0

VG-RAM Weightless Neural Networks for Face Recognition

Alberto F. De Souza, Claudine Badue, Felipe Pedroni, Stiven Schwanz Dias, Hallysson Oliveira and Soterio Ferreira de Souza

Departamento de Informática
Universidade Federal do Espírito Santo
Av. Fernando Ferrari, 514, 29075-910 - Vitória-ES

Brazil

1. Introduction

Computerized human face recognition has many practical applications, such as access control, security monitoring, and surveillance systems, and has been one of the most challenging and active research areas in computer vision for many decades (Zhao et al.; 2003). Even though current machine recognition systems have reached a certain level of maturity, the recognition of faces with different facial expressions, occlusions, and changes in illumination and/or pose is still a hard problem.

A general statement of the problem of machine recognition of faces can be formulated as follows: given an image of a scene, (i) identify or (ii) verify one or more persons in the scene using a database of faces. In identification problems, given a face as input, the system reports back the identity of an individual based on a database of known individuals; whereas in verification problems, the system confirms or rejects the claimed identity of the input face. In both cases, the solution typically involves segmentation of faces from scenes (face detection), feature extraction from the face regions, recognition, or verification. In this chapter, we examine the recognition of frontal face images required in the context of identification problems.

Many approaches have been proposed to tackle the problem of face recognition. One can roughly divide these into (i) *holistic* approaches, (ii) *feature-based* approaches, and (iii) *hybrid* approaches (Zhao et al.; 2003). Holistic approaches use the whole face region as the raw input to a recognition system (a classifier). In feature-based approaches, local features, such as the eyes, nose, and mouth, are first extracted and their locations and local statistics (geometric and/or appearance based) are fed into a classifier. Hybrid approaches use both local features and the whole face region to recognize a face.

Among holistic approaches, eigenfaces (Turk and Pentland; 1991) and fisherfaces (Belhumeur et al.; 1997; Etemad and Chellappa; 1997) have proved to be effective

in experiments with large databases. Feature-based approaches (Lee and Seung; 1999; Li et al.; 2001; Gao and Leung; 2002) have also been quite successful and, compared to holistic approaches, are less sensitive to facial expressions, variations in illumination and occlusion. Some of the hybrid approaches include the modular eigenface approach (Martinez; 2002), the Flexible Appearance Model approach (Lanitis et al.; 1995), an approach that combines component-based recognition with 3D morphable models (Huang et al.; 2003), and an approach that encodes geometric and structural information extracted from the face image in attributed relational graphs (ARG) and matches Face-ARG's for recognition (Park et al.; 2005). Experiments with hybrid approaches showed slight improvements over feature-based approaches.

Recently, Wright et al. (2009) proposed a new approach to face recognition named Sparse Representation-based Classification (SRC). SRC is based on the compressive sampling theory (Candès and Wakin; 2008) and can use the whole face, a combination of features, or both features and the whole face for recognition. In SRC, the recognition problem is casted as one of classifying among multiple linear regression models. Wright et al. (2009) argue that compressive sampling offers the key to address this problem and, based on a sparse representation computed by ℓ^1-minimization, they propose a general classification algorithm for face recognition that provides new insights into what kind of transformation one should perform on face images to extract data to use as the input of the classifier of the recognition system. They showed that, if sparsity in the recognition problem is properly harnessed, the choice of transformation is no longer critical. What they found that is critical is whether the size of the data vector extracted is sufficiently large and whether the sparse representation is properly selected. They discovered that unconventional image transformations such as downsampling and random projections perform just as well as conventional ones such as eigenfaces, as long as the dimension of the data vector extracted surpasses certain threshold, predicted by the theory of sparse representation (Wright et al.; 2009).

Virtual Generalizing Random Access Memory Weightless Neural Networks VG-RAM WNN (Aleksander; 1998) is an effective machine learning technique that offers simple implementation and fast training and test. In this chapter, we evaluated the performance of VG-RAM WNN on face recognition using the well known AR Face Database (Martinez and Benavente; 1998) and Extended Yale Face Database B (Georghiades et al.; 2001; Lee et al.; 2005). We examined two VG-RAM WNN architectures, one holistic and the other feature-based, each implemented with different numbers of neurons and synapses per neuron. Using the AR Face Database, we compared the best VG-RAM WNN performance with that of: (i) a holistic approach based on principal component analysis (PCA) (Turk and Pentland; 1991); (ii) feature-based approaches based on non-negative matrix factorization (NMF) (Lee and Seung; 1999), local non-negative matrix factorization (LNMF) (Li et al.; 2001), and line edge maps (LEM) (Gao and Leung; 2002); and (iii) hybrid approaches based on weighted eigenspace representation (WER) (Martinez; 2002) and attributed relational graph (ARG) matching (Park et al.; 2005). In addition, using both the AR Face Database and the Extended Yale Face Database B, we compared the best VG-RAM WNN performing architecture (feature-based) with that of SRC. We selected these approaches for comparison because they are representative of some of the best techniques for face recognition present in the literature. Our results showed that, even training with a single face image per person, VG-RAM WNN outperformed PCA, NMF, LNMF, LEM, WER, and ARG approaches under all face conditions tested. Also, training and testing in the same conditions as those employed by Wright et al. (2009) (downsampled face images), VG-RAM WNN outperformed

SRC. These results show that VG-RAM WNN is a powerful technique for tackling this and other important problems in the pattern recognition realm.

This chapter is organized as follows. Section 2 introduces VG-RAM WNN and Section 3 describes how we have used them for face recognition. Section 4 presents our experimental methodology and experimental results. Our conclusions follows in Section 5.

