Biologically-based Face Recognition using Gabor Filters and Log-Polar Images

María José Escobar*, Javier Ruiz-del-Solar**
*Department of Electronics, Universidad Téc. Fed. Santa María, Valparaíso, CHILE
Email: mjescobar@ieee.org
** Department of Electrical Engineering, Universidad de Chile, Santiago, CHILE.
Email: jruizd@cec.uchile.cl

Abstract. A face recognition system that uses Log-Polar Images and Gabor Filters is here proposed. This system models the way in which face images are processed between the retina and the areas V1 and V2 of our visual system. Some simulations of the recognition abilities of the system are also presented.

I. INTRODUCTION

Face Recognition is a very lively and expanding research field. The increasing interest in face recognition is mainly driven by applications like access control to buildings and (computational) systems, identification for law enforcement, borders control, identification and verification for credit cards and ATM, and very recently passive recognition of criminals (e.g. terrorists, hooligans) in public places or buildings (airports, train stations, stadiums, etc.). Many different approaches have been proposed to solve this task. Roughly speaking the most successful approaches can be divided into the ones that analyze the faces in an holistic sense (e.g. eigenfaces and fisherfaces) and the ones that analyze the constitutive parts of the faces (eyes, mouth, nose, etc.) as for example the so-called *dynamic link architecture*.

The dynamic link architecture is a general face recognition technique that represents the faces by projecting them onto an elastic grid where a Gabor filter bank response is measured at each grid node. The recognition of the faces is performed by measuring the similarity of the filter response at each node (Gabor-jets) between different face images. The nodes of the grid normally, but not necessary, correspond to so-called fiducial points (center of the eyes, top of the nose, corners of the lips, etc.). One of the most successful dynamic link architectures is the *Elastic Bunch Graph Matching - EBGM* (see for example [10]).

Gabor analysis, a particular case of *joint spatial/ frequency analysis*, is biologically based, and their oriented filters (Gabor filters) model the kind of visual processing carried out by the *simple* and *complex* cells of the primary visual cortex of higher mammals [6]. The shape of the receptive fields of these cells and their organization are the results of visual unsupervised learning during the development of the visual system in the first few months of life [9]. Based on these facts one can say that the dynamic link approach for face recognition is biologically motivated, at least in the stages where Gabor analysis is employed.

The aim of this work is to take a step further into the construction of a face recognition system that is biologically based. To achieve this objective *the Log-Polar Transformation - LPT* [5], which models the *Retinothopic Mapping* of the visual information between the retina and the area V1 of the visual cortex [6][9], is used. A face recognition system based on the dynamic link architecture, and in which input images are first processed using the LPT is here described (see block diagram in figure 1). The proposed system models the way in which face images are processed between the retina and the areas V1 and V2 of our visual system. The suppositions assumed are that in our visual system faces are stored as Gabor-jets, and that the face identification is performed by measuring the similarity between the Gabor-jets of an input face image and the Gabor-jets of the stored faces.

The use of the LPT in face recognition applications is not new, different research groups have employed it. Tistarelli and Grosso [7][8] have implemented an active face recognition system that uses the LPT together with Principal Component Analysis; Chien and Choi have used the LPT to locate landmarks in faces [1]; Minut *et al.* have used the LPT together with Hidden Markov Models for the recognition of faces [4]. However, to our knowledge our work is the first one where the LPT have been included in a dynamic-link face recognition architecture.

The article is structured as follows. The proposed face recognition architecture is described in section II. In section III some simulations of the recognition abilities of the proposed method using the Yale Face Database are presented, together with a comparison with the EBMG architecture. Finally, in section IV some conclusions and projections of this work are given.

