of AMs is presented. In section 5 the experimental results obtained with the proposal are given. In section 6, finally, the conclusions and several directions for further research in this direction are presented.

### 2. DYNAMIC ASSOCIATIVE MODEL

The brain is not a huge fixed neural network, as had been previously thought, but a dynamic, changing neural network that adapts continuously to meet the demands of communication and computational needs [9]. This fact suggests that some connections of the brain could change in response to some input stimuli.

Humans, in general, do not have problems to recognize patterns even if these are altered by noise. Several parts of the brain interact together in the process of learning and recalling a pattern. For example, when we read a word the information enters the eye and the word is transformed into electrical impulses. Then electrical signals are passed through the brain to the visual cortex, where information about space, orientation, form and color is analyzed. After that, specific information about the patterns passes on the other areas of the cortex that integrate visual and auditory information. From here information passes through the arcuate fasciculus, a path that connects a large network of interacting brain areas; paths of this pathway connect language areas with other areas involving in cognition, association and meaning, for details see [10] and [11].

Based upon the above example we have defined in our model several interacting areas, one per association we would like the memory to learn. Also we have integrated the capability to adjust synapses in response to an input stimulus.

As we could appreciate from the previous example, before an input pattern is learned or processed by the brain, it is hypothesized that it is transformed and codified by the brain. In our model, this process is simulated using the following procedure recently introduced in [7]:

**Procedure 1.** Transform the fundamental set of associations into codified patterns and de-codifying patterns:

Input: FS Fundamental set of associations:

1. Make \( d = \text{const} \) and make \([x, y]) = (x', y')\)

2. For the remaining couples do {
    For \( k = 2 \) to \( p \)
    For \( i = 1 \) to \( n \)

\[
\mathbf{\bar{x}}_i^k = \mathbf{\bar{x}}_i^{k-1} + d:
\mathbf{\bar{z}}_i^k = \mathbf{\bar{x}}_i^k - x_i^k;
\mathbf{\bar{y}}_i^k = \mathbf{\bar{y}}_i^{k-1} + d:
\mathbf{\bar{f}}_i^k = \mathbf{\bar{y}}_i^k - y_i^k
\]

}\}

Output: Set of codified and de-codifying patterns.

This procedure allows computing *codified patterns* from input and output patterns denoted by \( \mathbf{\bar{x}} \) and \( \mathbf{\bar{y}} \) respectively; \( \mathbf{\bar{z}} \) and \( \mathbf{\bar{f}} \) are *de-codifying patterns*. Codified and de-codifying patterns are allocated in different interacting areas and \( d \) defines of much these areas are separated. On the other hand, \( d \) determines the noise supported by our model. In addition a simplified version of \( x_i^k \) denoted by \( s_k \) is obtained as:

\[
s_k = s(x_i^k) = \text{mid} x_i^k
\]

where \( \text{mid} \) operator is defined as \( \text{mid} x = x_{(n+1)/2} \).

When the brain is stimulated by an input pattern, some regions of the brain (interacting areas) are stimulated and synapses belonging to those regions are modified.

In our model, we call these regions *active regions* and could be estimated as follows:

\[
ar = r(x) = \arg \left( \min_{i=1}^n |s(x) - s_i| \right)
\]

Once computed the *codified patterns*, the *de-codifying patterns* and \( s_k \) we can build the associative memory.

Let \([x, y])_k \in \text{R}^n, \bar{x}^k \in \text{R}^n, \bar{y}^k \in \text{R}^n \) a fundamental set of associations (codified patterns). Synapses of associative memory \( W \) are defined as:

\[
w_{ij} = \bar{y}_i - \bar{y}_j
\]

After computed the *codified patterns*, the *de-codifying patterns*, the reader can easily corroborate that any association can be used to compute the synapses of \( W \) without modifying the results. In short, building of the associative memory can be performed in three stages as:

1. Transform the fundamental set of association into codified and de-codifying patterns by means of previously described Procedure 1.
2. Compute simplified versions of input patterns by using equation 1.
3. Build \( W \) in terms of codified patterns by using equation 3.

