Feature Extraction Using Block-based Local Binary Pattern for Face Recognition

A. Moujahid, A. Abanda, F. Dornaika; University of the Basque Country UPV/EHU; San Sebastin, Spain

Abstract

It is widely assumed that texture is generally characterized locally by two complementary aspects, a pattern and its strength. Based on this assumption and using Local Binary Pattern (LBP) operator as texture descriptor, this work aims to implement an automatic weighting of the local blocks or regions characterizing a given face image. The work reports an improved version of the margin-based iterative search Simba algorithm to feature extraction for face recognition. The main contribution is twofold: (i) we extend the margin-based iterative search algorithm (Simba) to the Chi-square distance that computes dissimilarities between histograms. (ii) since we are interested in studying the relevance of individual blocks or local regions characterizing a given face image, we also extended the Simba algorithm so that one can compute the weights of each attribute as well as of subsets of attributes or blocks. The resulting weight vector has been used initially for an automatic selection of attributes and/or blocks for face recognition with supervised learning based on k-nearest neighbors classifier. Besides, in order to improve the performance of the face recognition task we also made use of the Simba weight vector to weight the distance measures adopted by the k-NN classifier. The experimental results clearly show that the selection based on the automatic weighting outperforms the classification based in all the features. Furthermore, selecting blocks is more effective than selecting attributes, and Chi-square distance performs appreciably better than Euclidean one.

Introduction

Face recognition has received a great deal of attention over the last years because of its many applications in various domains, and it presents a challenging problem in the field of image analysis and computer vision research. Most of the proposed methods for face recognition tasks perform well when the images are in controlled environment but return quite worse results when real environments are considered, where variations of different factors such as illumination and pose are present. In other words, current systems are still far away from the capability of the human perception system.

Biometrics-based authentication systems are becoming a popular option in recent years, changing the authentication based on Personal Identification Numbers, passwords or cards to an authentication based on physiological characteristics. Passwords and cards impose an obligation on the user to remember them or to carry them wherever in the wallet, and moreover, they can be stolen. In the age of comfort, this authentication methods fall short and new biometrics-based systems have started to pop up, systems which do not require the user any effort and in addition, cannot be misplaced, forgotten or stolen. Most common biometrics-based technologies include identification using face, finger geometry, palm, iris, retina or voice, but most of them require some voluntary action by the user, as placing his hand or finger on some machine or standing in a fixed position in front of a camera for its iris identification. Face recognition, on the other hand, can be done without any effort by the user, just acquiring his face image from a distance by a camera.

The applications of face recognition cover many fields from entertainment to law enforcement or surveillance. In the entertainment, it can be used for human-robot interaction or virtual reality or even as Facebook did, for automatic tag suggestion on the photos. In the security field, it has a wide variety of applications: national ID, voter registration, TV Parental control, personal device logon, advanced video surveillance.

Face recognition techniques can be grouped into two main groups: feature-based and holistic methods [16]. Feature-based methods process the face image to extract the relevant features like eyes, mouth or nose and compute the geometrical relationship among them to reduce the original image to a vector of geometrical features. One of those methods is the one carried out by Kanade [17], where a simple image processing is used to extract a vector of 16 parameters of the face image (including size of eyes, distances and angles) and used a simple Euclidean distance measure for matching. Holistic methods, conversely, try to represent a face using global descriptors instead of local feature and within them the most commonly used one is PCA, which was first used by Sirovich and Kirby [18]. In face recognition, PCA is more commonly named as eigenface method, and reduces the original image space to an orthogonal eigenspace with reduced dimensionality.

In this work we follow a holistic approach based on the spatially enhanced Local Binary Pattern (LBP) descriptor. This operator is more effective than the original LBP and it has proven to be highly discriminative allowing a description of the face image at three different levels of locality: (i) the labels of the histogram contain information about the micro-patterns; (ii) the summation of these labels over a small region produces information on a regional level; (iii) the regional histograms could be concatenated to build a global description of the face.

The rest of the paper is organized as follow. In Section 2 we summarize the main idea behind Local Binary Patters approach and its importance in face description. Section 3 introduces the margin-based iterative search algorithm (Simba), and describes the main extensions of Simba. Section 4 reports the experimental results. Finally, conclusions and discussions are drawn in Section 5.

