Face Recognition using 1DLBP, DWT and SVM

Amir Benzaoui, Abdelhani Boukrouche, Hakim Doghmane, and Houcine Bourouba

Laboratory of Inverse Problems, Modeling, Information and Systems (PI: MIS)
Department of Electronics and Telecommunication
University of 08 MAI 1945, P.O box 401, Guelma 24000, Algeria
amirbenzaoui@gmail.com Hani.Boukrouche@gmail.com bourouba2004@yahoo.fr doghmane_hakimaz@yahoo.fr

***Abstract*—The popular Local binary patterns (LBP) have been highly successful in describing and recognizing faces. However, the original LBP has several limitations which must to be optimized in order to improve its performances to make it suitable for the needs of different types of problems. In this paper, we investigate a new local texture descriptor for automated human identification using 2D facial imaging, this descriptor, denoted: One Dimensional Local Binary Pattern (1DLBP), produces binary code and inspired from classical LBP. The performances of the textural descriptor have been improved by the introduction of the wavelets in order to reduce the dimensionalities of the resulting vectors without losing information. The 1DLBP descriptor is assessed in comparison to the classical and the extended versions of the LBP descriptor. The experimental results applied on two publically datasets, which are the ORL and AR databases, show that this proposed approach of feature extraction, based on 1DLBP descriptor, given very significant improvements at the recognition rates, superiority in comparison to the state of the art, and a good effectiveness in the unconstrained cases.**

Keywords—biometrics; face recognition; Local Binary Patterns (LBP), One Dimensionnal Local Binary Patterns (1DLBP), Wavelets, DWT, SVM.

I. INTRODUCTION

Biometric systems have increasingly becoming important tool in the information and public security domains; this is because they provide an automatic identification or verification of the identity based on the analysis of physical or behavioral modalities of the human body. Several modalities have been used to recognize the human identity; we can cite fingerprint, voice, iris, palm-print, retina, computer keystroke, or signature [1, 2]… Especially, the automatic analysis of the human face has become an active research area in the artificial vision and pattern recognition domains, due to its important use in several applications such as electronic transactions, biometrics, forensics and video surveillance. The human face is dynamic entity, which changes under the influence of several factors as pose, size, occlusion, background complexity, lighting and the presence of some components such as mustaches, beard, and glasses. So, the essential key for any facial analysis problem is on how to find an efficient descriptor to represent and to model the face in a real context?

The crucial step in any problem of face analysis is the features extraction phase. In this phase, there are two major approaches, local and global approaches. Psychological and neuroscience studies have proved that the human visual system combines between local and global features to differentiate between persons [3]. Global approaches are based on pixel information; all pixels of the facial image are treated as a single vector; the vector size is the total number of the image's pixels [4]. Most of the methods of this approach use another space of representation (subspace) to reduce the number of pixels and to eliminate the redundancies. Principal Component Analysis (PCA) [5], Linear Discernment Analysis (LDA) [6] and Independent Component Analysis (ICA) [7] are the most popular methods used to reduce the dimensions and to select the useful information. However, these approaches are not effective in the unconstrained cases, i.e., situation where occlusion, lighting, pose and size of the face are uncontrolled.

Recently, the scientists concentrate on local approaches, which are considered as a robust approaches in the unconstrained cases compared with global approaches; in this case, the face analysis is given by the individual description of its parts and their relationships, this model corresponds to the manner of perception by the human visual system. The methods of this approach are based on the extraction of features from the facial image and the definition of an adequate model to represent the face [4]. Several methods and strategies have been proposed to model and classify faces essentially based on textures, normalized distances, angles and relations between eyes, mouth, nose and edge of the face. Local Binary Pattern (LBP) [8], Local Gabor Binary Pattern (LGBP) [9] and Oriented Edge Magnitudes (POEM) [10] are the recent methods in this approach.

The specific contributions of this paper are:

- An automated biometric system using 2D face imaging is developed. This work is the continuation of our previous works [11,12] and the objective is to improve the performances of recognition in the unconstrained situations.

