# Aguará: An Improved Face Recognition Algorithm through Gabor Filter Adaptation

C. Aguerrebere, G. Capdehourat, M. Delbracio, M. Mateu, A. Fernández and F. Lecumberry

{gcapde, mdelbra}@fing.edu.uy

Instituto de Ingeniería Eléctrica de la Facultad de Ingeniería

Universidad de la República, Montevideo, Uruguay

## Abstract

*We developed an* EBGM-*based algorithm that successfully implements face recognition under constrained conditions. A suitable adaptation of the Gabor filters was found through a power spectral analisys (*PSD*) of the face images. We outperformed the best-known implementations of the* EBGM *algorithm in the* FERET *database. The results are comparable with those of the state of the art.*

## 1. Introduction

Automatic Face Recognition is a challenging problem in computer vision. Face recognition is by itself an open problem for different sciences (e.g. psychology, neuroscience). One of its main goals is the understanding of the complex human visual system and the knowledge of how humans represent faces in order to discriminate different identities with high accuracy. On the other side, several emerging applications, from law enforcement to commercial tasks, demand the industry to develop new face recognition systems.

A general way of defining the problem is: given still images or a video sequence, recognize one or more people at the scene by using a face database as an identity register. This problem, taken into account without constraints, has several challenges to be solved. Different illumination of the scene; changes in pose, orientation, expression and face occlusions are some examples of the issues to deal with. In our case, the problem is simplified by working under controlled conditions, which can be summarized as follows: one person per image, known background, uniform illumination, frontal pose, neutral expression and no occlusions. However, the robustness to different face expressions and non-uniform illumination are also evaluated in section 4.

Two basic and conceptually independent problems have to be addressed by this kind of systems: face detection in the scene and recognition of the detected face. We focused our work on the recognition stage, taking the detected face as the input to the algorithm. This stage can be separated in two steps: *feature extraction*, where important information for discrimination is saved, and the *matching step*, where the recognition result is given with the aid of a face database.

Several face recognition methods have been proposed during the last thirty years. In the vast literature on the topic (see [15] for a review), there are different classifications of the existing techniques. The following is one posible high-level classification:

- *Holistic Methods* The whole face image is used as the raw input to the recognition system. An example is the well-kown PCA-based technique introduced by Kirby and Sirovich [4], followed by Turk and Pentland [11].

- *Local Feature-based Methods* Local features are extracted, such as eyes, nose and mouth. Their locations and local statistics (appearance) are the input to the recognition stage. An example of this methods is Elastic Bunch Graph Matching (EBGM) [13] studied in this work.

There are also hybrid methods, which combine holistic and local techniques. A description of the most representative algorithms in each category can be seen in [15].

A keypoint in pattern recognition techniques is the *descriptor* used to represent the target object. The descriptor is the way to link the real object with a suitable representation. In this work, local descriptors were chosen, which also consider the fact that we are dealing with faces. Following the EBGM algorithm our work uses responses of Gabor wavelets as descriptors.

EBGM is a technique which uses local information extracted with Gabor filters for discrimination. The implementation used in this work was based on the open-source package developed by David Bolme in Colorado State University [1].

In section 2 a high level description of the algorithm is presented. Next, our main contribution is described in section 3. The most significant results are shown in section 4. Some concluding remarks and future work are finally stated in section 5.

## 2. EBGM-based Algorithm Description

Here we present an EBGM-based algorithm, inspired on [13], which consists in localizing certain fiducial facial points in order to extract local information. For this purpose, a family of Gabor wavelet responses are computed at each of the fiducial points, where the Gabor wavelets can be tuned in frecuency, orientation and spatial range of action in the image. A graph structure is defined on the face, where the nodes are the previously selected fiducial points. In general, easily human identifiable points are chosen, such as mouth and eyes corners or face contour.