2. VG-RAM WNN

RAM-based neural networks, also known as n-tuple classifiers or weightless neural networks, do not store knowledge in their connections but in Random Access Memories (RAM) inside the network's nodes, or neurons. These neurons operate with binary input values and use RAM as lookup tables: the synapses of each neuron collect a vector of bits from the network's inputs that is used as the RAM address, and the value stored at this address is the neuron's output. Training can be made in one shot and basically consists of storing the desired output in the address associated with the input vector of the neuron (Aleksander; 1966) (see Figure 1).

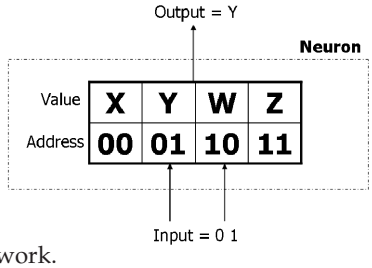


Fig. 1. Weightless neural network.

In spite of their remarkable simplicity, RAM-based neural networks are very effective as pattern recognition tools, offering fast training and test, in addition to easy implementation (Aleksander; 1998). However, if the network input is too large, the memory size becomes prohibitive, since it must be equal to 2^n, where n is the input size. Virtual Generalizing RAM (VG-RAM) Weightless Neural Networks (WNN) are RAM-based neural networks that only require memory capacity to store the data related to the training set (Ludermir et al.; 1999). In the neurons of these networks, the memory stores the input-output pairs shown during training, instead of only the output. In the test phase, the memory of VG-RAM WNN neurons is searched associatively by comparing the input presented to the network with all inputs in the input-output pairs learned. The output of each VG-RAM WNN neuron is taken from the pair whose input is nearest to the input presented—the distance function employed by VG-RAM WNN neurons is the *hamming distance*. If there is more than one pair at the same minimum distance from the input presented, the neuron's output is chosen randomly among these pairs.

Figure 2 shows the lookup table of a VG-RAM WNN neuron with three synapses (X_1, X_2 and X_3). This lookup table contains three entries (input-output pairs), which were stored during the training phase (entry #1, entry #2 and entry #3). During the test phase, when an input vector (input) is presented to the network, the VG-RAM WNN test algorithm calculates the distance between this input vector and each input of the input-output pairs stored in the lookup table. In the example of Figure 2, the *hamming distance* from the input to entry #1 is

lookup table	X_1	X_2	X_3	Y
entry #1	1	1	0	class 1
entry #2	0	0	1	class 2
entry #3	0	1	0	class 3
	↑	↑	↑	↓
input	1	0	1	**class 2**

Fig. 2. VG-RAM WNN neuron lookup table.

two, because both X_2 and X_3 bits do not match the input vector. The distance to entry #2 is one, because X_1 is the only non-matching bit. The distance to entry #3 is three, as the reader may easily verify. Hence, for this input vector, the algorithm evaluates the neuron's output, Y, as class 2, since it is the output value stored in entry #2.

3. Face Recognition with VG-RAM WNN

We examined the recognition part of the face identification problem only. That is, in the experiments described in this chapter, the segmentation of faces from images (face detection) is performed semi-automatically. Also, thanks to the properties of the VG-RAM WNN architectures employed, explicit feature extraction (e.g., line edge extraction; eye, nose, or mouth segmentation; etc.) is not required, even though in one of the two VG-RAM WNN architectures studied some neurons specializes in specific regions of the faces and, because of that, we say it is feature-based. The other VG-RAM WNN architecture studied is holistic.

3.1 Holistic Architecture

The holistic architecture has a single bidimensional array of $m \times n$ VG-RAM WNN neurons, N, where each neuron, $n_{i,j}$, has a set of synapses $W = \{w_1, \ldots, w_{|W|}\}$, which are randomly connected to the network's bidimensional input, Φ, of $u \times v$ inputs, $\varphi_{k,l}$ (see Figure 3 and Figure 4). The random synaptic interconnection pattern of each neuron $n_{i,j}$, $\Omega_{i,j}(W)$, is created when the network is built and does not change afterwards.

VG-RAM WNN synapses can only get a single bit from the input. Thus, in order to allow our VG-RAM WNN to deal with images, in which a pixel may assume a range of different values, we use *minchinton cells* (Mitchell et al.; 1998). In the proposed VG-RAM WNN architectures, each neuron's synapse, w_t, forms a minchinton cell with the next, w_{t+1} ($w_{|W|}$ forms a minchinton cell with w_1). The type of the minchinton cell we have used returns 1 if the synapse w_t of the cell is connected to an input element, $\varphi_{k,l}$, whose value is larger than the value of the element $\varphi_{r,s}$ to which the synapse w_{t+1} is connected, i.e., $\varphi_{k,l} > \varphi_{r,s}$; otherwise, it returns zero (see the synapses w_1 and w_2 of the neuron $n_{m,n}$ of Figure 4).

The input face images, I, of $\xi \times \eta$ pixels (Figure 4) must be transformed in order to fit into the network's input, Φ. In the case of the AR Face Database, the images are rotated, scaled and cropped (Figure 5); the rotation, scaling and cropping are performed semi-automatically, i.e., the position of the eyes are marked manually and, based on this marking, the face in the image is computationally adjusted to fit into Φ. Before being copied to Φ, the transformed image