- A new textural descriptor is proposed, this descriptor called: One Dimensional Local Binary Pattern (1DLBP). It is essentially inspired from the classical LBP, and projected in one dimensional space.
- The developed approach of feature extraction, based on 1DLBP, is characterized by a combination of the local and global features to analyze, describe, differentiate and recognize persons by using their faces.
- The performances of the realized system have been improved with the introduction of the 1D Wavelets as an efficient mathematical tool in dimensionalities reduction.
- The experimental results applied on two classical databases, the ORL and AR datasets, have showed that this proposed system has given very significant improvements at the recognition rates, superiority in comparison to well-known and classical feature extraction approaches, and a good effectiveness against deferent external factors as occlusion, illumination variation and noise.

This paper is organized as follows: in the next section, we describe the classical LBP, histogram features and the proposed descriptor 1DLBP. In Section 3, the proposed algorithm of feature extraction for face recognition is presented. In Section 4, we present our experimental results by applying the proposed algorithm on ORL and AR databases. Finally, a conclusion related to this work is given in Section 5.

II. LOCAL TEXTURAL DESCRIPTORS

Texture based feature extraction methods play an important role in the fields of computer vision and image processing. Several algorithms of textural features extraction have been proposed during the past years, which can be divided mainly into statistical approaches and structural approaches. More recently, the local texture descriptors have received considerable attention and have been used in several applications such as texture classification, image retrieval or in object recognition. They are distinctive, robust to occlusion, illumination variation, weak lighting, and do not require segmentation. The function of the local descriptor is to convert the pixel-level information into a useful form, which captures the most important contents but is insensitive to irrelevant aspects caused by varying environment. In contrast to global descriptors which compute features directly from the entire image, local descriptors, which have proved to be more effective in real world conditions, represent the features in small local image patches. In the following subsections, we separately discuss the description and the implementation of the LBP and 1DLBP in details.

A. Local Binary Patterns (LBP)

The LBP texture analysis operator, introduced by *Ojala et al.* [13], is defined as a gray-scale invariant texture measure, derived from a general definition of texture in a local neighborhood. It is a powerful mean of texture description and among its properties in real-world applications; we note its discriminative power, computational simplicity and tolerance against monotonic gray-scale changes. The original LBP operator forms labels for the image pixels by thresholding the 3×3 neighborhood of each pixel with the center value and considering the result as a binary number. The histogram of these $2^8 = 256$ different labels can then be used as a texture descriptor for further analysis. This process is illustrated in Fig. 1.

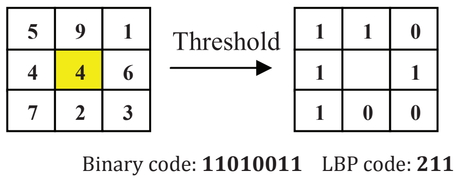

Fig. 1 Calculation of the original LBP operator.

The LBP operator has been extended to use neighborhoods of different sizes. Using a circular neighborhood and bilinearly interpolating values at non-integer pixel coordinates allow any radius and number of pixels in the neighborhood. Each LBP label (or code) can be regarded as a micro-texton. Local primitives which are codified by these labels include different types of curved edges, spots, flat areas etc. The notation (P, R) is generally used for pixel neighborhoods to refer to P sampling points on a circle of radius R as shown in Fig. 2. Examples of LBPs applications with different mask are shown in Fig. 3. The calculation of the LBP codes can be easily done in a single scan through the image. The value of the LBP code of a pixel (x_c, y_c) is given by:

$$LBP^{P,R}(x_c, y_c) = \sum_{i=1}^{p} S\left(g_i^{P,R} - g_c\right)2^{i-1} \quad (1)$$

$$S(x)\ is\ defined\ as\text{:}\ S(x) = \begin{cases} 1 & if\ x \geq 0; \\ 0 & otherwise; \end{cases} \quad (2)$$

Where g_c and $g_i^{P,R}$ are respectively the values of the central element and its 2D neighbors.

The occurrences of the LBP codes in the image can be collected into a histogram. The classification can then be performed by computing histogram similarities. For an efficient representation, facial images are first divided into several local regions from which LBP histograms are extracted and then concatenated into an enhanced feature histogram for classification [14].