A block diagram of the algorithm is depicted in figure 1. The first stage implements a normalization in order to compensate the different illumination and contrast (dynamic range) across the different face images. Also an affine transformation is applied to adjust all face images to have the same size. For this purpose, the location of the eyes is assumed known. Next, the locations of the fiducial
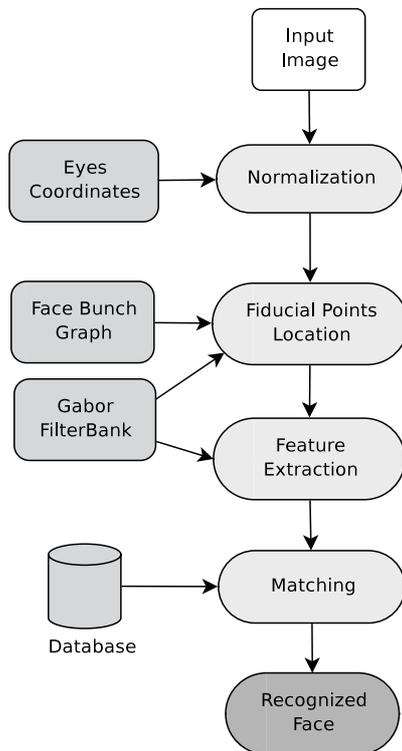
points in the normalized face are found using a statistical model (called *Face Bunch Graph* in [13]). Then, a set of Gabor filters is applied on each of the previously located fiducial points. Finally, in the matching stage, the obtained face descriptor is compared to the database for recognition. This stage depends on the system mode, identification or verification. In identification mode, the system must decide the input user identity comparing the extracted features with all the database entries. In verification mode, the user claims his identity with a PIN[1] and the system validates it or not, comparing only with the features in the database for the claimed identity.



Figure 2: Normalized face image examples: with eliptical mask (right) and without it (left).

### 2.1. Normalization

The normalization process is needed in order to reduce image variability, due to lighting and image acquisition. This preprocessing step is based on [1] and consist of the following steps:

1. A geometric affine transformation takes the eyes to predetermined fixed coordinates and the image is rescaled to $128 \times 128$ pixels.

2. Pixel values are adjusted to have zero mean and standard deviation one.

3. Image borders are smoothed across a band 30 pixels wide. In this way, possible artifacts due to the geometric transformation are avoided. Pixel values are weighted by a factor $\frac{d}{30}$, where $d$ is the distance to the image edge. Pixels farther than 30 pixels are not modified.

Optionally, between steps 2 and 3, illumination can be normalized by histogram equalization. An elliptic mask (see figure 2) can be used to keep the face and remove image background. This mask has an exponential decay to the image mean, which removes background in a smooth way.



Figure 1: System block diagram.

---
[1]Personal Identification Number

These variants did not show important differences in terms of the algorithm performance because of the database characteristics, but in a real environment, the results can change dramatically.

## 2.2. Local Descriptors: Gabor Wavelets

The basic expression of a Gabor wavelet is a Gaussian kernel function modulated by a sinusoid,

$$W(x,y) = f e^{-\frac{x'^2+y'^2}{2\sigma^2}} \left( \cos(2\pi f x' + \phi) - DC \right) \quad (1)$$

where:

$$DC = \cos\phi \; e^{-2\pi^2 \frac{\sigma^2}{f^2}} \;\; \text{and} \; \begin{cases} x' = x\cos\theta + y\sin\theta \\ y' = -x\sin\theta + y\cos\theta \end{cases}$$

The parameter $\sigma = kf$ ensures that the filter spatial-range of action is limited proportionally to the central frecuency $f$. We employed a discrete realization of (1) using 5 different frecuencies and 8 orientations, totalizing 40 complex-Gabor filters. The last term in equation (1) $DC$, makes the filter DC-free. A study of the suitable Gabor filters transfer function is presented in section 3.

## 2.3. Fiducial Points Location

A graph structure is defined over the face, where the nodes are the previously selected facial fiducial points. These points are chosen to be easily identifiable from face to face. Figure 3 shows an example of the fiducial points for a sample image.

The responses to the Gabor filter bank at each fiducial point is called *jet*. The graph labeled with the corresponding *jets* is called *Face Graph* and is used as a description of the face. For each node the coordinates of the fiducial point and its associated *jet* are saved. Graph edges are labeled with the distance between nodes.

For automatically locating fiducial points on new faces, the system needs a representative description of a generic face. This description has to be suitable for a wide range of variabilities like different eyes, mouth, presence of beard, different genders, etnias or ages. To acomplish this, a representative set of *Face Graphs* are combined in a stack-like structure called *Face Bunch Graph* (FBG). Each model graph has the same structure and fiducial points. Each node of the FBG is labeled with a set of *jets*, each one corresponding to the different face graph models and the average node coordinates in the set. The graph edges are labeled with the average distance between the nodes. In this way, the FBG captures the features of a general face and allows variations between people. In [1, 13] it is suggested that a model of
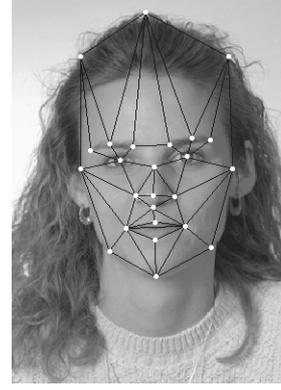


Figure 3: Fiducial points graph example.

