A New Local Binary Probabilistic Pattern (LBPP) and Subspace methods for face recognition

[1]*Abdellatif Dahmouni, [2]Karim El Moutaouakil, and [1]Khalid Satori
[1]LIIAN, Department of Mathematics and computer science
Faculty of Sciences, Dhar-Mahraz Sidi Mohamed Ben Abdellah
B.P 1796 Atlas-Fez
[2]INSAH, National School of Applied Sciences Al Hoceima
BP 03, Ajdir Al-Hoceima University Mohamed 1
Morocco
abdellatifdahmouni@gmail.com
karimmoutaouakil@yahoo.fr
khalidsatorim3i@yahoo.fr

Abstract: - In this paper, we present a new model of extraction of local characteristics, named Local Binary Probabilistic Pattern (LBPP), and based on the Local Binary Pattern (LBP). This model relates to a very important result of probability theory; it is the large great numbers. In this respect, the distribution of the gray levels on the areas (homogeneous texture) on a face image follows a law of probability, which is the sum of several normal laws. This vision allows evaluating the current pixel while basing on the concept of confidence interval, which permits to overcome some LBP shortcomings, especially the information losses associated to LBP deterministic nature. We have combined the proposed model with the most known algorithms of the dimensionality reduction of data analysis in the face recognition field. In order to evaluate our approach, various experiments are carried out on data bases ORL and YALE. In this context, we made a comparison between the performances of systems LBP+ACP, LBP+LDA, LBP+2DPCA, LBP+2DLDA, LBPP+ACP, LBPP+LDA, LBPP+2DPCA and LBPP+2DLDA. The obtained results show the effectiveness of our systems, in particular for systems LBPP+2DPCA and LBPP+2DLDA. The experimental exactitude observed is of 96.5%.

Key-Words: - LBP, LBPP, PCA, LDA, 2DPCA, 2DLDA, Confidence interval

1 Introduction

The face recognition is a biometric modality which involves several fields of research; it covers many applications such as biometric systems, computer security systems, access control systems, and surveillance systems [24]. Many approaches are developed for face recognition: Principal Components Analysis PCA and 2DPCA [1][2], Linear Discriminate Analysis LDA and 2DLDA [3][4], Independent Component Analysis ICA [5], Artificial Neural Networks ANN [6][22][25], Hidden Markov Models HMM [7][23], Gabor Wavelets [8], and Local Binary Pattern LBP [11]. The fusion between the dimensionality reduction methods and the local characteristics extraction methods is a key stage for face recognition. In this work, we focus on fusion between the PCA, LDA, 2DPCA, and 2DLDA methods, and some extended methods of local binary pattern LBP.

The local binary pattern proposed in [9] is a non parametric operator for texture analysis, and extended for face recognition [11], by describing the face like the primitives composition called micro-patterns; and for other implementations like: face detection [17], facial expression [18], age evaluation [19], Gender classification [20], and movement analysis [21]...

The original LBP uses a neighborhood of (3*3) pixels, centered on current pixel to calculate its gray level value. Each of the eight neighboring pixels has higher or equal value to the current pixel value is coded by « 1 », the others by « 0 ». The process is depicted in Fig 1. Where a byte of eight bits is generated, and then converted into gray level (value between 0 and 255) by the following mathematical equation:

$$LBP_{P,R}(x_c, y_c) = \sum_{n=1}^{P} s(i_n - i_c) * 2^{n-1} \quad (1)$$

Where: $s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$

The parameters R and P are respectively the neighborhood radius and the number of neighboring pixels around a current pixel.

The parameters i_c and i_n represent the gray level values of the current pixel and the neighboring pixels.

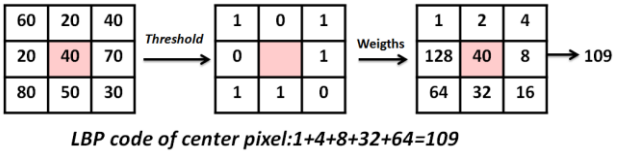


Fig.1 The method of LBP code Calculation

Because of its invariance to the monotonous lighting changes, and its low calculation complexity, the LBP became largely used. Several extensions of LBP are developed:

Ojala proposes a circular neighborhood (illustrated in Fig 2) to solve the multi-resolution problems; he noticed that more than 90% of the texture information present at most two transitions (01 or 10); for example, 00111111 and 10000011 are uniform patterns. It's the uniform variety $LBPU^2$, which reduces the range of the histogram values to 59 distinct values [10].

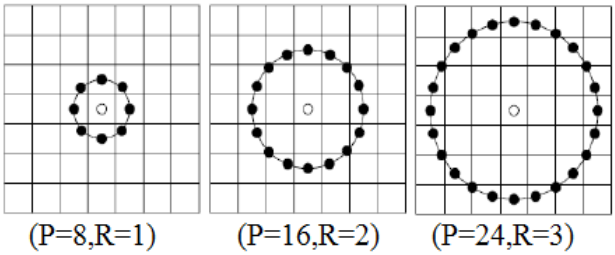


Fig.2 Different circular neighborhoods

To overcome the noise consequences Tan proposes the Local Ternary Pattern LTP, by using a mode of three bits coding [13]. Ahonen introduces the Soft-LBP by using fuzzy logic [12], Liao proposes the multi-blocks MB-LBP, he applies the LBP descriptor to the rectangular blocks intensities that consists to substitute the pixels of each blocks by their average pixel [14]. Jabid proposes the Local Directional Pattern LDP, which uses the Kirsch masks as detectors of the maximum information directions [15]. Bongjin uses the gradient values to correct the local intensity variations produce along the edge components, by introducing the Local Gradient Pattern LGP [16].