As we had already mentioned, synapses could change in response to an input stimulus; but which synapses should be modified? For example, a head injury might cause a brain lesion killing hundred of
neurons; this entails some synapses to reconnect with others neurons. This reconnection or modification of the synapses might cause that information allocated on brain will be preserved or will be lost, the reader could find more details concerning to this topic in [12] and [13].

This fact suggests there are synapses that can be drastically modified and they do not alter the behavior of the associative memory. In the contrary, there are synapses that only can be slightly modified to do not alter the behavior of the associative memory; we call this set of synapses the kernel of the associative memory and it is denoted by $K_w$.

In the model we can find two types of synapses: synapses that can be modified and do not alter the behavior of the associative memory; and synapses belonging to the kernel of the associative memory. These last synapses play an important role in recalling patterns altered by some kind of noise.

Let $K_w \in R^m$ the kernel of an associative memory $W$. A component of vector $K_w$ is defined as:

$$kw_j = \text{mid}(w_j), j = 1,\ldots,m$$

(4)

According to the original idea of our proposal, synapses that belong to $K_w$ are modified as a response to an input stimulus. Input patterns stimulate some active regions, interact with these regions and then, according to those interactions, the corresponding synapses are modified. Synapses belonging to $K_w$ are modified according to the stimulus generated by the input pattern. This adjusting factor is denoted by $\Delta w$ and can be computed as:

$$\Delta w = \Delta(x) = s(\xi) - s(x)$$

(5)

where $ar$ is the index of the active region.

Finally, synapses belonging to $K_w$ are modified as:

$$K_w = K_w \oplus (\Delta w - \Delta w_{old})$$

(6)

where operator $\oplus$ is defined as $x \oplus e = x_i + e \ \forall i = 1,\ldots,m$. As you can appreciate, modification of $K_w$ in equation 6 depends of the previous value of $\Delta w$ denoted by $\Delta w_{old}$ obtained with the previous input pattern. Once trained the AM, when it is used by first time, the value of $\Delta w_{old}$ is set to zero.

Once synapses of the associative memory have been modified in response to an input pattern, every component of vector $\tilde{y}$ can be recalled by using its corresponding input vector $\tilde{x}$ as:

$$\tilde{y}_i = \text{mid}(w_j + \tilde{x}_j), j = 1,\ldots,n$$

(7)

In short, pattern $\tilde{y}$ can be recalled by using its corresponding key vector $\tilde{x}$ or $\tilde{x}$ in six stages as follows:

1. Obtain index of the active region $ar$ by using equation 2.
2. Transform $\tilde{x}$ using de-codifying pattern $\xi$ by applying the following transformation: $\tilde{x}^k = x^k + \xi^k$.
3. Compute adjust factor $\Delta w = \Delta(x)$ by using equation 5.
4. Modify synapses of associative memory $w$ that belong to $K_w$ by using equation 6.
5. Recall pattern $\tilde{y}$ by using equation 7.
6. Obtain $\tilde{y}$ by transforming $\tilde{y}$ using de-codifying pattern $\xi$ by applying transformation: $\tilde{y}^k = \tilde{y}^k - \xi^k$.

The formal set of prepositions that support the correct functioning of this dynamic model can be found in [14].

3. MODIFIED DAM

In [15] the authors describe an interesting idea for recognizing faces based on the random selection of stimulating points (pixels), see figure 1.

![Fig. 1 – Random selection of stimulation points used to train an AM.](image)

In order to recognize complex images such as faces, we add to the DAM model a vector of stimulating points $SP$ where each stimulating point, given by $sp_i = \text{random}(n)$, is a random number between zero and the length of input pattern and $i = 1,\ldots,c$ where $c$ is the number of stimulating points used. To determine the active region we allocate in the DAM model an alternative simplified version of each pattern $x^k$ given by:

$$ss^k = sss(x^k) = x^k_{sp}$$

(8)

Once compute these simplified versions we could
estimate the active region as follows:

\[ ar = r(x) = \arg \max_{i=1}^{p} \left( \frac{x_i}{p} \right) \]  \hspace{1cm} (9)

where \( a_i = a_i + 1 \), \( b_i = \arg \min_{i=1}^{p} \left( \left| ss(x) - ss_i \right| \right) \) and \( i = 1, \ldots, c \).