70 images is a good balance between size and variability. A remarkable point is that facial fiducial points in model images can be manually marked or located in a semiautomatic way.

Once the FBG is constructed, the problem is to find the fiducial points for a given input image. This task is solved finding the input image *Face Graph* that maximizes the similarity with the considered FBG. Different ways of doing this were proposed [1, 13]. The main idea is summarized in the following steps:

1. An initial new node position is estimated from the previously located ones. For this, the averaged distance between the FBG nodes is used.

2. Given the initial estimation of a node, the corresponding *jet* is extracted from the normalized input image. The similarity with each *jet* in the FBG is calculated for that node. The objective is to maximize the similarity with the FBG in a neighborhood. Finally the coordinates where the maximun similarity was found are saved.

3. The procedure is iterated for the rest nodes of the graph.

The adjustment begins with the fiducial points located in the eyes that are always in the same place due to the normalization process. For step 2 there are several different ways of maximizing the similarity. For an exhaustive study on this topic see [1]. Another possibility is to make no refinement of the fiducial point estimation, so that nodes are placed considering the detected eyes coordinates and the averaged distances from the eyes given by the FBG. In the evaluation (see section 4) we saw that the refinement step does not improve the system performance.

## 2.4. Feature Extraction

After the fiducial points localization procedure, feature extraction comes up. The response to the Gabor filterbank is calculated on each located point. This information is saved for identity discrimination.

## 2.5. Matching

Several similarity measures were proposed in [1], all of them evaluating some similarity measure between *jets* and then averaging for all the nodes. The difference between them is the information considered from the *jets*. Some possibilities are: the module of the Gabor filter response, the phase or combinations of both. In this work, we present the results with the one that considers only the module. After testing the different options, we chose this one because it is the simplest one and achieves a good performance. The expression is the following:

$$L(G, G') = \frac{1}{n} \sum_{i=1}^{n} S_a(J_i, J_i')$$

where $J_i$ and $J_i'$ are *jets* $i$ of the graphs $G$ and $G'$ respectively; $n$ is the number of nodes for each graph and $S_a$ is the similarity measure between *jets* that considers only the module which has the expression:

$$S_a(J, J') = \frac{\sum_{j=1}^{N} a_j a_j'}{\sqrt{\sum_{j=1}^{N} a_j^2 \sum_{j=1}^{N} {a_j'}^2}}$$

with $a$ and $a'$ modules of each filter response and $N$ the number of complex Gabor filters.

## 3. Our contribution

Different experiments [2] show that natural images have a power spectral density (PSD) proportional to $1/f^2$. This fact is used in [13], where they included an extra factor $f$ multiplying equation (1), in order to take the decay of the natural images PSD into account. As it is stated by Lades *et al.* [5], the algorithm performance is seriously affected by this fact. In order to extend this power compensation to *face images*, the PSD of 50 FERET normalized images was estimated. The following procedure is based on [12],

1. Each image is preprocessed according to 2.1, without histogram equalization nor mask.

2. The square of the FFT module (PSD estimation) is computed after smoothing the normalized image with a Kaiser-Bessel window of parameter $\beta = 2$.

3. The PSD is circularly averaged in order to independize of a particular direction. This unidimensional vector is called $G(f)$, averaged power spectral density of this image. An example is presented in Figure 4.

4. A model of $G(f) = K f^\alpha$ is adjusted by minimum mean-square error to the image. Recall that for natural images $\alpha = -2$.

5. The above procedure is repeated for each image in the subset. The obtained values of $\alpha$ are averaged.
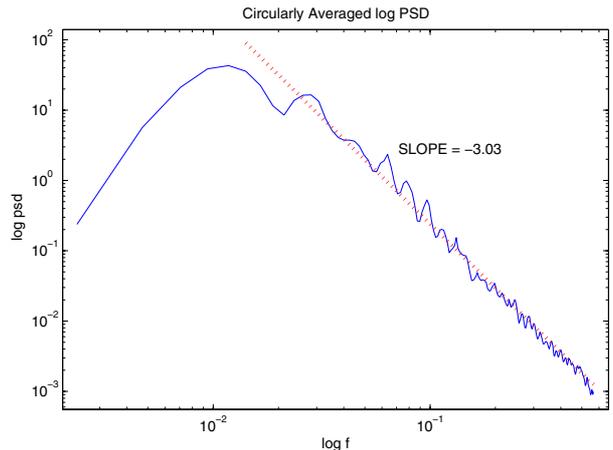


Figure 4: Circularly averaged log-PSD for a normalized face image example.