Face Description with enhanced Local Binary Patterns (LBP)

Local Binary Patterns have proved to be a good texture descriptor. The original LBP operator labels the pixels of an image with decimal numbers, which are called LBPs or LBP codes that encode the local structure around each pixel [1, 2]. It proceeds thus, as follow: Each pixel is compared with its eight neighbors in a neighborhood by subtracting the center pixel value; the resulting strictly negative values are encoded with 0,and the others with 1. For each given pixel, a binary number is obtained by concatenating all these binary values in a clockwise direction, which starts from the one of its top-left neighbor. The corresponding decimal value of the generated binary number is then used for labeling the given pixel. The histogram of LBP labels (the frequency of occurrence of each code) calculated over a region or an image can be used as a texture descriptor.

The size of the histogram is 2^P since the operator $LBP(P,R)$ produces 2^P different output values, corresponding to 2^P different binary patterns formed by P pixels in the neighborhood. LBP methodology has been developed recently with large number of variations to improve performance in different applicationsThese variations focus on different aspects of the original LBP operator. In this work we adopted eight points $P=8$ and radius $R=1$.

Given the original LBP operator, the natural next step would be to compute the histograms of the images and to define a distance between them. But as reported in [3], there is a more efficient representation of face images with LBP operator which encodes both the local and the spatial information of facial regions. This representation consists in dividing the image into local regions, extracting texture descriptors from each region independently, and then concatenating them to get the spatially enhanced descriptor of the face image (see Figure 1). The spatially enhanced descriptor has size $m \times n$, where n is the length of a single LBP histogram and m the number of regions composing the face image. Based on the fact that in human face recognition some features like eyes or nose are more important in discriminating faces than others, and tacking into account that each element of the enhanced histogram corresponds to a certain small region of the face, we reports an automatic weighting approach of the enhanced descriptor based on the information each region contains and based on maximizing a given margin. For that, the margin-based iterative search *Simba* algorithm [4, 5] has been considered to compute the weights of the elements composing the histogram-based enhanced descriptor. In our study, the elements can be either the histogram elements (histogram bins) or the block histograms (the entire local histogram).

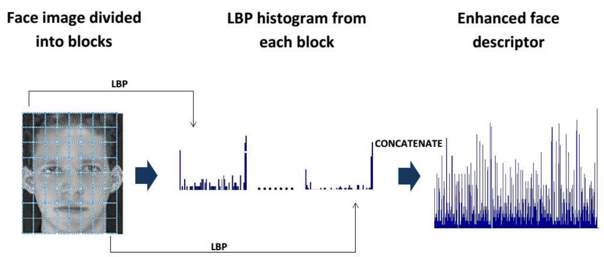

Figure 1: Enhanced Local Binary Pattern descriptor.

Margin-based iterative search algorithm (Simba)

In this section we give a brief description of the Simba algorithm [5]. The main idea is to obtain an effective subset of features such that the relatively significant features have relatively large weights by using hypothesis-margin criterion.

More specifically, the hypothesis-margin of an instance x with respect to a set of points P for the Nearest Neighbor classifier is defined by the following formula:

$$\theta_P(x) = \frac{1}{2}(||x - nearmiss(x)|| - ||x - nearhit(x)||) \tag{1}$$

where $nearhit(x)$ and $nearmiss(x)$ denote the nearest point to x in P with the same and different label, respectively. To choose a subset of features that makes the sum of margins in equation (1) as large as possible, an evaluation function which assigns score to any set of features is required. The hypothesis-margin as a function of the chosen set of features is given by:

$$\theta_P^w(x) = \frac{1}{2}(\parallel x - nearmiss(x) \parallel_w - \parallel x - nearhit(x) \parallel_w) \tag{2}$$

where w is the weight vector of the features and $\parallel x \parallel_w = \sqrt{\sum_i w_i^2 x_i^2}$