The different varieties of LBP suffer from the loss of information related to the hard thresholding mode for evaluating a current pixel. What affects the representation of principal components which characterize a face (eyes, nose, mouth…).

In this work, we present a new extension of the local binary pattern LBP, called local binary probabilistic pattern LBPP. This model leads to a very important result of probability theory; it's the large of great numbers. In this context, the distribution of the gray levels on a face image follows a law of probability, which is the sum of several normal laws. This vision allows evaluating the current pixel while basing itself on the confidence interval concept.

Initially, the proposed model is tested with the spatially enhanced local binary pattern histogram (eLBPH) presented by Ahonen [11][12], and combined, in second time, with the most known algorithms of dimensionality reduction PCA, LDA, 2DPCA and 2DLDA, for face recognition on images extracted from the databases ORL and YALE.

Also, we compare the performances of the systems LBP+ (eLBPH, ACP, LDA, 2DPCA and 2DLDA) and LBPP+ (eLBPH, ACP, LDA, 2DPCA and 2DLDA). The obtained results show the effectiveness of the proposed approach for face recognition compared to some existing techniques, especially for systems LBPP+2DPCA and LBPP+2DLDA. The observed experimental exactitude is of 96.5%.

The Following of this paper consists of four sections followed by a conclusion. The section 2 presents in detail the LBPP descriptor variety based on the confidence interval. The face recognition systems based on LBPP are examined in the section 3. The section 4 presents the experimental results which evaluate the effectiveness of the suggested approach.

2 Proposed Method: LBPP

As many distributions observed in reality, the distribution in levels of gray in a face follows approximately a sum of several normal laws. Recall that the normal law $N(\mu,\sigma)$ is defined by the probability density f and the distribution function F given by:

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} \quad \forall t \in R \qquad (2)$$

$$F(x) = \int_{-\infty}^{x} f(t)dt \quad \forall x \in R \qquad (3)$$

It should be noted this kind of law describes phenomena which result from the addition of a large number of independent fluctuations. For areas, almost homogeneous, exist small variations of intensity between neighboring pixels, we can remove these intensities variations by assimilating

the distribution of the pixels in these areas to a normal law. Therefore, we have from the mode of fixed and signed thresholding for LBP [9][11], absolute and locally adapted for LGP [16], to a mode of thresholding per confidence interval depending on the distribution of the pixels in a neighborhood of a current pixel; see Fig.3 and Fig.4.

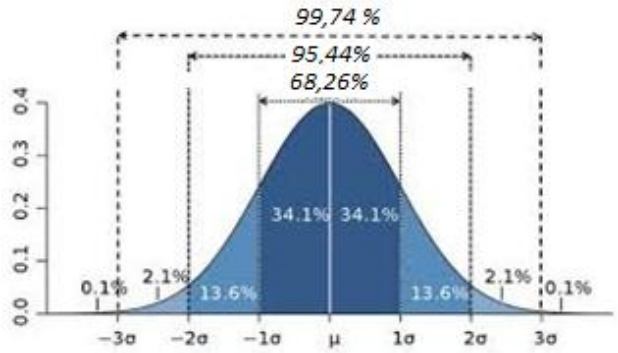


Fig.3 Probability density of various confidence intervals

We propose to generate a confidence interval $[\alpha_1, \alpha_2]$, based on statistical moments in a neighborhood of $N=p^2$ pixels centered on current pixel. Each of the eight neighboring pixels having a value located inside the confidence interval is coded by 1, the others by 0. A byte of eight bits is generated, and converted into gray level.

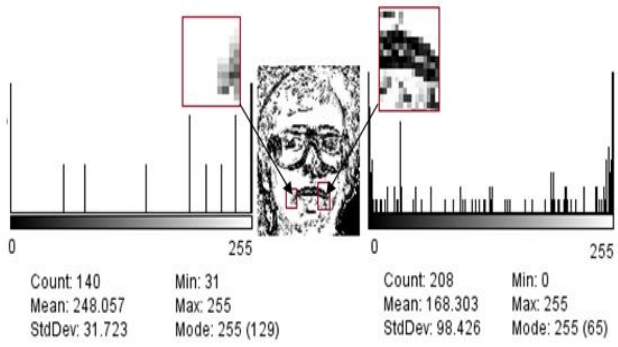


Fig.4 Histograms of various areas of LBPP image

2.1 Confidence interval evaluation

To evaluate the confidence interval, a random variable X is associated to the pixel gray level value. The following statistical moments are calculated:

Average: $$\mu = \frac{1}{N}\sum_{i=1}^{N} n_i \quad (4)$$

Standard deviation: $$\sigma = \sqrt{\frac{1}{N}\sum_{1}^{N}(n_i - \mu)^2} \quad (5)$$

Variation coefficient: $$\delta = \frac{\sigma}{\mu} \quad (6)$$

Skewness coefficient: $$S = \frac{1}{N}\sum_{1}^{N}(n_i - \mu)^3/\sigma^3 \quad (7)$$

The normal law is characterized by the symmetry coefficient null. It appears natural to calculate this indicator to measure the rapprochement of an empirical distribution $D(\mu,\sigma)$ with the normal law $N(\mu,\sigma)$. The evaluation of the fluctuations related to data dispersion is ensured by constraints applied on the variation coefficient δ.