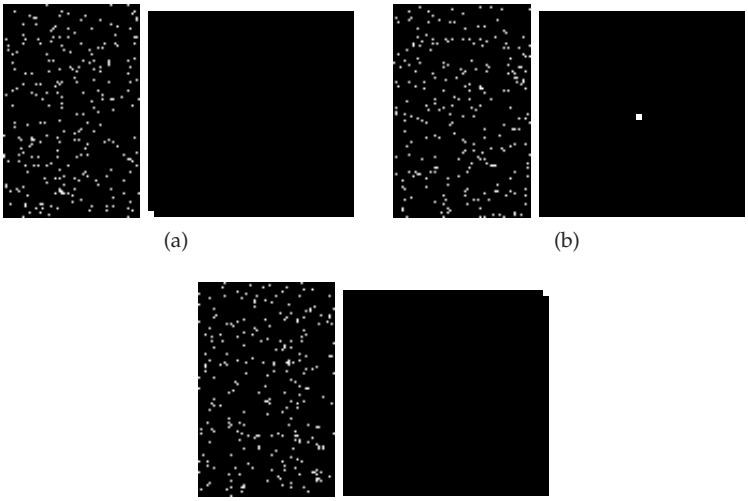


Fig. 3. The synaptic interconnection pattern of the holistic architecture. (a) Left, input Φ: in white, the elements $\varphi_{k,l}$ of the input Φ that are connected to neuron $n_{1,1}$ of N via $\Omega_{1,1}(W)$. Right, neuron array N: in white, the neuron $n_{1,1}$ of N. (b) Left: in white, the elements $\varphi_{k,l}$ of Φ connected to $n_{\frac{m}{2},\frac{n}{2}}$ via $\Omega_{\frac{m}{2},\frac{n}{2}}(W)$. Right: in white, the neuron $n_{\frac{m}{2},\frac{n}{2}}$ of N. (c) Left: in white, the elements of Φ connected to $n_{m,n}$ via $\Omega_{m,n}(W)$. Right: in white, the neuron $n_{m,n}$.

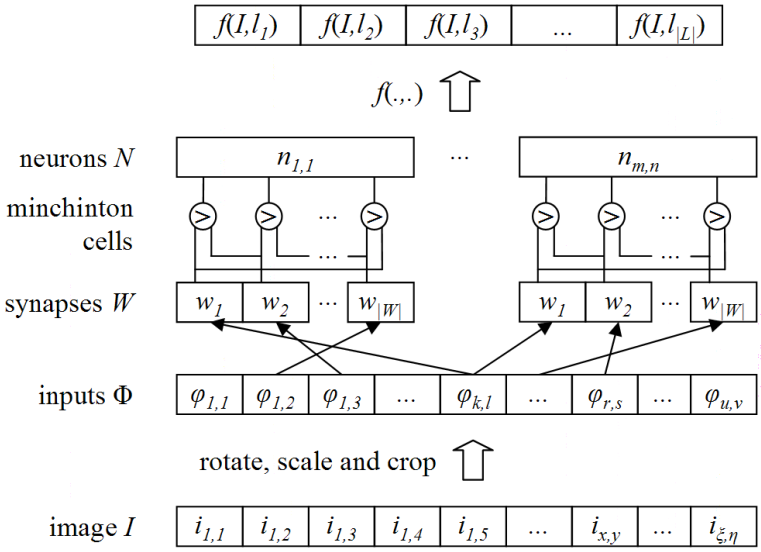


Fig. 4. Schematic diagram of the holistic and feature-based VG-RAM WNN architectures.

(a) (b) (c)

Fig. 5. Face image and its preprocessing. (a) Original image; (b) rotated, scaled and cropped image; and (c) filtered image.

is filtered by a Gaussian filter to smooth out artifacts produced by the transformations (Figure 5(c)). In the case of the Extended Yale Face Database B, only scaling and filtering are necessary, since this database includes versions of the images already properly cropped (Lee et al.; 2005).

During training, the face image I_x of a person p is transformed and filtered, and its pixels are copied to the VG-RAM WNN's input Φ and all $n_{i,j}$ neurons' outputs are set to the value of the label $l_p \in L = \{l_1, \ldots, l_{|L|}\}$, associated with the face of the person p ($|L|$ is equal to the number of known persons). All neurons are then trained to output this label with this input image. This procedure is repeated for all images I_x of the person p and, likewise, for all persons in the training data set. During testing, each face image I_y is also transformed, filtered, and copied to the VG-RAM WNN's input Φ. After that, all neurons' outputs are computed and the number of neurons outputting each label is counted by a function $f(I_y, l_p)$ for all $l_p \in L = \{l_1, \ldots, l_{|L|}\}$. The network's output is the label with the largest count.

3.2 Feature-Based Architecture

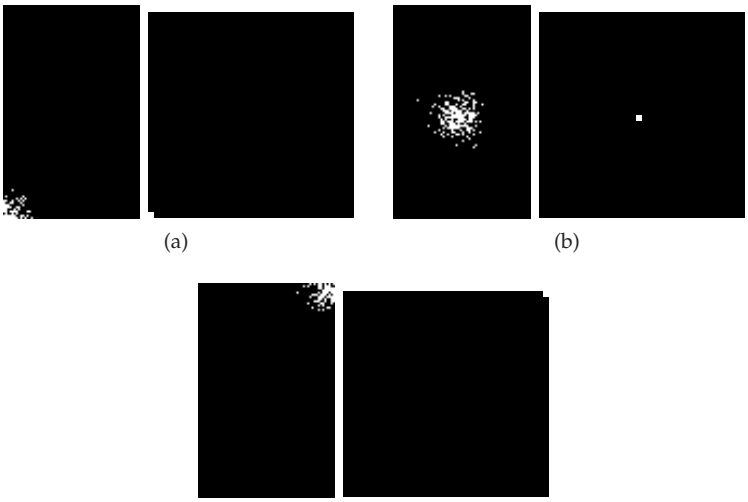
As the holistic architecture, the feature-based architecture has a single bidimensional array of $m \times n$ VG-RAM WNN neurons, N, where each neuron, $n_{i,j}$, has a set of synapses, $W = \{w_1, \ldots, w_{|W|}\}$, which are connected to the network's bidimensional input, Φ, of $u \times v$ inputs. The synaptic interconnection pattern of each neuron $n_{i,j}$, $\Omega_{i,j,\sigma}(W)$, is, however, different (Figure 6). In the feature-based architecture, $\Omega_{i,j,\sigma}(W)$ follows a bidimensional Normal distribution with variance σ^2 centered at φ_{μ_k,μ_l}, where $\mu_k = \frac{i.u}{m}$ and $\mu_l = \frac{j.v}{n}$; i.e., the coordinates k and l of the elements of Φ to which $n_{i,j}$ connects via W follow the probability density functions:

$$\omega_{\mu_k,\sigma^2}(k) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(k-\mu_k)^2}{2\sigma^2}} \tag{1}$$

$$\omega_{\mu_l,\sigma^2}(l) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(l-\mu_l)^2}{2\sigma^2}} \tag{2}$$

where σ is a parameter of the architecture. This synaptic interconnection pattern mimics that observed in many classes of biological neurons (Kandel et al.; 2000), and is also created when the network is built and does not change afterwards.

A comparison between Figure 3 and Figure 6 illustrates the difference between the interconnection patterns of the holistic and feature-based architectures. In the feature-based architecture (Figure 6), each neuron $n_{i,j}$ monitors a region of the input Φ and, therefore, specializes in



(a) (b)

(c)

Fig. 6. The synaptic interconnection pattern of the feature-based architecture. (a) Left, input Φ: in white, the elements $\varphi_{k,l}$ of the input Φ that are connected to neuron $n_{1,1}$ of N via $\Omega_{1,1,\sigma}(W)$. Right, neuron array N: in white, the neuron $n_{1,1}$ of N. (b) Left: in white, the elements $\varphi_{k,l}$ of Φ connected to $n_{\frac{m}{2},\frac{n}{2}}$ via $\Omega_{\frac{m}{2},\frac{n}{2},\sigma}(W)$. Right: in white, the neuron $n_{\frac{m}{2},\frac{n}{2}}$ of N. (c) Left: in white, the elements of Φ connected to $n_{m,n}$ via $\Omega_{n,n,\sigma}(W)$. Right: in white, the neuron $n_{m,n}$.

the face features that are mapped to that region. On the other hand, each neuron $n_{i,j}$ of the holistic architecture monitors the whole face (Figure 3).

As in the holistic architecture, in the feature-based architecture each neuron's synapse, w_t, forms a minchinton cell with the next, w_{t+1}, and, before training or testing, the input face images, I, are transformed and only then copied to the VG-RAM WNN input Φ. Training and testing are performed the same way as in the holistic architecture.

4. Experimental Evaluation

We used the AR Face Database (Martinez and Benavente; 1998) and the Extended Yale Face Database B (Georghiades et al.; 2001; Lee et al.; 2005) to evaluate the performance of VG-RAM WNN on face recognition. The AR Face Database contains over 4,000 color images corresponding to 135 people's faces (76 men and 59 women). Images feature frontal view faces with different facial expressions, illumination conditions, and occlusions (sun glasses and scarf). Its 768×576 pixels pictures were taken under strictly controlled conditions, but no restrictions on wear (clothes, glasses, etc.), make-up, hair style, etc. were imposed to participants. Each person participated in two sessions, separated by two weeks (14 days) time. On each of those sessions, thirteen images of each person were taken: four with variations of expression (neutral expression, smile, anger and scream—first session Figure 7(a) and second session Figure 7(e)), three with different illumination conditions (left light on, right light on, all side lights on—Figure 7(b) and Figure 7(f)), three wearing large sun glasses in differ-

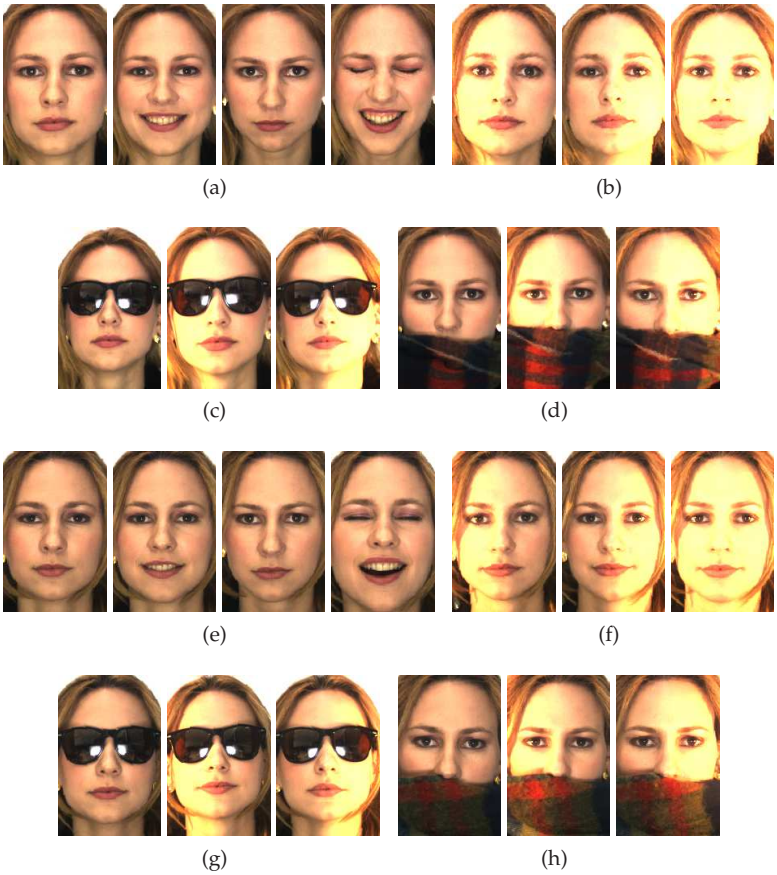


Fig. 7. Rotated, scaled and croped images of one person of the AR Face Database.

ent illumination conditions (Figure 7(c) and Figure 7(g)), and three wearing scarf in different illumination conditions (Figure 7(d) and Figure 7(h)).