Finally, given a training set S and a weight vector w, the evaluation function is defined as follow:

$$e(w) = \sum_{x \in S} \theta_{(S-\{x\})}^w(x) \tag{3}$$

It is natural to look at the evaluation function only for weight vectors w such that $\max w_i^2 = 1$. In order to effectively compute the optimal weight vector, the so-called gradient ascent strategy [4] was used for maximizing $e(w)$. The gradient of $e(w)$ when evaluated on a sample x is given by

$$(\bigtriangledown e(w))_i = \frac{1}{2}\sum_{x \in S}\left(\frac{(x_i - nearmiss(x)_i)^2}{\parallel x - nearmiss(x) \parallel_w} - \frac{(x_i - nearhit(x)_i)^2}{\parallel x - nearhit(x) \parallel_w}\right)w_i \tag{4}$$

The iterative search algorithm is given by Algorithm 1 where N denotes the dimension of the feature vector x.

Algorithm 1 Simba

Initialize $w = (1, 1, ..., 1)$
for $t = 1, ..., T$ **do**
Pick randomly a sample x from S
Calculate $nearmiss(x)$ and $nearhit(x)$ with respect to $(S - \{x\})$ and the weights w
for $i = 1, ..., N$
$$\Delta_i = \frac{1}{2}\left(\frac{(x_i - nearmiss(x)_i)^2}{\|x - nearmiss(x)\|_w} - \frac{(x_i - nearhit(x)_i)^2}{\|x - nearhit(x)\|_w}\right)w_i$$
End
$w = w + \Delta$
End
$w = w^2 / \parallel w^2 \parallel_\infty$

To compute both $nearmiss(x)$ and $nearhit(x)$, and the associated norms $||x - nearmiss(x)||_w$ and $||x - nearhit(x)||_w$, the original Simba algorithm uses the weighted Euclidean distance, that is, $d_w(x,y) = \sqrt{\sum_i w_i^2 (x_i - y_i)^2}$.

But, it is well known that χ^2 distance is more appropriate to identify dissimilarities between histograms. So, in this work, we have adopted a modified version of the Simba algorithm that includes χ^2 distance. The weighted χ^2 distance between x and y is given by

$$\chi_w^2(x,y) = \sum_i w_i \frac{(x_i - y_i)^2}{x_i + y_i}.$$

This distance is used in the new Simba to calculate $nearmiss(x)$, $nearhit(x)$ and the norms $||x - nearmiss(x)||_w$, $||x - nearhit(x)||_w$. But, in addition to that, the increment Δ_i inside the algorithm also changes, since it is based on the hypothesis margin which depends on the distance. The adapted version of the Simba algorithm based on χ^2 distance is given by Algorithm 2.

Algorithm 2 The adapted version of Simba based on χ^2 distance

Initialize $w = (1, 1, ..., 1)$
for $t = 1, ..., T$ **do**
  Pick randomly a sample x from S
  Calculate $nearmiss(x)$ and $nearhit(x)$ with respect to $(S - \{x\})$ and the weights w
  for $i = 1, ..., N$
    $\Delta_i = \frac{1}{2}\left[\chi^2(x_i, nearmiss(x)_i) - \chi^2(x_i, nearhit(x)_i)\right]$
  End
  $w = w + \Delta$
End
$w = w^2 / ||w^2||_\infty$

where N is the length of a sample or number of attributes of a sample.

The resulting Simba weight vector has size given by the number of attributes characterizing a given sample.We point out that in this work a sample is an enhanced LBP histogram, which is a concatenation of the individual histograms describing the different regions composing the face image. Therefore, to compute the weights of these regions based on the importance of the information they contain, the Simba algorithm is also extended to include block-based distances. We have considered both Euclidean and χ^2 block-based distances.

Given two samples x and y, which are divided into m regions with n attributes per region. Given also a region-weight vector w of size m, the **block-weighted Euclidean** and **block-weighted χ^2** distances between x and y are given respectively by

$$d_w(x,y) = \sqrt{\sum_{j=1}^{m} \sum_{i=1}^{n} w_j^2 (x_{i,j} - y_{i,j})^2}$$

and

$$\chi_w^2(x,y) = \sum_{j=1}^{m} \sum_{i=1}^{n} w_j \frac{(x_{i,j} - y_{i,j})^2}{x_{i,j} + y_{i,j}}.$$

Experimental Results

Dataset

To perform the proposed framework we have considered the Yale face dataset [13] which is characterized by lighting variations and different facial expressions. The Yale dataset was constructed at the Yale Center for the Computational Vision and Control. It contains 165 gray-scale images of 15 subjects (11 images per subject), under various facial expressions or configuration, specifically: center-light, with glasses, happy, left-light, without glasses, normal, right-light, sad, sleepy, surprised and winking. In addition, the faces are not centred. The size of the images is 320x243 pixels, with 256 grey levels per pixel and they are in GIF format.