Let us suppose that any variable X fulfilling the condition:

$$(S = 0 \; or \; \delta < \beta \quad with \quad 0.1 < \beta < 0.2) \quad (8)$$

Where, β is experimentally determined (β depends of databases ($\beta_{orl} = 0{,}1$); follows the normal law $N(\mu,\sigma)$, it is a homogeneous neighborhood in which the probability that X is inside the confidence interval $[\alpha_1, \alpha_2]$, is given by:

$$P(\alpha_1 \le X \le \alpha_2) = \frac{1}{\sigma\sqrt{2\pi}}\int_{\mu-k\sigma}^{\mu+k\sigma} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt \; \forall k \epsilon \mathbf{R}$$

$$Where: \quad (9)$$

$$\alpha_1 = \mu - K \; and \; \alpha_2 = \mu + K \quad with \; K = k\sigma$$

The normal variable X is almost certainly located in the interval: $[\mu - 3\sigma, \mu + 3\sigma]$; we define a convergence limit of the normality in the confidence degree to 99.99% for the interval: $[\mu - 4\sigma, \mu + 4\sigma]$ therefore $P(\mu - 4\sigma \le X \le \mu + 4\sigma) \approx 1$. See, Fig.3. For any variable X not fulfilling (8) one takes:

$$\alpha_1 = \mu - K \; and \; \alpha_2 = \mu + K$$

$$Where: \; K = \begin{cases} \sigma & if \; \beta < \delta < 0{,}2 \\ 0{,}2 * \mu & others \end{cases} \quad (10)$$

Such choices permit to define a limit to radius of data convergence, and correct the asymmetry and kurtosis problems. See, Fig.5.

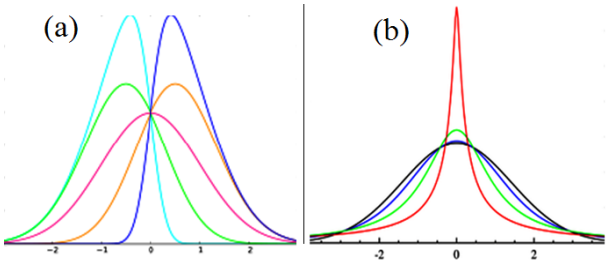


Fig.5 Asymmetry (a) and Kurtosis (b) fluctuations

2.2 LBPP mathematical formulation

The LBPP approach based on the confidence interval defined in (9) and (10) is given by the following mathematical formulation:

$$LBPP_{P,R,P',R'}(x_c, y_c) = \sum_{n=1}^{P} s(i_n) * 2^{n-1} \quad (11)$$

Where:

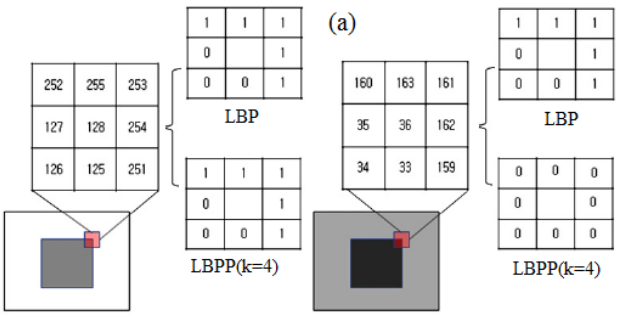
$$s(x) = \begin{cases} 1, if\ \alpha_1(\mu, \sigma, K) \le x \le \alpha_2(\mu, \sigma, K) \\ 0, others \end{cases} \quad (12)$$

The parameter K is defined in (9) and (10).

The parameters R and P are respectively the neighborhood radius and the number of neighboring pixels that contribute to binary code calculation.

The parameters R' and P' are respectively the neighborhood radius and the number of neighboring pixels that contribute to the statistical moments: (μ, σ, δ and S) calculation.

Contrary to LBP, which is a deterministic model, our approach considers the face as distributions of pixels which follow a law of probability. We thus pass to a probabilistic and statistical description where the notion of the confidence interval allows assigning values near 255 at all pixels of almost homogeneous areas (forehead, cheeks….). And the values near 0 at all corners. The nose, eyes and mouth having of strong curves, will have different values, will be easily located Fig.4.The LBPP is robust to local and global intensity variations. However, in Fig 7(a) where the background and the foreground together change (global), LBP, and LBPP codes have the same pattern. By against, when the background and the foreground separately change (local), the LBPP code remains invariant, but LBP code changes Fig 7((b) and (c)).



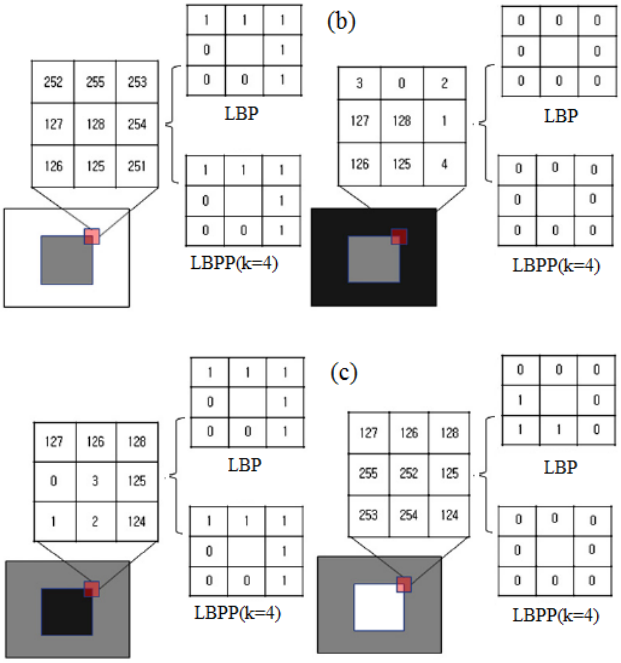


Fig.6 LBP and LBPP for different changes of the gray levels intensity: (a) Monotonic change, (b) Background change, and (c) Foreground change.

3 FACE RECOGNITION SYSTEMS

Our approach is tested on different systems. Initially, we use the histograms information to establish the feature vector. Afterwards, we use several dimensionality reduction methods of face recognition.