Building phase of the DAM is done as follows:
1. Let \( I^x = I^y \) an association of images and \( c \) be the number of stimulating points.
2. Take at random a stimulating point \( s_{p_i}, i = 1, \ldots, c \).
3. For each association:
   4. Select filter size and apply it to the stimulating points in the images.
   5. Transform the images into a vector \( (x^i, y^i) \) by means of the standard image scan method.
   6. Train the DAM as in building procedure and compute the alternative simplified version of the patterns by using equation 8.

Pattern \( I^y \) can be recalled by using its corresponding key image \( I^x \) or distorted version \( \widetilde{I}^y \) as follows:
1. Use the same stimulating point, \( s_{p_i}, i = 1, \ldots, c \) and filter size as in building phase.
2. Apply filter to the stimulating points in the images.
3. Transform the images into a vector by means of the standard image scan method.
5. Apply steps from two to six as described in recalling procedure.

**4. ARCHITECTURE OF THE NETWORK**

Classical AMs (see for example [1], [2], [3], [4], [5], [6], [7] and [8]) are able to recover a pattern (an image) from a noisy version of it. In their original form classical AMs are not useful when image is altered by image transformations, such as translations, rotations, and so on.

The network of AMs proposed in this paper is robust under some of these transformations. Taking advantage of this fact, we can associate different versions of an image (rotated, translated and deformed) to an image.

Our task is to propose a network of AMs aimed to associate an image with other images belonging to the same collection. In order to achieve this, first suppose we want to associate images belonging to a collection with an image of the same collection using an AM. A good solution could be to compute the average image of whole images belonging to the collection and then associate the average image with any image that belongs to the collection. The same solution can be applied to other collections. Once computed the average images from different collections and chosen the images to be associated, we can train the AM as was described in section 2.

Until this point the AM only can recover an association between a collection of input patterns \( X \) and output pattern \( y \) denoted as, \( \{(X^k, y^k)\}_{k=1, \ldots, p} \) where \( p \) is the number of association, \( X^k = \{x^i, \ldots, x^r\} \) is a collection of input patterns and \( r \) is the number of patterns belonging to collection \( X \). This means that it can only be recovered the associated image using any image from a collection. However, we would like to get the inverse result; instead of recovering the associated image using any image from a collection, we would like to recover all the images belonging to the collection using any image of the collection.

To achieve this goal we will train a network of AMs built as in previous sections. Each AM will associate all the images of a collection with one image of this collection. This implies that for recovering all images of a given collection, we would need \( r \) AMs, where \( r \) is the number of images belonging to the collection. The network architecture of AMs needed for recovering a collection of images is shown in Fig. 2.

![Fig. 2 – Architecture of a network of AMs for recalling a collection of patterns using an input pattern.](image)

In order to train the network of \( r \) AMs, first of all we need to know the number of collections we want to recover. Training phase is done as follows:
1. Transform each image into a vector.
2. Build \( n \) collections of images \( \{CI^i\}_{i=1}^q \) where \( q \) is the number of pixels of each image and \( r \) the number of images.
3. Let \( [AI]_{i=1}^n \) a matrix of average images. For \( k = 1 \) to \( n \) compute the average image as:

\[ AI^k = \frac{\sum_i CI^i_k}{r} \]  \hspace{1cm} (10)

4. For \( s = 1 \) to \( r \) build an AM\(^s\) as described in Section 4 (building phase). For \( k = 1 \) to \( n \)
\[ x^i = A I^i \quad \text{and} \quad y^i = C I^i \]

Once trained the network of AMs, when is fed with any image of a collection, each AM will respond with an image that belongs to the collection. To recover a collection of images we just operate each AM with the input image as described in Section 3 (recalling phase).

Some important to remark is that once the collection of associative memories has been stimulated by an input pattern, we do not need any decision rule to determine the final class. Each associative memory will recall the corresponding output pattern which belongs to the same collection.

6. EXPERIMENTAL RESULTS

In this section the accuracy of the proposal is tested using two different benchmarks of images, see Fig. 3 and Fig. 4.