Experimental setup

In this work, the enhanced LBP descriptor has been built as follows: we have divided the face image into different regions, and for each one we have computed the individual LBP histogram. The spatially enhanced histogram is then computed by concatenating all the individual histograms. For example, a division of the face image into 8×8 blocks (64 in all) leads to an individual LBP histogram with size 59, and to an enhanced descriptor with size of 59*64 = 3776. Different partitions of the input images into blocks of different sizes have been analyzed.

Two main configurations have been considered to implement the Simba algorithms: (i) Simba by attributes, which means that we compute the weight vectors in order to characterize the individual elements of the enhanced LBP descriptor; and (ii) Simba by blocks, where the weight vectors are assigned to blocks or subsets of feature of the enhanced LBP descriptor. In both configurations Euclidean and Chi-square distances have been considered to compute dissimilarities between histograms.

On the other hand, we have considered different uses of the weight vectors obtained from the Simba algorithms. First, we have performed feature selection to identify the relevant regions or blocks composing the face images. The protocol followed is depicted in Figure 2. The original dataset is randomly divided into training and test sets. For each split, we compute the Simba weight vector to score each feature providing both a ranking of the elements composing the enhanced histogram or the blocks representing individual regions of the face image. From this obtained ordering, several feature subsets can be chosen by setting a cutoff for the selected features (attributes or blocks). In this work we have adopted threshold-based criterion. In fact, we have analyzed different cutoff values ranging from 10% to 90% of the features. Once the selection is fixed; it is applied on both the training and test sets. Finally, we evaluate the recognition rate on the test set using the selected features. Since we use several random splits, the final rate is given by the average over all used splits.

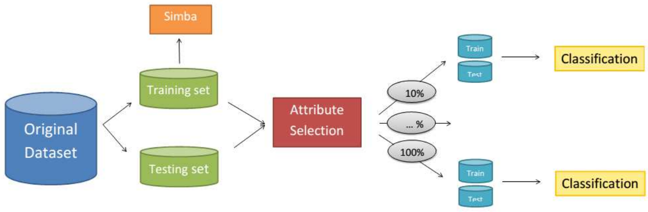

Figure 2: The evaluation protocol followed to carry out feature selection and classification tasks.

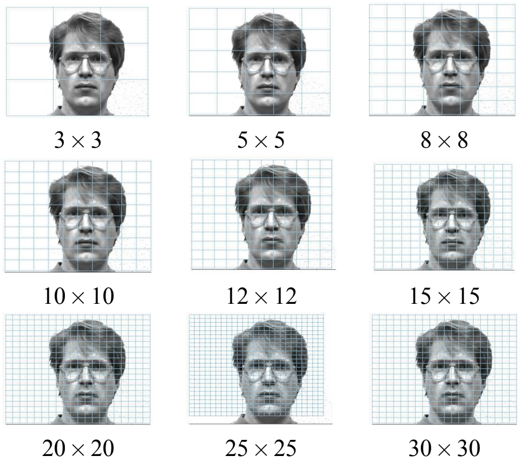

Figure 3: Partitions of the face images into blocks of different sizes.

Performance evaluation

As described before the enhanced LBP descriptor has been computed for partitions of the input images into blocks of different sizes. The blocks size ranges from 3×3 to 30×30 (See Figure 3) . The mean accuracy achieved for the different partitions using k-NN classifiers with both Euclidean and Chi-square distances is reported in Figure 4. We can observe that the accuracy increases as the number of blocks increases for almost all considered cases. The 1-NN classifier based on Chi-square distance gives better accuracy. The highest accuracy is obtained with a partition of 25 blocks.