3.1 Building eLBPPH Feature Vector

Generally, the face recognition approaches based on LBP descriptor, use the enhanced local binary pattern histogram (eLBPH) proposed by Ahonen.

To build the eLBPPH feature, the facial image is divided into R1...Rd not-overlapped rectangular areas, the histograms of these areas are individually calculated, and then are together concatenated to establish a global face description. After, a classifier based on the histogram matching techniques such as the popular Chi square statistic ($\chi 2$) metric is used for computation feature vector similarity. See, Fig.7.

$$\chi^2_{\omega} = \sum_{i,j} \omega_j \frac{(S_{ij} - M_{ij})^2}{S_{ij} + M_{ij}} \quad (13)$$

In which S_{ij} and M_{ij} are images histograms, and ω_j is weight for region j.

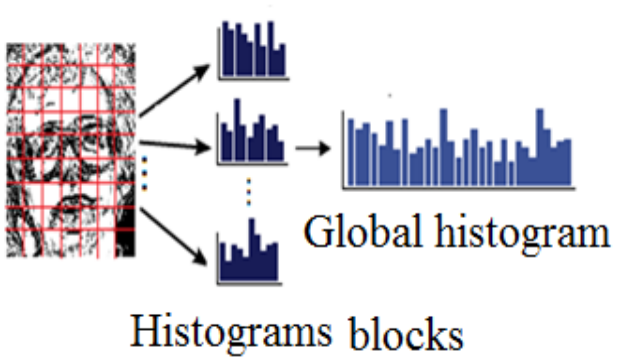


Fig.7 The eLBPPH feature vector extracting

For the suggested approach, the analysis of the histograms of images on ORL and YALE databases; show that the histogram values are concentrated in the interval [0-15; 240-255].What reduces the eLBPPH feature vector. The length of the eLBPPH is $32*d_n^2$, instead of $59*d_n^2$ of the uniform LBP, and $256*d_n^2$of the original LBP. Where, d_n^2 represent the size of each block. See, Fig.8

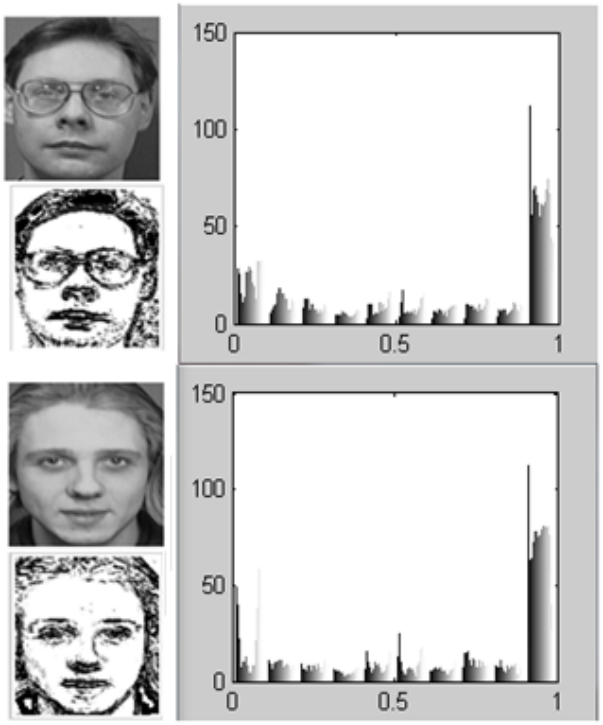


Fig.8 The histograms for LBPP(k=4) images

3.2 Methods of dimensionality reduction

To profit from the advantages of the local and global approaches. The LBPP descriptor is combined with different dimensionality reduction algorithms. The dimensionality reduction algorithms consist in determining an optimal orthogonal basis. This basis is formed by Eigen vectors of the covariance matrix constructed starting from the training images. In contrast, the conventional algorithms PCA and LDA which transform each image by a column vector; 2DPCA and 2DLDA algorithms employ directly the two-dimensional data to form the covariance matrix. The size of this matrix is (n,n) instead of (n^2,n^2).

Consequently, 2DPCA and 2DLDA have significant advantages compared to PCA and LDA. First of all, the face images are less distorted; the image details are better preserved. In addition it is easier to evaluate with precision the covariance matrix, less time is necessary to determine the Eigen vectors. Moreover, 2DLDA dominates the singularity problem (the size of the database can produce of an amendment of covariance not diagonalizable). The operations of PCA, LDA, 2DPCA and 2DLDA are presented below:

3.2.1 PCA algorithm Stages

The PCA algorithm is presented to the works of Turk and Pentland. The stages of this algorithm are:
Each image $\Gamma_{j(m,n)}$ is transformed into a vector column $\varphi_{j(mn)}$ of m*n size.
The average face of the set is defined by:

$$\varphi_{\text{mean}} = \frac{1}{M}\sum_{j=1}^{M} \varphi_j \quad (14)$$

The data set is adjusted to the average with the following formula:

$$\psi_j = \varphi_j - \varphi_{mean} \quad (15)$$

The covariance matrix who estimates the dispersion of all features vectors at their average is defined by:

$$C = \frac{1}{M}\sum_{j=1}^{M} \psi_j {\psi_j}^t = AA^t, A = [\psi_1, \dots, \psi_M] \quad (16)$$

The Eigen-vectors e_i of C, corresponding to Eigen-values λ_i are computed by using: $Ce_i = \lambda_i e_i$
Sort the Eigen-vectors according to their corresponding Eigen-values.