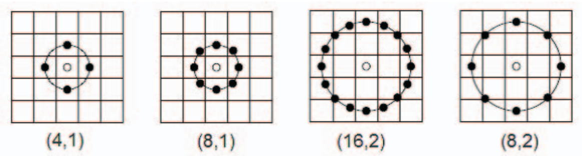

Fig. 2 Neighborhood set for different (P, R).

Original Image **LBP (8,1)** **LBP (8,2)** **LBP (16,2)**

Fig. 3 Examples of LBPs application with different masks.

In [3], psychological and neuroscience studies have showed that the human visual system combines between local and global features to recognize and differentiate between peoples. In the other hand, the extend versions of LBP operators present a good results by capturing the local patterns and the micro features of the human face, but they are not performed for capturing the global patterns that can be considered as dominants structures in the image, which is in contradictory to the theory of recognition demonstrated in neuroscience and psychological sciences.

B. One Dimensinnal Local Binary Patterns (1DLBP)

The Projected Local Binary Pattern in One dimensional space (1DLBP) was introduced for the first time by *L. Houam et al.* [15]; it has been combined with wavelets to classify X-ray bone images for bone disease diagnosis. The concept of the 1DLBP method consists in a binary code describing the local agitation of a segment in 1D signal. It is calculated by thresholding of the neighborhood values with the central value. All neighbors get the value 1 if they are greater or equal to the current element and 0 otherwise. Then, each element of the resulting vector is multiplied by a weight according to its position (see Fig. 4.c). Finally, the current element is replaced by the sum of the resulting vector. This can be recapitulated as follows [16]:

$$1DLBP\ (x_c) = \sum_{n=0}^{N-1} S(g_n - g_0)2^n \quad (3)$$

$$S(x)\ function\ is\ defined\ as\ (2)$$

Where g_0 and g_n are respectively the values of the central element and its 1D neighbors. The index *n* increases from the left to the right in the 1D string as shown in Figure 4.c.The 1DLBP descriptor is defined by the histogram of the 1D patterns.

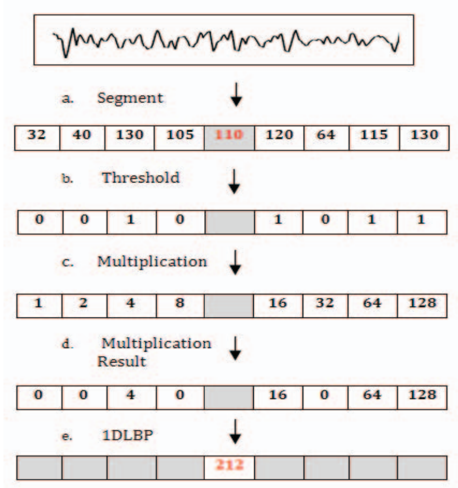

Fig. 4 Example of 1DLBP Application.

III. PROPOSED APPROACH

The proposed biometric system requires two phases of operations. The first phase is called the training which consists of recording faces features from each individual in order to create his own biometric template; this latter is stored in the database. The second phase is the test, which consists in recording the same features and compares them to the biometric templates stored in the database. If the recorded data match a biometric template from the database; the individual in such a case is considered identified. The proposed algorithm used to extract information for face description and recognition is described in the following recapitulation:

- *Preprocessing.*
- *Multi-block decomposition of the image.*
- *Projection of each decomposed block in one dimensional space.*
- *Application of the proposed descriptor 1DLBP for each projected block.*
- *Concatenation of the resulting vectors from each block in one global vector.*
- *Dimensionalities Reduction of each global vector using 1D Wavelets.*
- *Classification / Template Matching*

A. Preprocessing

The objective of the preprocessing is the preparation of the source's image representation to facilitate the task of the following steps and to improve the rate of recognition. First, the facial image is converted into grayscale image. Next, every grayscale image is filtered by median filter to reduce noise. Lastly, the noise suppression image is then adjusted to improve the contrast of the image.

B. *Image decomposition*

Most LBP operators characterize the face texture distribution of each pixel with its neighborhood only. But, the differences between two faces cannot only demonstrate by the texture distribution of each pixel with its neighborhood, but also by the relative connection with other pixels. With this intention, we have decomposed the original image into several sub-images (see Fig.5) to characterize better the details and the relationships between all the image pixels. Next, the extracted histograms will be concatenated in one global vector in the next stages. With this technique, we can obtain the fine details and the relative connections between all pixels.