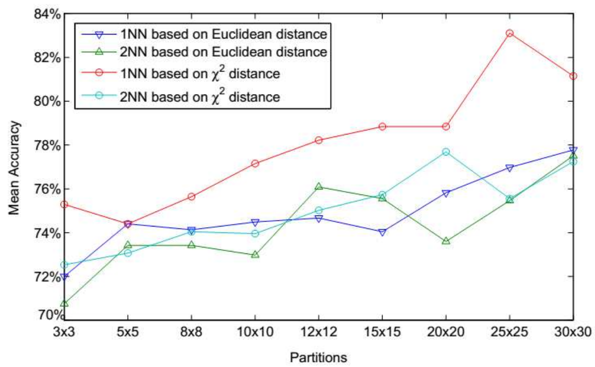

Figure 4: The mean accuracy corresponding to different partitions of the face images into blocks whose number ranges from 3×3 to 30×30.

Once the optimal number of blocks has been fixed, the first group of experiments performed feature selection on the attributes making use of the Simba weight vector considering both Euclidean and Chi-square distances. That is, we compute Simba weight vectors to score the attributes of the enhanced LBP descriptor, then we perform subset selection based on percentiles of order k as threshold values. Figures 5 and 6 show the accuracy of two individual splits and the mean accuracy over 20 splits of a 1-NN classifier as a function of the number of selected attributes. As it can be appreciated, higher accuracy have been achieved by few relevant attributes. For Euclidean Simba the optimal subset corresponds to the percentile 30th, while for Chi-square Simba are required only 10% of the relevant attributes. These results have been improved weighting the distance measures adopted by the k-NN classifiers with the Simba weight vectors.

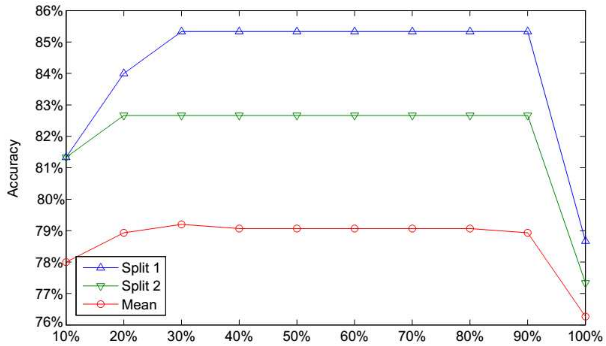

Figure 5: The accuracy of two individual splits and the average accuracy corresponding to 20 random splits training/test as a function of selected attribute by Euclidean-distance-based Simba. The classification has been achieved by 1-NN based on Euclidean distance.

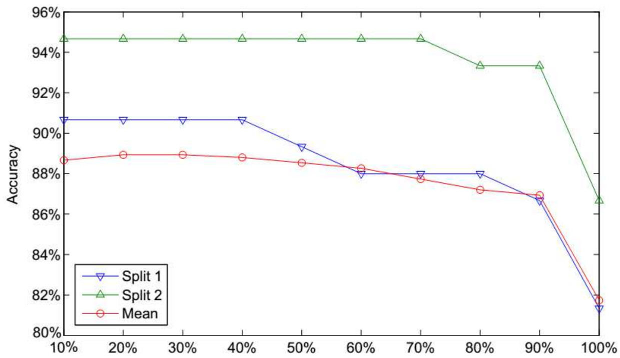

Figure 6: The accuracy of two individual splits and the average accuracy corresponding to 20 random splits training/test as a function of selected attribute by χ^2-distance-based Simba. The classification has been achieved by 1-NN based on χ^2 distance.

In the second group of experiments, we have implemented the Simba algorithm by blocks. This allows a selection of relevant blocks or local regions of the face image. As in the previous phase, the Simba weight vector has been computed using the training set, then we performed the classification using 1-NN classifier for different numbers of selected blocks. Results of accuracy corresponding to Simba by blocks based on Chi-square distance are depicted in figure 8. The inset panel shows the shape of the weight vector. The relevant blocks are those shifted to the

Table 1: Average accuracy (%) corresponding to 20 random splits training/test achieved when considering respectively all the blocks and 20% of relevant blocks. The results are obtained with 1-NN classifier.