Finally, each testing image is projected in the training Eigen-space. The images classification is ensured by the Euclidean distance:

$$d_2(x,y) = \sum_{i=1}^{k} \sqrt{(x_i - y_i)^2} \quad (17)$$

Suppose that, $e_i = Av_i$ and $L = A^tA$ one has:

$$Ce_i = \lambda_i e_i \rightarrow Lv_i = \lambda_i v_i \quad (18)$$

The reduction of the big-size matrix C is reduced to the reduction of the matrix L much smaller. To find the Eigen-vectors of C it is enough to multiply by A the Eigen-vectors of L. The maximum of information is contained in the first Eigen-values.

The optimal basis of projection is formed by K better Eigen-vectors associated to K greater Eigen-values.

3.2.2 LDA algorithm Stages

The LDA algorithm is presented to the works of Belhumeur.al. Like the PCA, the stages of LDA algorithm are:
The average per class is defined by:

$$\varphi_{C_i} = \frac{1}{q_i}\sum_{j=1}^{q_i} \varphi_j \quad (19)$$

The images for each class are adjusted to their average:

$$\psi_j = \varphi_j - \varphi_{C_i} \quad (20)$$

Then we calculate the intra-class dispersion matrix Sw, and inter-class dispersion matrix Sb:

$$S_w = \sum_{i=1}^{p}\sum_{\varphi_j \in C_i}(\varphi_j - \varphi_{C_i})\varphi_j - \varphi_{C_i})^t \quad (21)$$

$$S_b = \sum_{i=1}^{p} q_i(\varphi_{mean} - \varphi_{C_i})(\varphi_{mean} - \varphi_{C_i})^t \quad (22)$$

The parameter p is total number of class.
The parameter q_i is number images in the class C_i

The optimal projection W, which maximizes the intra class dispersion, and minimizes the interclass dispersion, must check the following Eigen-values equation:

$$S_b W = \lambda_\omega S_\omega W \quad (23)$$

Thus, the projection of an image is: $g(\varphi_j) = W^t\varphi_j$
The classification is ensured with the Euclidean distance:

$$d_2(x,y) = \sum_{k=1}^{c}\sqrt{(g(\varphi_i) - g(\varphi_j))^2} \quad (24)$$

3.2.3 2DPCA algorithm Stages

The 2DPCA method, proposed by Yang.al is based on the PCA algorithm.
The idea of this method is to project a matrix Γ(m,n) via a linear transformation such as:

$$Y_i = \Gamma R_i \quad (25)$$

Where: Y_i is known as principal component vector of dimension (n,1), and Y_i is the vector projection of dimension (m,1). For M images of training the covariance matrix of size (n, n) is computed as:

$$C = \frac{1}{M}\sum_{j=1}^{M} (\Gamma_j - \psi_{mean})^t (\Gamma_j - \psi_{mean}) \quad (26)$$

Where: ψ_{mean} is the average matrix of the M training images (14).

The optimal projection basis *Ropt=[R1,…,Rg]*, obtained by maximization of generalized variance criterion: $J(R)=R^TCR$, is composed by the first Eigenvectors corresponding to g greater Eigen values of the covariance matrix C.

Then we obtain for each image $\Gamma_{(m,n)}$, a characteristic matrix $Y_{(m,g)}=[Y_1,\ldots,Y_g]$, whose components $Y=\Gamma R_{opt}$.

Finally, the similarity between the test image Y_{test} and training image Y_{train} is measured by the distance: $d=argmin\ ||Y_{test}-Y_{train}||$.

3.2.4 2DLDA algorithm Stages

Inspired by 2DPCA, Visani.al extended the LDA algorithm to two-dimensional LDA (2DLDA). Suppose that images are separate to p-class of q-images, one appointed as:
The average image of i^{éme} class ψ_{C_i}:

$$\psi_{C_i} = \frac{1}{q_i}\sum_{k=1}^{q_i}\Gamma_k \quad (27)$$

The intra-class dispersion matrix S_w

$$S_w = \sum_{i=1}^{p}\sum_{\Gamma_k \in C_i}(\Gamma_k - \psi_{C_i})(\Gamma_k - \psi_{C_i})^t \quad (28)$$

The inter-class dispersion matrix S_b

$$S_b = \sum_{i=1}^{p} q_i(\psi_{mean} - \psi_{C_i})(\psi_{mean} - \psi_{C_i})^t \quad (29)$$

The optimal projection basis $R_{opt}=[R_1,\ldots,R_g]$, obtained by maximization of Fisher's criterion: $J(R)=R^TS_wR/R^TS_bR$, is composed by the first Eigenvectors corresponding to g greater Eigen values of the $S_w^{-1}S_b$ matrix. Then we obtain for each image $\Gamma_{(m,n)}$, a characteristic matrix $Y_{(m,g)}=[Y_1,..,Y_g]$, whose components $Y=\Gamma R_{opt}$

Finally, the similarity between the test image Y_{test} and training image Y_{train} is measured by the distance: $d=argmin\ ||Y_{test}-Y_{train}||$.

4 Experimentations

To optimize the recognition rate, we propose the following strategy:

Initially, we use the Ahonen method who measure the images similarity based on the distance between histograms.

Secondly, we use the obtained image from LBPP to direct input of algorithms PCA, LDA, 2DPCA and 2DLDA.

To evaluate the performances of the proposed approach, we have used the databases ORL and YALE.

To harmonize the gray levels distribution of each image, one pre-treatment with histogram equalization is necessary.