II. PROPOSED ARCHITECTURE

A block diagram of the proposed architecture is presented in figure 1. When a new face image arrives to the system, the *Face Alignment* block finds the *Fp* fiducial points to be used for the recognition (pupils, top of the nose, corners of the lips, etc.). In our current implementation this block is not implemented in an automatic way, but manually. Afterwards the *Log-Polar Transformation* is applied over the input face image using as origin of the transformation the coordinates of a fiducial point (the top of nose in our case). Before applying the *Gabor Filtering* block, the coordinates of the *Fp* fiducial points are transformed into the new log-polar domain (*Log-Polar Coordinate Transformation* block). Thereafter, the log-polar image is convoluted with each of the *Gf* Gabor Filters composing the *Filter Bank*. The convolution is applied in each of the *Fp* fiducial points (log-polar coordinates). The results of the convolution are the *Gabor-jets*. Finally, to determine the identity of the input face, its Gabor Jets are compared with all the Gabor-jets that form the *Face Graphs Database*, using the *Similarity Matching* block.

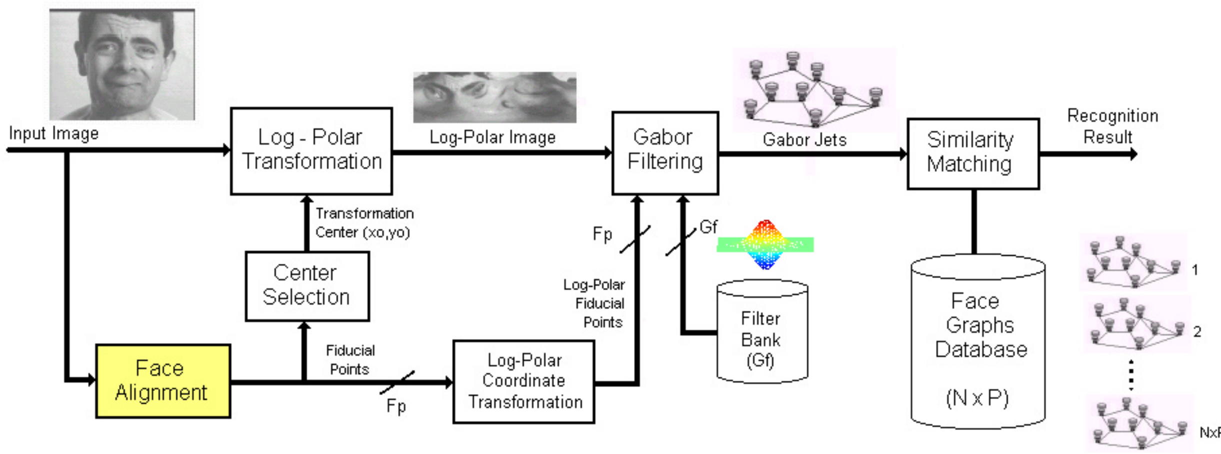

Figure 1. Block diagram of the proposed face recognition architecture.

A. Log-Polar Transformation

Studies of optical nerves and visual image projection in the cerebellar cortex show that the global retinothopic structure of the cortex can be characterized in terms of the geometrical properties of the so-called *retinothopic mapping*. This mapping can be modeled by the Log-Polar Transformation – LPT [6][9].

If $I(x,y)$ is a rectangular image in Cartesian coordinates, then the LPT with origin (x_0, y_0) will be given by:

$$I^*(\rho,\phi) = L\{I(x,y)(x_0,y_0)\} \quad (1)$$

where

$$\rho = M\log(r) \quad (2)$$

$$r = \sqrt{(x-x_0)^2 + (y-y_0)^2}, \quad (3)$$

$$\phi = \tan^{-1}\left(\frac{y-y_0}{x-x_0}\right) \quad (4)$$

The parameter M is determined by the selected height of the transformed image. The angle resolution is given by the selected width of the transformed image. An example of the application of the LPT is shown in figure 2.

The LPT allows a significant reduction of the visual data to be processed. This data diminution is produced by the logarithmic sampling of the input signal in the radial direction and by the constant sampling (the same number of points is taken) in each angular sector to be transformed. Additionally, this mapping provides an invariant representation of the objects, because rotations and scalings of the input signal are transformed into translations [5], which can be easily compensated.

Figure 2. The image on the top corresponds to the original image, while the image on the bottom was obtained applying the LPT. The size of the original image is 128x128 pixels, while the size of the transformed one is 60x219 pixels.