	Chi-square	Euclidean
All block	80.53	74.27
20% of relevant blocks	94.04	81.87

red color and correspond to hair, front and neck. For a better appreciation, we represented the location of these blocks in the face image for Euclidean and Chi-square distances (See figure 7). As it can appreciated, higher values of the accuracy are achieved when only 10% of relevant blocks are selected. The maximum improvement reached is of about 15% compared with the accuracy when all the blocks are selected. Table 1 summarizes these results.

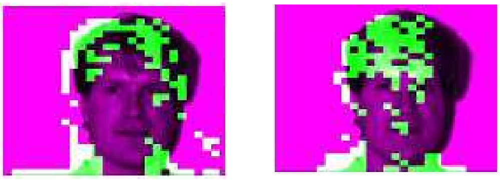

Figure 7: Blocks selection based on the Simba weight vectors. (Left) Simba based on Euclidean distance. (Right) Simba based on Chi-square distance. Only 20% of relevant blocks are represented.

Conclusions

In this work, an updated version of the margin-based iterative search algorithm has been used to feature extraction for face recognition. Firstly, an enhanced block-based face descriptor is used to represent face images. The images are divided into blocks, and the individual LBP histograms of these blocks are computed and then concatenated to get the enhanced descriptor. In human perception some features of faces play more relevant role in recognizing a face than others, for example eyes or mouth. Thus, it is expected that the same will happens in artificial perception. Under this assumption and since the notion of block is an inherent feature of the LBP descriptor, the main goal of this work is to implement an automatic weighting of these blocks. The main contribution is twofold: (i) we have implemented an updated version of the Simba algorithm that include Chi square distance to compute dissimilarities between histograms. (ii) since we are interested in studying the relevance of individual blocks or subsets of attribute to characterize a given face image, we have extended the Simba algorithm so that one can compute the weights of each attribute as well as for subsets of attributes or blocks. The resulting weight vector has been used initially for an automatic selection of attributes and/or blocks for face recognition with supervised learning based on k-nearest neighbors classifier. But, to improve the performance of the face recognition task we also made use of the Simba weight vector to weight the distance measures adopted by the k-NN classifier.

To determine the optimal size of blocks, we have analyzed the performance of k-NN classifier for different partitions of the images into blocks. The results show that for the Yale dataset, a division of the face images into 25×25 blocks seems to be the optimal partition. Once the number of blocks has been fixed, we performed attribute as well as block selection. the experimental results clearly show that the selection based on the automatic weighting outperforms the classification based on all the features. Furthermore, selecting blocks is more effective than selecting attributes, and Chi-square distance performs appreciably better than Euclidean one.

Future works may investigate the use of others distance measures such as Jeffery or Kolmogorov-Smirnov distances, and to improve the quality of the LBP enhanced descriptor others modes and radii may be considered.

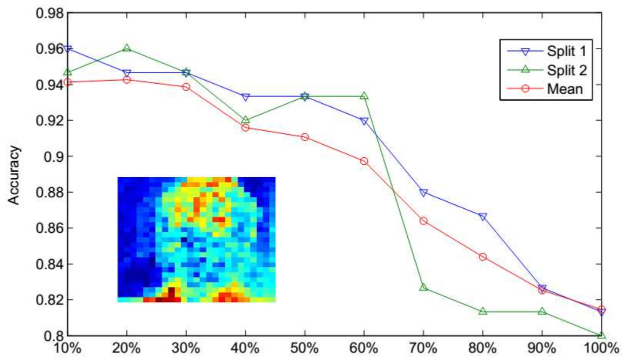

Figure 8: The accuracy of two individual splits and the mean accuracy as a function of selected blocks by χ^2-distance-based Simba. The classification corresponds to 1-NN based on χ^2 distance. The inset shows the shape of the block weight vector.

References

[1] Ojala T., Pietikinen M., Harwood D.: *A comparative study of texture measures with classification based on feature distributions*. Pattern Recognition 29 (1996) 51–59.

[2] Ojala T., Pietikinen M., Menp T.: *Multiresolution gray-scale and rotation invariant texture classification with local binary patterns*. IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (2002) 971–987.

[3] Ahonen T., Hadid A. , Pietikinen M.: *Face Recognition with Local Binary Patterns*. Computer Vision Using Local Binary Patterns, Computational Imaging and Vision (2004) 40.