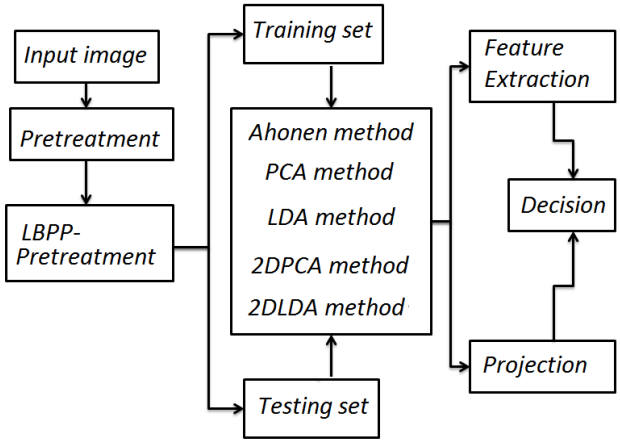


Fig.9 Diagram block of approach steps

The ORL database is made of 40 subjects having each one 10 different views [26]; the images in gray levels have the same size (92x112) pixels. Some images were taken at different times, containing variations of lighting, facial expressions, and details of the face.

The YALE database is made up of 165 images of 15 subjects representing 11 conditions of lighting, pose and facial expressions of size (320x243) [27], centered and scaled to size (112x92) pixels.

As it is shown, in the Fig 3, formed by images respectively obtained by application of LBP, LBPP (k=1), LBPP (k=2), LBPP (k=3) and LBPP (k=4). The pixels of the almost homogeneous areas, and the pixels of the peaks areas are perfectly separate in the case of LBPP (k>2).

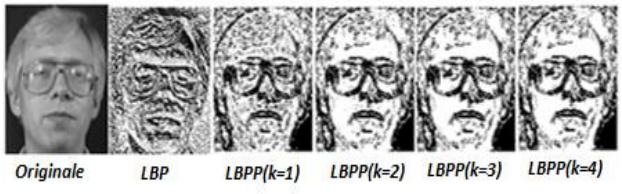


Fig.10 ORL image obtained by LBP and LBPP(k)

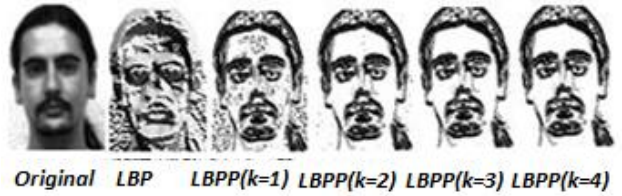


Fig.11 YALE image obtained by LBP and LBPP

4.1 Results of eLBPPH approach

Table2 shows the results to eLBPPH applied on ORL database. To reduce the values of histogram is a widely justified hypothesis.

Table 1 Rate recognition eLBPPH on ORL

	Dimension	LBP	LBPP K=2	LBPP K=3	LBPP K=4
Rate (%)	**[0-15,240-255]**	**90**	**96**	**95,5**	**95,5**
	[0-255]	**92,5**	**94,5**	**94**	**94**
Time (ms)	**[0-15,240-255]**	**59**	**59**	**58**	**58**
	[0-255]	**388**	**365**	**364**	**364**

4.2 PCA, LDA, 2DPCA, 2DLDA and LBPP hybrids Methods

To evaluate the performances of the proposed approach, we use the obtained image from $LBPP(\mathrm{R}=1,\mathrm{P}=8,\mathrm{R}'=2,\mathrm{P}'=25 \text{ and } \beta=0{,}1)$, to direct input of algorithms PCA, LDA, 2DPCA and 2DLDA.

Initially, we compare the performances (rate and time recognition) of 2DPCA, 2DLDA and PCA in terms of rate and time recognition for different dimensions of Eigen-space projection, for a number of 5 training images per class of ORL and YALE databases.

Table 2 Rate and Time recognition for different Eigen-space dimensions on database ORL

Dimension of Eigen-base		**4**	**8**	**12**	**16**	**20**
Rate (%)	LBPP+PCA	**57**	**78,5**	**85,5**	**86,5**	**87,5**
	LBPP+2DPCA	**95**	**96,5**	**95**	**95**	**94**
	LBPP+2DLDA	**95,5**	**96,5**	**96,5**	**96**	**95,5**
Time (ms)	LBPP+PCA	**10,92**	**10,78**	**10,82**	**10,84**	**10,93**
	LBPP+2DPCA	**6,34**	**7,59**	**9,17**	**10,44**	**11,95**
	LBPP+2DLDA	**7,99**	**9,48**	**11,19**	**12,16**	**13,8**

Tables 2 and 3, shows that LBPP+2DPCA and LBPP+2DLDA algorithms reach their maximum rate for a small Eigen-base size:

On ORL database:

LBPP+2DPCA: (dim=8; rate=96, 5; time=7, 59) and LBPP+2DLDA: (dim=8; rate=96, 5; time=9, 48), differently to LBPP+PCA: (dim=18; rate =87,5; time = 10, 93).

On Yale database:
LBPP+2DPCA: (dim=4;rate=95,56 ; time=1,95) and LBPP+2DLDA: (dim=4;rate=96,67; time=2,58), differently to LBPP+PCA: (dim=12; rate =91,11; time = 2,84).

Table 3 Rate and Time recognition for different Eigen-space dimensions on database YALE

Dimension of Eigen-base		4	8	12	14	15
Rate (%)	LBPP+PCA	**66,67**	**85,56**	**91,11**	**91,11**	**91,11**
	LBPP+2DPCA	**95,56**	**95,56**	**94,44**	**94,44**	**92,22**
	LBPP+2DLDA	**96,67**	**94,44**	**94,44**	**93,33**	**93,33**
Time (ms)	LBPP+PCA	**2,83**	**2,84**	**2,84**	**2,87**	**2,89**
	LBPP+2DPCA	**1,95**	**2,18**	**2,44**	**2,5**	**2,59**
	LBPP+2DLDA	**2,58**	**2,88**	**3,18**	**3,42**	**3,41**

We compare the performances of hybridizations of 2DPCA, 2DLDA, PCA and LDA with LBP and LBPP (k):

ORL database: for a random set of 5 training images and 5 testing images per class, and for optimal base of projection; the top results, illustrated in Table 4, show that the association of algorithms (2DPCA, 2DLDA and LBPP (k>2)) gives the best recognition rate (2DPCA (96, 5), 2DLDA (96,5)).