The results show that $\alpha = -3.05$, with a standard deviation $\sigma_\alpha = 0.12$. In order to compensate this decay, the suitable factor to multiply equation (1) is $f^{1,5}$. By this way, computing the *jet* similarity ($S_a$), every Gabor filter response is taken into account with the same weight for all the different frecuencies.

## 4. Evaluation and Results

### 4.1. Evaluation Protocol

We evaluated the algorithm with the FERET database[2], widely accepted in the scientific community, for which there is an standard evaluation procedure defined [8, 9]. In order to test the algorithm performance, we followed the semi-automatic FERET evaluation protocol, in which eye coordinates are manually marked. In this way, the face detection stage is avoided in the evaluation.

---

The FERET database has a total of 14051 images, corresponding to more than 1000 people, with complexity according to our needs. For testing the robustness in different conditions, five sets are provided, one is used as the gallery set and the other four ones are used as probe sets.

**Gallery set**

- **fa** - Regular facial expression. 1196 images.

**Probe sets**

- **fb** - Alternative facial expression. 1195 images.

- **fc** - Different illumination. 194 images.

- **dup I** - Taken between 1 minute and 1031 days after **fa**. 722 images.

- **dup II** - Taken at least 18 months after **fa**. 234 images.

With this image sets, four different tests are defined. Each one considers the gallery set (**fa**) and one of the probe sets (**fb**, **fc**, **dup I**, **dup II**). They are summarized in table 1.

| Task | Test Name | Gallery Set | Probe Set |
|---|---|---|---|
| Aging | Duplicate I | fa | dup I |
| Aging | Duplicate II | fa | dup II |
| Expression | FB | fa | fb |
| Illumination | FC | fa | fc |

Table 1: Different tests with FERET database.

Our system has mainly two flavours, **aguará_06_F** and **aguará_06_NF**. The first one uses the refinement step in the fiducial points location stage, while the second one does not. Both of them incorporate Gabor filters compensation according to normalized faces PSD.

## 4.2. EBGM-based Algorithms Comparison

We compare our algorithm with previous EBGM implementations, developed by Wiskott et al. [13] and David Bolme [1].

- bolme_03 - Colorado State University, D. Bolme [1].

- usc_mar_97 - University of Southern California, Wiskott et al. [13].

Table 2 shows a comparison of the RR values for the identification case. In figure 5 rank curves are presented for the **FB** test.

|  | RR | | | |
|---|---|---|---|---|
| **Algorithm** | **D-I** | **D-II** | **FB** | **FC** |
| usc_mar_97 | 59.1 | 52.1 | 95.0 | 82.0 |
| bolme_03 | 46.3 | 24.4 | 89.8 | 41.8 |
| **aguará_06_F** | **66.3** | **59.0** | **96.4** | **92.3** |
| **aguará_06_NF** | **68.8** | **63.2** | **95.5** | **91.2** |

Table 2: RR for different EBGM-based algorithms.
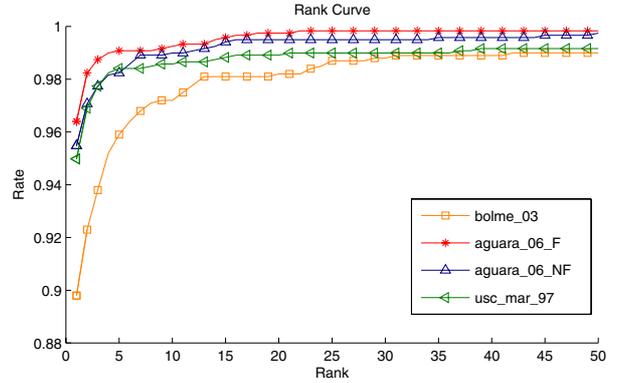


Figure 5: Rank curve for different EBGM-based algorithms for FERET **FB** test.

## 4.3. Comparison to Other Algorithms

In order to compare the results of this work with those of the state of the art, we choose a set of representative algorithms tested with the FERET database. Between them we can find some of the participants in the *FERET Test* [8, 9] (developed between September 1996 and March 1997) and other more recent ones.

- mit_sep_96 - MIT Media Lab, Moghaddam et al. 1996 [6].

- umd_mar_97 - University of Maryland, Zhao et al. 1998 [14].

- gabor_02 - Middle East Technical University, Kepenekci et al. 2002 [3].

- bayesian_03 - Colorado State University, M. Teixeira 2003 [10].

- log-G_PCA_06 - Kaunas University of Technology, V. Perlibakas 2006 [7].