[4] Yang M., Wang F., Yang P.: *A novel feature selection algorithm based on hypothesis-margin*. Journal Of Computers (2008) vol. 3, NO. 12

[5] Gilad-Bachrach R., Navot A., Tishby, N.: *Margin Based Feature Selection-Theory and Algorithms*. In Proc. of the 21st ICML, Banff, Canada, (2004) 43–50.

[6] Crammer K., Gilad-Bachrach R., Navot A., Tishby N.: *Margin analysis of the lvq algorithm*. Proc. of 17th CNIPS, (2002).

[7] Vapnik V.. *The nature of statistical learning theory*. New York, Springer-Verlag, (1995).

[8] Phillips P., Grother P., Micheals R.J., Blackburn D.M., Tabassi E., Bone J.M.: *Face recognition vendor test 2002 results*. Technical report (2003).

[9] Zhao W., Chellappa R., Rosenfeld A., Phillips P.J.: *Face recognition: a literature survey*. Technical Report CAR-TR-948, Center for Automation Research, University of Maryland (2002).

[10] Fukunaga K.: *Introduction of Statistical Pattern Recognition*. Second ed. Academic Press (1991).

[11] Guyon I., Elisseeff, A.: *An Introduction to Variable and Feature Selection*. JMLR, 2003(3): 1157–1182.

[12] Rodriguez Y., Marcel S.: *Face Authentication Using Adapted Local*

Binary Pattern Histograms. Proc. Ninth European Conf. Computer Vision, IV: (2006) 321–332.

[13] Belhumeur P. N., Hespanha J. P., Kriegman D. J.: *Eigenfaces vs. fisherfaces: recognition using class specific linear projection*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):711720, July (1997)

[14] Yang B, Songcan C.: *A comparative study on local binary pattern (LBP) based face recognition: LBP histogram versus LBP image*. Neurocomputing 120 (2013) 365–379.

[15] Jafri* R., Arabnia H.: *A Survey of Face Recognition Techniques*. Journal of Information Processing Systems, Vol.5, No.2, June (2009) 41

[16] Heisele B., Ho P., Wu J., Poggio T.: *Face recognition: component-based versus global approaches*. Computer Vision and Image Understanding, Vol.91, (2003) 6–21

[17] Kanade T.: *Picture Processing System by Computer Complex and Recognition of Human Faces*. Kyoto University, Japan, PhD. Thesis (1973).

[18] Sirovich L., Kirby M.: *Low-dimensional Procedure for the Characterization of Human Faces*. Journal of the Optical Society of America A: Optics, Image Science, and Vision, Vol.4, (1987) 519–524

[19] He X., Cai D., Niyogi P.: *Laplacian Score for Feature Selection*. Advances in Neural Information Processing Systems 18, Vancouver, Canada (2005)

Author Biography

F. Dornaika received his Master degree in Electrical Engineering from the Lebanese University, in 1990, an M.S. degree in signal, image and speech processing from Grenoble Institute of Technology, France, in 1992, and a Ph.D. degree in computer science from Grenoble Institute of Technology and INRIA, in 1995. Currently, he is a Research Professor at IKERBASQUE (Basque Foundation for Science) and the University of the Basque Country (UPV/EHU). Prior to joining IKERBASQUE, he held numerous research positions at several research centers and universities in Europe, China, and Canada. His research covers a broad spectrum in computer vision and pattern recognition. He has published more than 190 papers in the field of computer vision and pattern recognition. His current research includes pattern recognition, machine learning and data mining.

A. Moujahid received his BS in physics from the University of Rabat (Morocco) (1995) and his PhD in Computer Science from Basque Country University UPV/EHU (2005). Since then he has worked in the Department of Computer Science and Artificial Intelligence at the University of the Basque Country. His work has focused on topics that range from machine learning and data mining techniques to mathematical modelling and simulation of biological dynamical systems.

A. Abanda received her BS in mathematics from the Autonomous University of Barcelona (Spain) (2013), during which she stayed in the University of Granada for a period of one academic year. She then worked as software developer while studying the MSc of Computational Engineering and Intelligent Systems at the University of the Basque Country (2015). Her interests include applied mathematics, computer vision and pattern recognition.