The Extended Yale Face Database B consists of 2,414 frontal-face images of 38 individuals (Georghiades et al.; 2001). The manually cropped and 192×168 sized face images were captured under 64 different laboratory-controlled lighting conditions (Lee et al.; 2005). Figure 7 shows the 64 face images of one person of the Extended Yale Face Database B.

We used these face databases to perform two sets of experiments. In the first set, we used the AR Face Database to compare the performance of VG-RAM WNN with that of: (i) a holistic method based on principal component analysis (PCA) (Turk and Pentland; 1991); (ii) feature-based methods based on non-negative matrix factorization (NMF) (Lee and Seung; 1999), local non-negative matrix factorization (LNMF) (Li et al.; 2001), and line edge maps (LEM) (Gao and Leung; 2002); and (iii) hybrid methods based on weighted eigenspace representation (WER) (Martinez; 2002) and attributed relational graph (ARG) matching (Park et al.;

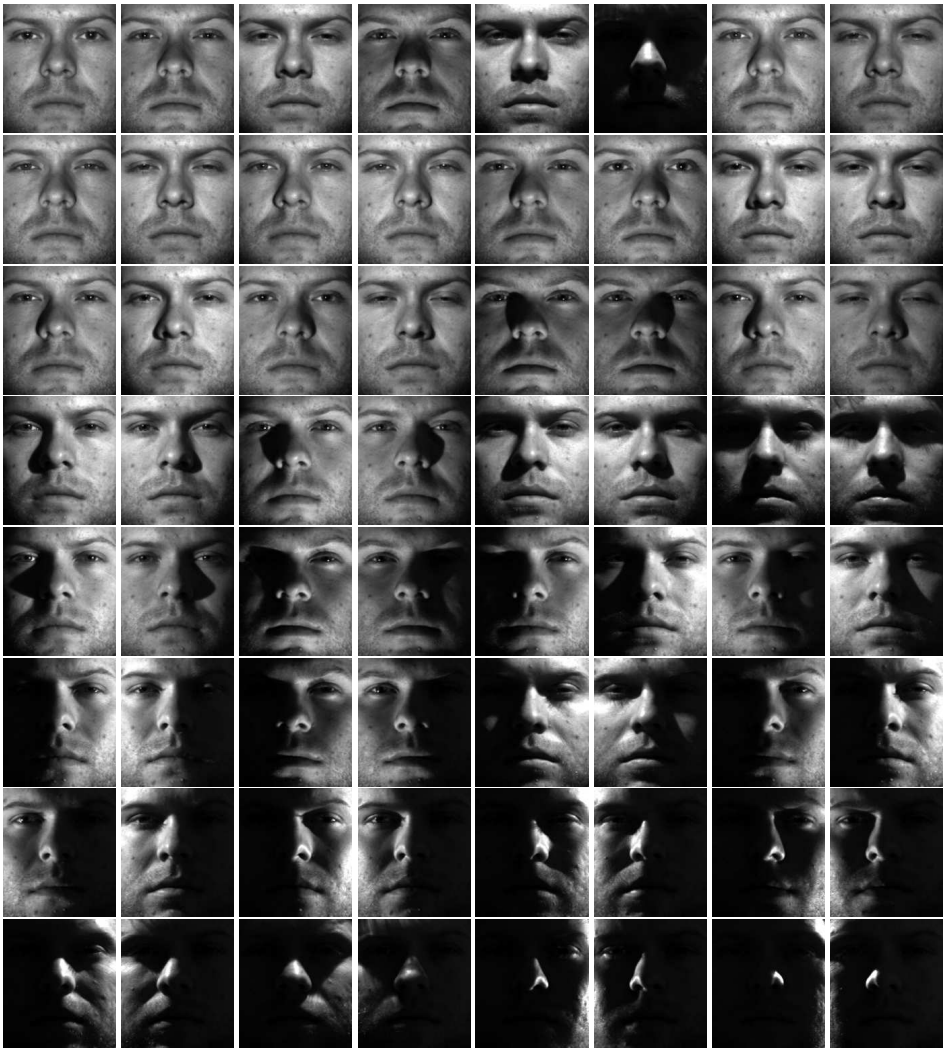


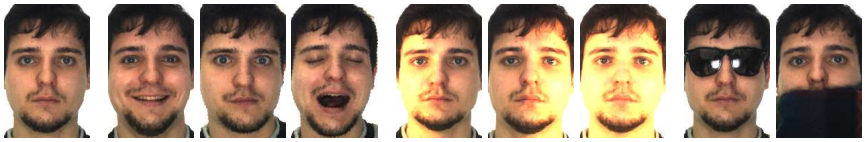
Fig. 8. Images of one person of the Extended Yale Face Database B.

2005). In the second set of experiments, we compared the performance of VG-RAM WNN with that of Sparse Representation-based Classification (SRC) (Wright et al.; 2009) using both the AR Face Database and the Extended Yale Face Database B. In the following sections we present these experiments.

4.1 VG-RAM WNN versus PCA, NMF, LNMF, LEM, WER, and ARG

In order to allow the comparison of VG-RAM WNN with that of PCA, NMF, LNMF, LEM, WER, and ARG, we used an experimental setup equivalent to that of Park et al. (2005). Park et al. (2005) proposed the ARG approach and compared it with PCA, NMF, LNMF, LEM, and WER. By using an equivalent experimental setup, we can compare VG-RAM WNN with his approach and the others mentioned.