Table 4 Recognition rate for various LBPP methods on ORL database

ORL	PCA	LDA	2DPCA	2DLDA
LBP	70%	82,5%	88%	90,5%
LBPP K=1	88%	88,5%	95%	95,5%
LBPP K=2	88,5%	91%	96%	96%
LBPP K=3	87,5%	91%	96,5%	96,5%
LBPP K=4	87,5%	91%	96,5%	96,5%

YALE database: for a random set of 5 training images and 6 testing images per class, and for optimal base of projection, the top results illustrated in table 5 show that the association of algorithms (2DPCA, 2DLDA and LBPP (k>2)) gives the best recognition rate (2DPCA (96, 67), 2DLDA (93, 33)).

Table 5 Recognition rate for various LBPP methods on YALE database

YALE	PCA	LDA	2DPCA	2DLDA
LBP	75,56	91,11%	90%	91,11%
LBPP K=1	94,44	93,33%	94,44%	95,56%
LBPP K=2	91,11	93,33%	95,56%	96,67%
LBPP K=3	91,11	93,33%	95,56%	96,67%
LBPP K=4	91,11	93,33%	95,56%	96,67%

Eigen-face, Fisher- face, ICA, 2DPCA and 2DLDA. The strategy of comparison uses 5 training images and 5 testing images of each class of database ORL; the experimental results are illustrated in table4.

Table 6 Comparison with various methods of face recognition on ORL

Methods	Recognition rate
PCA[2]	93,50%
LDA[4]	91,50%
ICA[2]	85,00%
2DPCA[2]	96,00%
2DLDA[4]	92,50%
2DPCA+LBPP(k=4)	96,5%
2DLDA+LBPP(k=4)	96,5%

The Figures 13 and 14, expose the recognition rates for different number of training images from the ORL and Yale databases. The number of erroneous testing images depends on the training images number, on used database, on global methods selected, and on the LBPP confidence interval size.

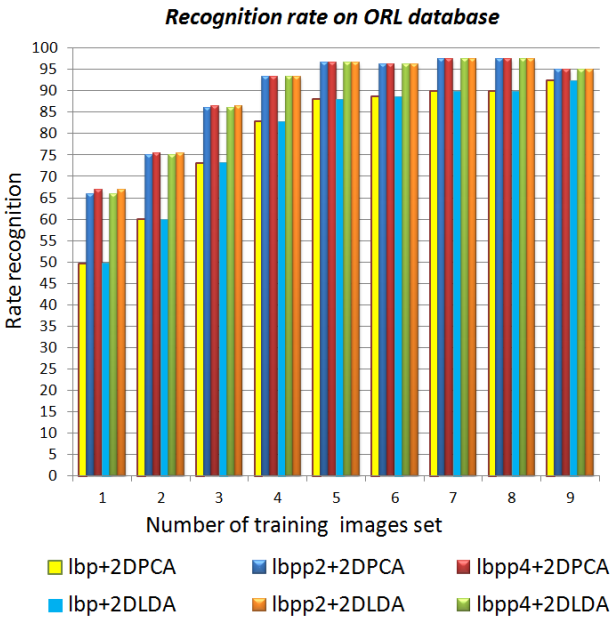


Fig.12 Recognition rate with 2DPCA, 2DLDA, LBP and LBPP(k=2 and k=4) on ORL data

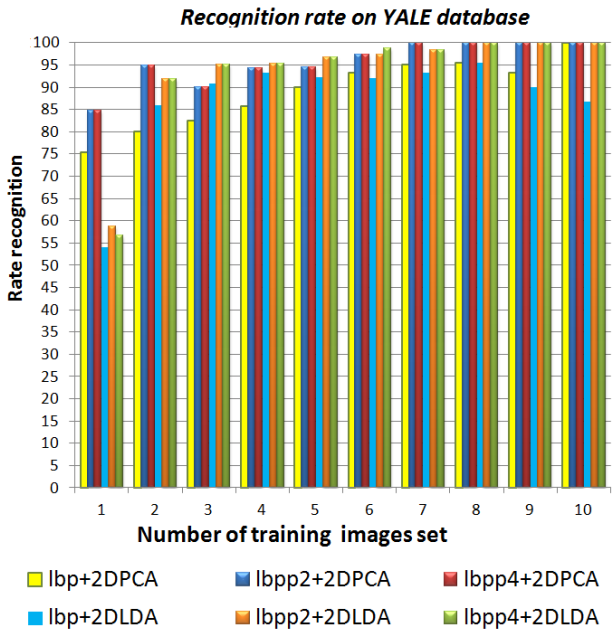


Fig.13 Recognition rate with 2DPCA, 2DLDA, LBP and LBPP(k=2 and k=4) on YALE data

5 Conclusion

In this paper, we have proposed a new model for local characteristics extraction, named LBPP. This probabilistic LBP variety is based on the idea that the face pixels distribution follows a probability law. The normality test estimates the proximity of this low with the sum normal laws. This vision minimizes the specific variance to each face by separating the homogeneous areas from the peaks areas. The covariance matrix of sub-space methods is simplified; the margin of histogram values is reduced; the execution time is improved. It is possible to replace the x2 metric with the Euclidean distance in the Ahonen method.