![Fig. 3 – (a-e) Collections of images taken from the Amsterdam Library of Objects Images (ALOI).](image)

![Fig. 4. (a-e) Collections of images taken from the Essex Collection of facial images.](image)

From the first benchmark 20 complex images were grouped into 5 collections composed by 4 images. After that, we proceeded to train the network of AMs as was explained in Section 5. For this problem our network is composed of four AMs. It is important to say that the number of images composing a collection could by any, the only restrictions to guaranty perfect and robust recall is that patterns (images) satisfy propositions described in [14].

Once trained the network, five experiments were performed to test the accuracy of the proposal. The first four experiments use the original DAM and the last experiment uses the modified DAM. The first experiment consisted on recovering a collection of images using any image of the collection in order to verify how much robust is the proposal under image deformations. Second experiment consisted on recovering a collection of images using any image of the collection altered by mixed noise in order to verify how much robust is the proposal under deformations and noisy version of the images. In the third experiment each image of the training set was rotated (from 0 to 360). We then used them to fed the network of AMs in order to verify how much robust is the proposal under deformations and rotations. Finally for forth and five experiments the images of the training set were rotated (from 0 to 360) and translated, and then used them to fed the network of AMs in order to verify how much robust is the proposal under deformations, rotations and translations.

The accuracy of the proposal was of 100% in the first experiment. The five collections of images where perfectly recovered by using any image of the collection (20 images), in Table 1 are shown some results obtained in this experiment. Remember that we train the network of AMs with average images, so then; when we fed the network with an image of any collection this image could be seen as a deformed version of the average images.

The results provided by our proposal in this experiment show that the associative model used to train the network of AMs are robust under deformations. Something important to say is that if we use other associative models for training the network of AMs such as morphological or median AMs, the collections might not be correctly recovered due to they are not robust under this kind of transformations or deformations.
The accuracy of the proposal, in the second experiment, was of 100%. The five collections of images where recovered perfectly by using any image of the collection as input, even when these images were altered by noise (200 images). As you can see in Table 2, despite of the level of noise added to the images, the collections were correctly recovered. Despite of other associative models are robust to this kind of noise, they might not recover the all collections due to they are only robust under additive, subtractive and mixed noise but not to image transformations.

The accuracy of the proposal, in the third experiment, was also of 100%. The five collections of images where recovered perfectly even when rotated version of the images were used (700 images), see in Table 3. Some important to say is, to our knowledge, neither morphological AMs nor other classical models are robust under rotations. Due to we used simplified patterns using mid operator and due to this operator is invariant to rotations the accuracy of the proposal was of 100%.

The accuracy of the proposal, in the fourth experiment was of 40%. In this experiment with 700 images; with some images we recalled a collection, but with some other images (when patterns did not satisfied the proposition that guarantees robust recall) the collections were not recalled, see Table 4. However, the results obtained by our proposal are acceptable if they are compared with the results provided by order associative models (less of 10% of accuracy).

Table 2. Some results obtained for the second experiment. Despite of the noise added to the images, all sets of images were again correctly recovered.

<table>
<thead>
<tr>
<th>Input image</th>
<th>Image recovered by each AM</th>
</tr>
</thead>
</table>

Table 3. Some results obtained for the third experiment. Despite of the noise added to the images and rotations, all sets of images were again correctly recovered.

<table>
<thead>
<tr>
<th>Input image</th>
<th>Image recovered by each AM</th>
</tr>
</thead>
</table>

Table 4. Some results obtained for the fourth experiment. With some images we recalled a collection, but with some other images (when patterns did not satisfied the proposition that guarantees robust recall) the collections were not recalled.

<table>
<thead>
<tr>
<th>Input image</th>
<th>Image recovered by each AM</th>
</tr>
</thead>
</table>

Finally, if we use the modified DAM the accuracy increases, in average, to more than 80%. In this experiment we selected different numbers of stimulation points, as in [15].

For the second benchmark 300 complex images of faces were grouped into 15 collections composed by 20 images. After that, we proceeded to train the network of AMs as was explained in Section 5. For this problem our network was composed of 15 AMs.