B. Gabor Filtering

Bidimensional Gabor Filters, originally proposed by Daugman [2][3], correspond to a family of bidimensional Gaussian functions modulated by a *cosine* function (real part) and a *sinus* function (imaginary part). These filters are given by ($\vec{x} = (x,y)$):

$$\psi_j(\vec{x}) = \frac{\vec{k}_j^2}{\sigma^2}\exp\left(-\frac{\vec{k}_j^2\vec{x}^2}{2\sigma^2}\right)\left[\exp\left(i\vec{k}_j\vec{x}\right) - \exp\left(-\frac{\sigma^2}{2}\right)\right] \quad (5)$$

$$\vec{k}_j = \begin{pmatrix} k_{jx} \\ k_{jy} \end{pmatrix} = \begin{pmatrix} k_v \cos\varphi_m \\ k_v sen\varphi_m \end{pmatrix}, \qquad k_v = 2^{-\frac{v+2}{2}}\pi, \qquad \varphi_m = m\frac{\pi}{8}$$

with s the standard deviation of the Gaussian, n the frequency parameter, m the orientation parameter.

The second term in the square brackets in (5) is used to avoid the offset, in other words, make that $\int y_j(\vec{x})d^2\vec{x}$ converges. In our implementation 40 filters (5 frequencies and 8 orientations) compose the filter bank. The filter parameters are given by:

n=0, ..., 4 and m=0, ..., 7. (6)

In our current implementation the Gabor filtering is performed using 16 different fiducial points (left pupil, right pupil, top of nose, corners of lips, corners of eyebrows, etc.), which are shown in figure 4. Of course the coordinates of these fiducial points are log-polar transformed before to apply the filtering

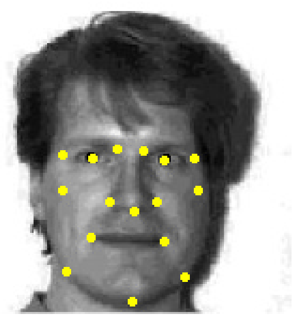

Figure 4. Fiducial Points chosen in a face image

A *Gabor-jet* describes the behavior of the image around a given pixel $\vec{x} = (x, y)$, and it is obtained doing the convolution between the image and each of the 40 Gabor Filters in $\vec{x}$ as shown in equation (7), obtaining a complex number as result. In that way, a *Gabor-jet* is a vector formed by the resulting of the convolution of the image with each of the 40 Gabor Filters and will have 40 components. Each component is given by the magnitude of the complex response of the corresponding filter.

$$J_j(\vec{x}) = \int I(\vec{x}')y_j(\vec{x} - \vec{x}')d^2\vec{x}' \qquad (7)$$

where y_j is one of the Gabor kernel described in (5).

As mentioned, before applying the Gabor Filters, the coordinates of the fiducial points (x_{fi}, y_{fi}) (*i=1, ..., 16*), manually located, must be transformed into the log-polar coordinate system. This transformation is given by:

$$u_{fi} = \frac{s_h}{\ln(r_{max})}\ln\left(\sqrt{(x_{fi} - x_0)^2 + (y_{fi} - y_0)^2}\right)$$

$$v_{fi} = \frac{s_w}{2p}\tan\_sp^{-1}\left(\frac{y_{fi} - y_0}{x_{fi} - x_0}\right) \qquad (8)$$

where (u_{fi}, v_{fi}) are the new log polar coordinates, s_h and s_w are the height and width of the log-polar images, r_{max} is the maximal radius, and the function $\tan\_sp^{-1}$ is like a $\tan^{-1}$ but with range between 0 and 2π.

The face database is build up by applying the LPT and the Gabor Filtering to all training images, and by storing the obtained Gabor-jets in a so-called *Face Graph* structure in the database (see details in [10]).