The results for identification and verification are presented in tables 3 and 4.

| | RR | | | |
|---|---|---|---|---|
| **Algorithm** | **D-I** | **D-II** | **FB** | **FC** |
| mit_sep_96 | 57.6 | 34.2 | 94.8 | 32.0 |
| umd_mar_97 | 47.2 | 20.9 | 96.2 | 58.8 |
| gabor_02 | 58.3 | 47.7 | 96.3 | 69.6 |
| bayesian_03 | 52.4 | 31.6 | 81.9 | 37.1 |
| log-G_PCA | 72.4 | 65.8 | 98.0 | 90.2 |
| **aguará_06_F** | **66.3** | **59.0** | **96.4** | **92.3** |
| **aguará_06_NF** | **68.8** | **63.2** | **95.5** | **91.2** |

Table 3: RR for different state of the art algorithms.

| | EER | | | |
|---|---|---|---|---|
| **Algorithm** | **D-I** | **D-II** | **FB** | **FC** |
| mit_sep_96 | 17.7 | 21.2 | 4.9 | 18.0 |
| umd_mar_97 | 12.6 | 13.4 | 1.2 | 10.0 |
| gabor_02 | - | - | - | - |
| bayesian_03 | - | - | - | - |
| log-G_PCA | 3.6 | 4.7 | 1.0 | 0.3 |
| **aguará_06_F** | **11.2** | **12.4** | **1.70** | **3.6** |
| **aguará_06_NF** | **9.0** | **10.0** | **1.6** | **2.7** |

Table 4: EER for different state of the art algorithms.

## 5. Conclusions and Future Work

We presented an EBGM-based system for face recognition. The algorithm uses local information, extracted from representative facial fiducial points, with a set of Gabor filters. We showed a considerable increase in recognition accuracy with respect to the original EBGM by adjusting the transfer function of Gabor filters. This was done in order to compensate the normalized face images PSD which was estimated to fall off as $1/f^3$.

We outperformed the best-known implementations of the EBGM algorithm in the FERET database, and results are comparable with those of the state of the art. Experimental results also indicate that the algorithm performance is robust under different face expressions or non uniform illumination. Finally, it was shown that the refinement step, in the fiducial points location stage, does not offer any advantage in terms of the algorithm performance.

We are currently working on compensating the Gabor functions taking into account not only the dependency of the face images PSD with the frecuency but also with the orientation. We are also investigating other kind of filters that may help to reduce computational costs.

## Acknowledgments

## References

[1] D. S. Bolme. Elastic bunch graph matching. Master's thesis, Colorado State University, Fort Collins, Colorado, USA, 2003.

[2] D. J. Field. Relations between the statistics of natural images and the response profiles of cortical cells. In *Journal of the Optical Society of America.*, volume 4, pages 2379–2394, 1987.

[3] B. Kepenekci, F. B. Tek, and G. B. Akar. Occluded face recognition by using gabor features. In *3rd COST 276 Workshop on Information and Knowledge Management for Integrated Media Communication*, Budapest, October 2002.

[4] M. Kirby and L. Sirovich. Application of the Karhunen-Loève procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1):103–108, 1990.

[5] M. Lades, J. C. Vorbrüggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Würtz, and W. Konen. Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers*, 42:300–311, 1993.

[6] B. Moghaddam, C. Nastar, and A. Pentland. Bayesian face recognition using deformable intensity surfaces. In *Proceedings of the 15th Computer Vision and Pattern Recognition Conference, ICVPR '96*, 1996.

[7] V. Perlibakas. Face recognition using principal component analysis and log-gabor filters. *Image Processing and Analysis Laboratory, Computational Technologies Centre*, 2006.

[8] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss. The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1090–1104, 2000.

[9] S. Rizvi, P. Phillips, and H. Moon. The FERET verification testing protocol for face recognition algorithms. *Image and Vision Computing Journal*, 1999.

[10] M. L. Teixeira. The bayesian intrapersonal/extrapersonal classifier. Master's thesis, Colorado State University, Fort Collins, Colorado, USA, 2003.

[11] M. Turk and A. P. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

[12] A. van der Schaaf. Modelling the power spectra of natural images: Statistics and information. In *Natural image statistics and visual processing*, chapter 2. Graduate School for Behavioural and Cognitive Neurosciences, University of Groningen.

[13] L. Wiskott, J.-M. Fellous, N. Krüger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. In L. C. Jain, U. Halici, I. Hayashi, and S. B. Lee, editors, *Intelligent Biometric Techniques in Fingerprint and Face Recognition*, chapter 11, pages 355–396. CRC Press, 1999.

[14] W. Zhao, R. Chellappa, and A. Krishnaswamy. Discriminant analysis of principal components for face recognition. In *Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition, FG '98*, pages 336–341, Nara, Japan, April 14-16 1998.

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