Once trained the network, two experiments were performed to test the accuracy of the proposal. The first experiment uses the original DAM and the second experiment uses the modified DAM. The first experiment consisted on recovering a collection of faces using any face of the collection in order to verify how much robust is the proposal under face recognition. In these experiments we did not include
noise or affine transformation because of the different gesticulations in the faces.

The accuracy of the proposal using the original DAM was of 20%. In this experiment with 300 images; with some images we recalled a collection, but with some other images (when patterns did not satisfy the proposition that guarantees robust recall), due to the gesticulations and illumination changes, the collections were not recalled.

For the last experiment, using the modified DAM, the accuracy increased to more than 95%. These results were obtained with more than 100 stimulating points.

In general, the accuracy of the proposal with different banks of images (altered by noisy and rotated) was of 100%. This was due to the input patterns (the images) satisfy the propositions presented in [14]. If these patterns do not satisfy these propositions, as images used in experiment four (translated and rotated images), the accuracy of the proposal diminish. However, the results provided by our proposal using the modified DAM, up-performed the results provide by the original DAM and other associative models.

5. CONCLUSION

In this paper we have proposed a network of AMs. This network is useful for recalling a collection of output patterns using an input pattern as a key. The network is composed by several dynamic associative memories (DAM). This DAM is inspired in some aspects of human brain. The model, due to plasticity of its synapses and functioning, is robust under some transformations as rotation, translation and deformations.

In addition, we describe an algorithm for training a network of AMs codifying the images of a collection by using an average image. Once computed the average images we proceed to training the network of AMs.

The network is capable to recall a collection of images (patterns) even if images are altered by noise or suffer some deformations and rotations.

Due to the original DAM present some problems with translations and face recognition we modified the DAM using a random selection technique. This technique allow the DAM up-perform the results obtained with the original DAM.

Through several experiments we have shown the efficiency of the proposal. In the first three experiments the proposal provided an accuracy of 100%. Even when the images were altered with mixed noise and rotated, the network of AMs recovered the corresponding collection. When object in images suffer translations or the object are faces, the accuracy of the proposal diminished. This is because most of the patterns under this transformation do not satisfied the propositions that guarantee robust recall. By using the modified DAM the accuracy of the proposal increases to more than 80% with translated objects and almost the 100% with complex faces. In general, the results provided by our proposal up-performed the results provide by other associative models.

The performed experiments could be seen as an application in image retrieval problems. We could say that we have developed a small system able to recover a collection of images (previously organized), even in the presence of altered versions of the images.

Nowadays we are working and directed this research to solve real problems related with signal and image retrieval. We are focusing our efforts to propose new associative models able to associate and recall images under more complex transformations. Furthermore, this new models have to work with images of much more complicated objects such as flowers, animals, cars, etc.

ACKNOWLEDGMENTS

This work was economically supported by SIP-IPN under grants 20071438 and 20082948 and CONACYT under grant 46805.

6. REFERENCES


Roberto A. Vazquez, was born in Mexico City, Mexico, 1981. He received his B. Sc. degree from the School of Computer Sciences, National Polytechnic Institute (ESCOM-IPN), Mexico City, Mexico, 2003. He received his M. Sc. degree from the Center for Computing Research, National Polytechnic Institute (CIC-IPN), Mexico City, Mexico, 2005.

He is now pursuing the Ph. D. degree at the Center for Computing Research, National Polytechnic Institute (CIC-IPN), Mexico City, Mexico.

His main research interests are Artificial Intelligence, Neurocomputing, Computational Neuroscience, Associative Memories, Pattern Recognition and Image Analysis.

Humberto Sossa, was born in Guadalajara, Jalisco, Mexico in 1956. He received a B. Sc. Degree in Electronics from the University of Guadalajara in 1981, a M.Sc. in Electrical Engineering from CINVESTAV-IPN in 1987 and a Ph.D. in Informatics from the National Polytechnic Institute of Grenoble, France in 1992.

He is Full Time Professor at the Center for Computing Research of the National Polytechnic Institute of Mexico.

His main research interests are in Pattern Recognition, Neurocomputing, Computational Neuroscience, Associative Memories, Image Analysis, and Robot Control using Image Analysis.