C. *Similarity Matching*

This block is in charge of the similarity between an array of Gabor-Jets with another one. It compares the Gabor Jets of an input image with the existing ones in the face database. To compare two jets of different images (J and J'), we define the following similarity function:

$$S_a(J, J') = \frac{\sum_{j=1}^{Gf} a_j a'_j}{\sqrt{\sum_{j=1}^{Gf} a_j^2 \sum_{j=1}^{Gf} a'^2_j}} \qquad (9)$$

where a_j, $j = 1,..., G_f$ is the magnitud of the result of realize the convolution between the real and imaginary part of Gabor Filter j and the image.

If now we want to compare all the jets of an image G^T with all the jets of another image G^M, the following expression is used:

$$S_G(G^T, G^M) = \frac{1}{N}\sum_{n=1}^{N} S_a(J_n^T, J_n^M) \qquad (10)$$

III. SIMULATIONS

We performed a comparison between our proposed system (Log-Polar Gabor) and the Elastic Bunch Graph Matching (EBGM). In order to test both methods we made several simulations using the Yale University - Face Image Database. We employed 165 images of 15 different classes. The size of the images was 128 x 128 pixels. In table 1 we show the results of several simulations. For each simulation we used a fixed number of training images (*N*), using the same type of images per class, according with the Yale database specification. In order to obtain representative results we take the average of 20 different set of images for each fixed number of training images. All the images not used for training are used for testing.

In the EBGM implementation 16 fiducial points were used. These points are shown in figure 4. The *Gabor-jets* for each one of these fiducial points were obtained applying 40 Gabor Filters (see definition in (5)), with frequencies and orientations given by (6). After that, the complete graph is compared with the graphs belonging to the database using the similarity function defined by (10).

For the "Log-Polar Gabor" architecture we employed the same 16 fiducial points used for EBMG, but transforming their coordinates using (8). We implemented two recognition systems. In the first one the size of the resulting log-polar was selected to be 60x219 pixels, while in the second one it was selected to be 40x141 pixels.

Table 1. Mean recognition rates using different numbers of training images per class (N), and taking the average of 20 different training sets (small numbers correspond to the standard deviations).

	EBGM	**Log-Polar Gabor**	
N	Normal Image 128x128 pixels	Polar Images of 60x219 pixels	Polar Images of 40x141 pixels
2	74.41% (9.8%)	67.9% (14.0%)	66.7% (15.2%)
3	74.83% (8.4%)	81.4% (7.0%)	78.2% (9.4%)
4	78.52% (8.1%)	86.1% (5.1%)	83.9% (6.8%)
5	81.94% (7.3%)	85.3% (3.4%)	85.72% (5.4%)
6	83.1% (6.7%)	88.7% (3.3%)	88.93% (4.4%)

In Table 1 we can see that the simulation results obtained with each of the two proposed systems are better than the ones obtained with the EBGM, when the number of training images per class is larger than 2. Otherwise EMGM is better. It should be also noted that using the LPT, the number of pixels to be processed with each of the proposed systems is smaller than the number of pixels to be processed with EMGM. In the first case the 80.2% of the pixels must be processed (60x219/128/128), while in the second case just 34.4% (40x141/128/128).

IV. CONCLUSIONS AND PROJECTIONS

A new biologically based approach for face recognition was here presented. Under this approach, the way in which face images are processed between the retina and the primary visual cortex of our visual system, is modeled using the Log-Polar Transformation and Gabor Filtering

Some simulations of the recognition abilities of the proposed architecture using the Yale Face Database were presented, together with a comparison with the EBMG architecture. The simulations results obtained with the proposed architecture are better than the ones obtained using EBGM, when the number of training images per class is larger than 2. The main advantage of using our log-polar approach is the fact that face images before Gabor Filtering have a smaller size, which speeds up the processing.

At the moment we are testing a variant of the proposed architecture, in which a different log-polar image is used for each fiducial point (see block diagram in figure 5). In this case the coordinates of each fiducial point is used as origin to apply the LPT. Thereafter, the Gabor Filters are applied over each obtained log-polar image. The similarity measurement is the standard one (see (10)). Soon we are going to publish the first simulation results of this new architecture.