As Park et al. (2005), we used only the following subset of image types of the AR Face Database: neutral expression, smile, anger, scream, left light on, right light on, all side lights on, wearing sun glasses (with a single illumination condition), and wearing scarf (with a single illumination condition). These can be divided into four groups: (i) normal (neutral expression); (ii) under expression variation (smile, anger, scream); (iii) under illumination changes (left light on, right light on, all side lights on); and (iv) with occlusion (wearing sun glasses, wearing scarf). We took these types of 768×576 sized face image of all persons in the AR Face Database and rotated, scaled, cropped and filtered them to obtain 128×200 face images that we used as the input Φ of our VG-RAM WNN. Figure 9 shows a set of transformed images of one subject of the AR Face Database (rotated, scaled and cropped to 128×200 sized images).



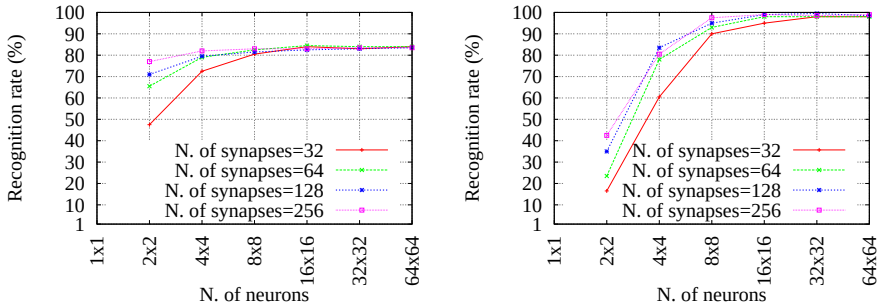
(a) (b) (c) (d)

Fig. 9. The AR face database: (a) normal (neutral expression); (b) under expression variation (smile, anger, scream); (c) under illumination changes (left light on, right light on, all side lights on); and (d) with occlusion (wearing sun glasses, wearing scarf).

We randomly selected 50 people from the database to tune the parameters of the VG-RAM WNN architectures (25 men and 25 women). We used one normal face image of each person to train (50 images), and the smile, anger, wearing sun glasses, and wearing scarf to evaluate the architectures (200 images) while varying their parameters. Below, we describe the experiments we performed to tune the parameters of the architectures.

4.1.1 Holistic Architecture Parameter Tunning

The holistic architecture has three parameters: (i) the number of neurons, $m \times n$; (ii) the number of synapses per neuron, $|W|$; and (iii) the size of the network input, $u \times v$. We tested networks with: $m \times n$ equal to 2×2, 4×4, 16×16, 32×32 and 64×64; number of synapses per neuron equal to 32, 64, 128 and 256; and $u \times v$ equal to 128×200 (we did not vary $u \times v$ to reduce the parameter search space). Figure 10(a) presents the results of the experiments we carried out to tune the parameters of the holistic architecture.



(a) (b)

Fig. 10. Performance tunning: (a) holistic architecture and (b) feature-based architecture.

As Figure 10(a) shows, the performance, i.e., the percentage of correctly recognized faces (recognition rate) of the holistic architecture grows with the number of neurons and synapses per neuron; however, as these numbers increase, the gains in performance decrease forming a plateau towards the maximum performance. The simplest configuration in the plateau has around 16×16 neurons and 64 synapses.

4.1.2 Feature-Based Architecture Parameter Tunning

The feature-based architecture has four parameters: (i) the number of neurons; (ii) the number of synapses per neuron; (iii) the size of the network input; and (iv) σ (see Section 3.2). We tested networks with: $m \times n$ equal to 2×2, 4×4, 16×16, 32×32 and 64×64; number of synapses per neuron equal to 32, 64, 128 and 256; $u \times v$ equal to 128×200; and σ equal to 10 (we did not vary $u \times v$ and σ to reduce the parameter search space).

Figure 10(b) presents the results of the experiments we conducted to tune the parameters of the feature-based architecture. As Figure 10(b) shows, the performance of the feature-based architecture also grows with the number of neurons and synapses per neuron, and again reaches a plateau at about 32×32 neurons and 128 synapses. However, it is important to note that, in this case, the plateau is very close to a recognition rate of 100%—the best performing configuration achieved a recognition rate of 99.5%.

4.1.3 Performance Comparison

We compared the performances of the holistic and feature-based VG-RAM WNN architectures with that of PCA, NMF, LNMF, LEM, WER, and ARG approaches. For that, we took the best VG-RAM WNN architectures configurations (holistic: 16×16 neurons and 64 synapses per neuron; feature-based: 32×32 neurons and 128 synapses per neuron), trained them with the normal face image of all people in the database (135 images), and tested them with the remaining face image categories of Figure 9 of all people in the database (135 images of each face image category). Table 1 summarizes this comparison, showing one technique on each line, grouped by type, and the corresponding performance for each face image category on each column.

Table 1. Comparison of the performance on the AR Face Database of the holistic (VWH) and feature-based (VWF) VG-RAM WNN architectures with that of: (i) PCA: principal component analysis (Turk and Pentland; 1991) (results obtained from (Park et al.; 2005)); NMF: non-negative matrix factorization (Lee and Seung; 1999) (results from (Park et al.; 2005)); LNMF: local non-negative matrix factorization (Li et al.; 2001) (results from (Park et al.; 2005)); LEM: line edge maps (Gao and Leung; 2002) (results from (Gao and Leung; 2002) with only 112 people of the AR Face Database); WER: weighted eigenspace representation (Martinez; 2002) (results from (Martinez; 2002) with only 50 people of the AR Face Database); and ARG: attributed relational graph matching (Park et al.; 2005) (results from (Park et al.; 2005)).