The obtained results by the face recognition system based on fusion the LBPP with 2DPCA and 2DLDA methods, prove that the approach based on the LBPP is robust that based on LBP. The suggested system reach their best performances for the confidence interval of k>2.In the future, we will continue us search in the face recognition field. Then we will pass to face detection and face localization fields, in order to lead at robust biometric system. We would like to implement this concept to various LBP-varieties, and to merge them with artificial intelligence and frequency approaches.

References:

[1] M. Turk, and A. Pentland, Eigenfaces for Recognition, J.Cognitive Neuroscience, Vol.3, No.1, 1991, pp. 71-86.

[2] J. Yang, D. Zhang, and A.F. Frangi, TwoDimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.26, No.1, 2004, pp. 131-137.

[3] K.Etemad and R.Chellappa, Discriminant Analysis for Recognition of Human Face Images, J. Optical Soc. Am, Vol.14, 1997, pp. 1724-1733.

[4] D.Zhou, X.Yang, N.Peng, and Y.Wang, Improved-LDA based face recognition using both facial global and local information, Pattern Recognition Letters, Vol.27, No.6, 2006, pp. 536-543.

[5] Ms. Bartlett, J.R. Movellan, and T.J. Sejnowski, Face recognition by independent component analysis, Neural Networks, IEEE Transactions on, Vol.13, No.6, 2002, pp. 1450-1464.

[6] HM. El-Bakry, and N. Mastorakis, A novel model of neural networks for fast data detection, WSEAS Transactions on Computers, Vol.5, No.8, 2006, pp. 1773-1780. 113.

[7] D.J. Kim, K.W. Chung , and K.S. Hong, Person authentication using face, teeth, and voice modalities for mobile device security, IEEE T,Consum,Electr,Vol.56, No.4, 2010, pp. 2678-2685.

[8] X. Tan, and B. Triggs, Fusing Gabor and LBP feature sets for kernel-based face recognition, Analysis and Modelling of Faces and Gestures, Springer Berlin Heidelberg, Vol.4778, 2007, pp. 235–249.

[9] T. Ojala, M. P, and D. H ,A comparative study of texture measures with classification based on feature distributions, Pattern Recognition, Vol.29, No.1, 1996, pp. 51–59.

[10] T. Ojala, M. Pietikäinen, and T. Mäenpää, Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns, IEEE Trans, Pattern Analysis and Machine Intelligence, Vol.24, No.7, 2002, pp. 971-987.

[11] T. Ahonen, A. Hadid, and M. Pietikainen, Face Description with Local Binary Patterns: Application to Face Recognition, IEEE T. Pattern Anal, Vol.28, No.12, 2006, pp. 2037-2041.

[12] T. Ahonen, and M. Pietikäinen, Soft histograms for local binary patterns, In Proceedings of the Finnish signal processing symposium, FINSIG, Vol.5, No.9, 2007, pp. 1.
[13] X. Tan, and B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, IEEE Trans. Image Process, Vol.1, No.6, 2007,pp. 1635–1650.
[14] S. Liao, X. Zhu, Z. Lei, L. Zhang, and S.Z. Li, Learning multi-scale block local binary patterns for face recognition, In ICB, 2007, pp. 828–837.
[15] T. Jabid, M.H. Kabir, and O.S. Chae, Local Directional Pattern (LDP) for Face Recognition, Proceeding of the IEEE International Conference of Consumer Electronics, 2010, pp.329-330.
[16] B. Jun, and D. Kim, Robust face detection using local gradient patterns and evidence accumulation, Pattern Recognition, Vol.45, No.9, 2012, pp. 3304-3316.
[17] L. Zhang, R. Chu, S. Xiang, S. Liao, and S. Li, Face detection based on mb-lbp representation, Proceedings of Advances in Biometrics, 2007, pp. 11-18.
[18] S. Zhang, X. Zhao, and B. Lei, Facial expression recognition based on local binary patterns and local fisher discriminant analysis, WSEAS Trans. Signal Process, Vol.8, No.1, 2012, pp. 21-31.
[19] Z. Yang, and H. Ai, Demographic classification with local binary patterns, Advances in Biometrics Springer Berlin Heidelberg, 2007, pp. 464-473.
[20] M.A. Rodrigues, M. Kormann, and P. Tomek, A Comparative Analysis of Binary Patterns with Discrete Cosine Transform for Gender Classification, 2014, Vol.1, pp 33-37.
[21] C.L. Devasena, R. Revathí, and M. Hemalatha, Video Surveillance Systems–A Survey, International Journal of Computer Science, Vol.8, No.4, 2011, pp. 635-642
[22] M. Ettaouil, M. Lazaar, K. Elmoutaouakil, and KA. Haddouch, New Algorithm for Optimization of the Kohonen Network Architectures Using the Continuous Hopfield Networks, Wseas Transactions On Computers, Vol.12, No.4, 2013, pp. 155-165.
[23] A. Maqqor. A. Halli, K. Satori, and H. Tairi, Using HMM Toolkit (HTK) for recognition of arabic manuscripts characters, Multimedia Computing and Systems, International Conference, IEEE, 2014, pp. 475-479.
[24] M. MERRAS, N. EL AKKAD, A. SAAIDI, G. NAZIH, and K. SATORI, Camera Calibration with Varying Parameters Based On Improved Genetic Algorithm, WSEAS Transactions on Computers, Vol.13, 2014, pp.129-137.
[25] N. AHARRANE, K. EL MOUTAOUAKIL, and K. SATORI, Recognition of handwritten Amazigh characters based on zoning methods and MLP, Vol.14, 2015, pp.178-185.
[26] The ORL face database at the AT&T http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase. Site last visited April 2015.
[27] The Yale Face Database, http://vision.ucsd.edu/content/yale-face-database. Site last visited April 2015.