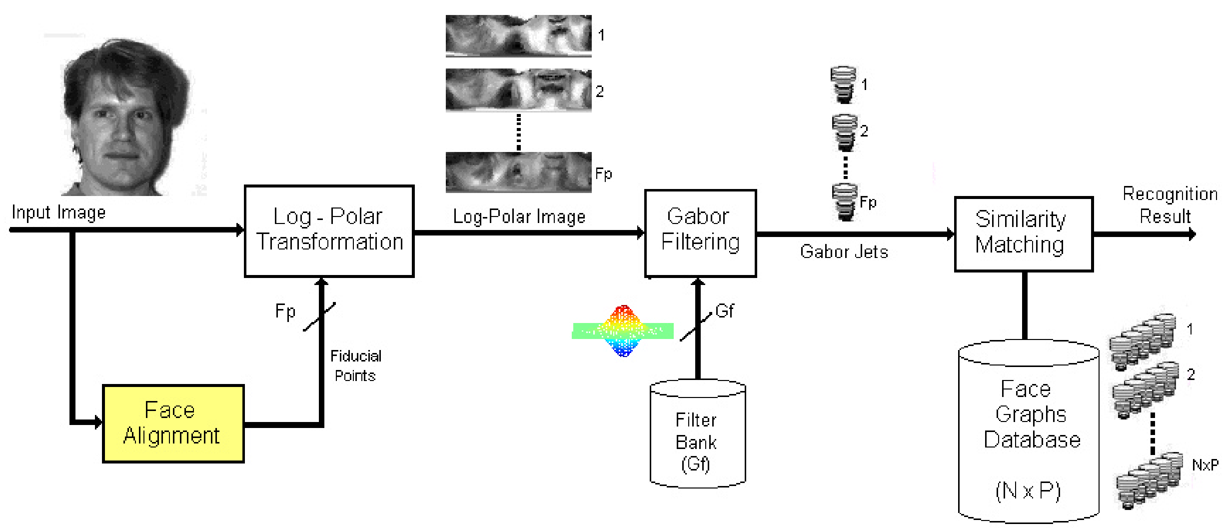

Figure 5. Block diagram of variant of the proposed architecture. In this case different log-polar images are used for each fiducial point.

V. REFERENCES

[1] S. Chien, and I. Choi, "Face and Facial Landmarks location based on Log-Polar Mapping", *Lecture Notes in Computer Science – LNCS* 1811, pp. 379-386, 2000.

[2] J. G. Daugman, "An Information-Theoric View of Analog Representation in Striate Cortex", *Computational Neuroscience*, MIT Press, pp. 403-424, 1990.

[3] J. G. Daugman, "Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 7, pp. 1169-1179, 1997.

[4] S. Minut, S. Mahadevan, J. Henderson, and F. Dyer, "Face Recognition using Foveal Vision", *Lecture Notes in Computer Science – LNCS* 1811, pp. 424-433, 2000.

[5] E.L. Schwartz, Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding, *Vision Research*, 20, pp. 645-669, 1980.

[6] J. Ruiz-del-Solar, *Biologisch basierte Verfahren zur Objekterkennung und Texturanalyse*, Doctoral Degree Thesis, Technical University of Berlin, Germany (ISBN 3-8167-4647-0).

[7] M. Tistarelli, and E. Grosso, "Active vision-based face authentication", *Image and Vision Computing*, no. 18, pp. 299-314, 2000.

[8] M. Tistarelli, and E. Grosso, "Active Vision-based Face Recognition: Issues, Applications and Techniques", in *Face Recognition: From Theory to Applications*, Springer, pp. 262-286, 1998.

[9] H. R. Wilson, D. Levi, L. Maffei, J. Rovamo, and R. DeValois, "The Perception of Form: Retina to Striate Cortex", *Visual Perception: The Neurophisiologcal Foundations*, Academic Press, 1990. /*falta Chap*/

[10] L. Wiskott, J. M. Fellous, N. Krüger, and C. Von der Malsburg, "Face Recognition by Elastic Bunch Graph Matching", *IEEE Trans. Patt. Analysis and Machine Intell.*, vol. 19, no. 7, pp. 775-779, July 1997.