Type	Technique	Category							
		Smile	Anger	Scream	Glasses	Scarf	Left light	Right light	All side lights
HOL[a]	PCA	94.1%	79.3%	44.4%	32.9%	2.2%	7.4%	7.4%	2.2%
	VWH	**98.5%**	**97.8%**	**91.1%**	**66.7%**	**25.2%**	**97.8%**	**95.6%**	**95.6%**
FBA[b]	NMF	68.1%	50.4%	18.5%	3.7%	0.7%	N/A[d]	N/A	N/A
	LNMF	94.8%	76.3%	44.4%	18.5%	9.6%	N/A	N/A	N/A
	LEM	78.6%	92.9%	31.3%	N/A	N/A	92.9%	91.1%	74.1%
	VWF	**99.3%**	**99.3%**	**93.3%**	**85.2%**	**98.5%**	**99.3%**	**98.5%**	**99.3%**
HYB[c]	WER	84.0%	94.0%	32.0%	80.0%	82.0%	N/A	N/A	N/A
	ARG	97.8%	96.3%	66.7%	80.7%	85.2%	98.5%	96.3%	91.1%

[a]HOL: holistic techniques. [b]FBA: feature-based techniques. [c]HYB: hybrid techniques. [d]N/A: not available.

As the results in Table 1 show, the VG-RAM WNN holistic (VWH) architecture outperformed all holistic and feature-based techniques examined (except the VG-RAM WNN feature-based architecture - VWF) in all face image categories. It also performed better than the hybrid techniques, except for the categories with occlusion and single side illumination. That was expected, since occlusions and single side illumination compromise eventual similarities between the input patterns learned by the VWH neurons and those collected by its synapses from a partially occluded or illuminated face. Nevertheless, it is important to note the overall performance achieved by VWH, which is better than that of several other relevant techniques from literature.

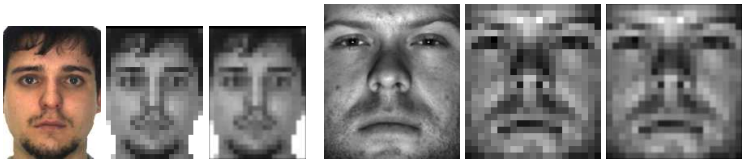
As Table 1 also shows, the VG-RAM WNN feature-based (VWF) architecture performed better than all other techniques examined in all categories and, in many cases, by a large margin.

4.2 VG-RAM WNN versus SRC

The central point of the SRC approach is that there are a lot of redundancy of information in face images, i.e., for the purpose of face recognition, the dimensionality of face images is typically too large because they are frequently oversampled. One can appreciate this by reasoning about the fact that images can be compacted; i.e., images sampled (or either oversampled) with 8 megapixels—which would result in a file with 8 megapixels $\times$ 3 bytes, one byte for each color, that is, 3 $\times$ 8 megabytes—can typically be compacted into a file of about one megabyte.

In the work of Wright et al. (2009), they studied several methods to reduce the dimensionality of the information extracted from face images for being used as input of the face recognition systems' classifiers. Therefore, in order to allow the comparison of VG-RAM WNN with that of SRC, we used an experimental setup equivalent to that of Wright et al. (2009).

We compared the best VG-RAM WNN performing architecture (feature-based) with that of SRC. For the experiments with the AR Face Database, as Wright et al. (2009) did, we rotated, scaled, and cropped the 768×576 sized face images to 120×165 sized images and, after that, downsampled the images at a ratio of $1/6$. The downsampled images have size 20×27, or 540 dimensions, which was the same used by Wright et al. (2009). After downsampling the images, we rescaled them back to 120×165 to use as the input Φ of the VG-RAM WNN (about the same size we used in the previous experiments, 128×200). Note that this does not add any information to the images; we did that in order to not change the parameters we have found in the tuning of the VG-RAM WNN feature-based architecture. After rescaling the images, we filtered them with a Gaussian filter to smooth out artifacts produced by the transformations. Again, it is important to note that this does not add any information to the images; it is required only for the proper work of our VG-RAM WNN. Figure 11(a) shows a transformed face image (rotated, scaled, and cropped), the downsampled version of this image, and the filtered version of this same image.



(a) (b)

Fig. 11. Face image subsampling. (a) AR Face Database. (b) Extended Yale Face Database B.

For the experiments with the Extended Yale Face Database B, also used by Wright et al. (2009), only scaling and filtering were necessary, since this database includes versions of the images already properly cropped. The image sizes in the Extended Yale Face Database B are different from those in the AR Face Database; we downsampled the 168×192 sized face images to 21×24 and, as we did with the AR Face Database, we rescaled these images back to 168×192 and filtered them. Figure 11(b) shows an original face image, the downsampled version of this image, and the filtered version of this same image.

In the case of the AR Face Database, following the same procedure of Wright et al. (2009), we randomly selected 50 men and 50 women. For each person, fourteen images with variations of expression and different illumination conditions were selected; the seven images from the first session were used for training and the other seven from the second session for testing. In the case of the Extended Yale Face Database B, for each person, we randomly selected half of the face images for training (i.e., about 32 images for each of the 38 people) and the other half for testing.

Table 2 summarizes the comparison of the performance of the VG-RAM WNN feature-based architecture with that of SRC, following the same format of the Table 1. The kind of face recognition system of Wright et al. (2009) is a holistic type.

Table 2. Comparison of the performance on the AR Face Database and the Extended Yale Face Database B of the VG-RAM WNN feature-based (VWF) architecture with that of the Sparse Representation-based Classification (SRC). Results for SRC were obtained from Wright et al. (2009)
.

Type	Technique	Database	
		AR Face Database	Extended Yale Face Database B
HOL[a]	SRC (Random)	94.70%	98.1%
FBA[b]	**VWF**	**98.86%**	**99.34%**

[a]HOL: holistic techniques. [b]FBA: feature-based techniques.

As the results in Table 2 show, VG-RAM WNN feature-based (VWF) architecture outperformed SRC for both databases and, in the case of the AR Face Database, by a large margin. The VWF superior performance, shown in both the Table 1 and Table 2, is the result of two factors. First, each VWF (or VWH) synapse collects the result of a comparison between two pixels, executed by its corresponding minchinton cell. Our best VWF has 128 synapses per neuron and 32×32 neurons. Therefore, during test, 131072 ($128 \times 32 \times 32$) such comparisons are executed on an input face image and the results are checked against equivalent results learned from training images. This amount of pixel comparisons allows not only high discrimination capability but also generalization. Second, thanks to the characteristics of the VWF architecture, i.e., its synaptic interconnection pattern, each VWF neuron monitors a specific region of the face only, which reduces the overall impact of occlusions and varying illumination conditions on recognition performance.

5. Conclusions and Future Work

In this work, we presented an experimental evaluation of the performance of Virtual Generalizing Random Access Memory Weightless Neural Networks (VG-RAM WNN Aleksander (1998)) on face recognition. We presented two VG-RAM WNN face recognition architectures, one holistic and the other feature-based, and examined its performance with two well known face database: the AR Face Database and the Extended Yale Face Database B. The AR Face Database is challenging for face recognition systems because it has images with different facial expressions, occlusions, and varying illumination conditions. The best performing architecture (feature-based) showed robustness in all image conditions and better performance than many other techniques from literature, even when trained with a single sample per person.

In future works, we will examine the performance of VG-RAM WNN with other databases and use it to tackle other problems associated with face recognition systems, such as face detection, face alignment, face recognition in video, etc.

6. Acknowledgments

We would like to thank *Receita Federal do Brasil*, *Conselho Nacional de Desenvolvimento Científico e Tecnológico*—CNPq-Brasil (grants 308207/2004-1, 471898/2004-0, 620165/2006-5), *Fundação Espírito Santense de Tecnologia* — FAPES-Brasil (grant 37711393/2007), and *Financiadora de Estudos e Projetos*—FINEP-Brasil (grants CT-INFRA-PRO-UFES/2005, CT-INFRA-PRO-UFES/2006) for their support to this research work.

7. References

Aleksander, I. (1966). Self-adaptive universal logic circuits, *IEE Electronic Letters* **2**(8): 231–232.

Aleksander, I. (1998). *RAM-Based Neural Networks*, World Scientific, chapter From WISARD to MAGNUS: a Family of Weightless Virtual Neural Machines, pp. 18–30.

Belhumeur, P. N., Hespanha, J. P. and Kriegman, D. J. (1997). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19**(7): 711–720.

Candès, E. and Wakin, M. (2008). An introduction to compressive sampling, *IEEE Signal Processing Magazine* **2**(25): 21–30.

Etemad, K. and Chellappa, R. (1997). Discriminant analysis for recognition of human face images, *Journal of the Optical Society of America A* **14**(8): 1724–1733.

Gao, Y. and Leung, M. K. (2002). Face recognition using line edge map, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(6): 764–779.

Georghiades, A. S., Belhumeur, P. N. and Kriegman, D. J. (2001). From few to many: Illumination cone models for face recognition under variable lighting and pose, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23**(6): 643–660.

Huang, J., Heisele, B. and Blanz, V. (2003). Component based face recognition with 3d morphable models, *Proceedings of the International Conference on Audio- and Video-Based Person Authentication*, pp. 27–34.

Kandel, E. R., Schwartz, J. H. and Jessell, T. M. (2000). *Principles of Neural Science*, 4th edn, Prentice-Hall International Inc.

Lanitis, A., Taylor, C. J. and Cootes, T. F. (1995). Automatic face identification system using flexible appearance models, *Image Vision Computing* **13**(5): 393–401.

Lee, D. D. and Seung, S. H. (1999). Learning the parts of objects by non-negative matrix factorization, *Nature* **401**: 788–791.

Lee, K. C., Ho, J. and Kriegman, D. (2005). Acquiring linear subspaces for face recognition under variable lighting, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**(5): 684–698.

Li, S. Z., Hou, X. W., Zhang, H. and Cheng, Q. (2001). Learning spatially localized, parts-based representation, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 207–212.

Ludermir, T. B., Carvalho, A. C. P. L. F., Braga, A. P. and Souto, M. D. (1999). Weightless neural models: a review of current and past works, *Neural Computing Surveys* **2**: 41–61.

Martinez, A. and Benavente, R. (1998). The AR face database, *Technical Report #24*, Computer Vision Center (CVC), Universitat Autònoma de Barcelona.

Martinez, A. M. (2002). Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(6): 748–763.

Mitchell, R. J., Bishop, J. M., Box, S. K. and Hawker, J. F. (1998). *RAM-Based Neural Networks*, World Scientific, chapter Comparison of Some Methods for Processing Grey Level Data in Weightless Networks, pp. 61–70.

Park, B.-G., Lee, K.-M. and Lee, S.-U. (2005). Face recognition using face-arg matching, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**(12): 1982–1988.

Turk, M. and Pentland, A. (1991). Eigenfaces for recognition, *Journal of Cognitive Neuroscience* **3**: 71–86.

Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S. and Ma, Y. (2009). Robust face recognition via sparse representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**(2): 210–227.

Zhao, W.-Y., Chellappa, R., Phillips, P. J. and Rosenfeld, A. (2003). Face recognition: A literature survey, *ACM Computing Surveys* **35**(4): 399–458.