Fig. 5 The decomposition of the image into differents blocks.

C. *1D vertical Projection*

The 1D projection of rows or columns of each level provides an effective mean to describe better the local and global patterns. Fig.6 presents an example of vertical projection. The objective of the projection is to validate the descriptor LBP in one dimensional space to find another mean for describing and analyzing better the human face's texture.

D. *1DLBP Application and Histogram's Concatenation*

To project a block from two dimensional representations (matrix) in one dimensional space (vector), the average values of all pixels from each column are calculated to create the projected vector. Next, the proposed descriptor 1DLBP is applied on all projected blocks of the decomposed image with the different resolutions, as presented in Fig.6. Then, the extracted vectors from each decomposed block will be concatenated in one global vector representing the face's template.

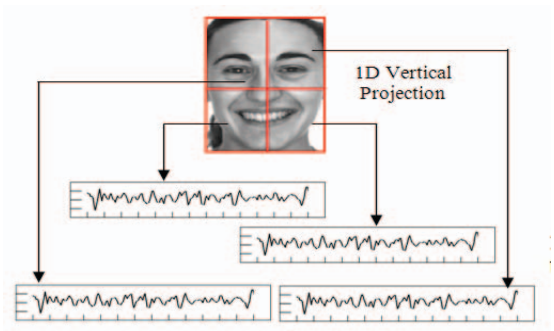

Fig. 6 Example of Vertical Projection in One Dimensional Space.

A problem of redundancies in information occurred due to the important size of the templates, which should contain essential information to make the classification task easier. For that purpose, we have used the one dimensional *Wavelets* as a robust mathematical tool, in dimensionalities reduction without losing information; this can be viewed as: given p-dimensional vectors $Y = (y_1, y_2, \dots, y_p)$, the objective is to find a representation in the lower dimension $Z = (z_1, z_2, \dots, z_q)$, where $q<p$ which preserves the content of the original data as much as possible.

The Discrete Wavelet Transform (DWT) of a sequence consists of two series expansions, one corresponding to the approximation and the other to the details of the sequence. The formal definition of DWT of an *N*-point sequence $x[n]$, $0 \le n \le N-1$ is given by [17, 18]:

$$\mathrm{DWT}\,\{f(t)\} = W_{\emptyset}(j_0, k) + W_{\Psi}(j, k) \tag{4}$$

Where:

$$W_{\emptyset}(j_0, k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n]\, \emptyset_{j_0,k}[n] \tag{5}$$

$$W_{\Psi}(j, k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n]\, \Psi_{j,k}[n], \; j \ge j_0 \tag{6}$$

The sequence $x[n]$, $0 \le n \le N-1$ can be recovered from the DWT coefficients $W\varphi$ and $W\psi$ as given by:

$$x[n] = \frac{1}{\sqrt{N}} \sum_k W_{\emptyset}(j_0, k)\, \emptyset_{j_0,k}[n] + \frac{1}{\sqrt{N}} \sum_{j=j_0}^{\infty} \sum_k W_{\Psi}(j,k)\, \Psi_{j,k}[n] \tag{7}$$

The wavelet's analysis is the breaking up of a signal into a set of scaled and translated versions of an original (or mother) wavelet. The wavelet transform of a signal decomposes the original signal into wavelets coefficients at different scales and positions, as shown in Fig.7. These coefficients represent the signal in the wavelet domain and all data operations can be performed using just the; corresponding wavelet coefficients.

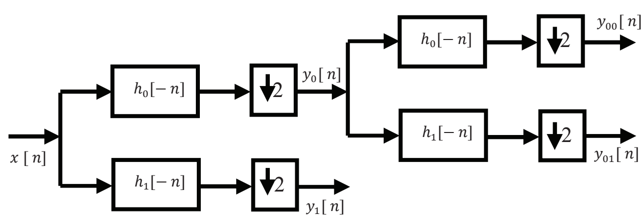

Fig. 7. Signal's decomposion using the 1D discrete wavelets transform [17, 18].

Wavelets work by decomposing the signal (concatenated histograms in our case) into different resolutions or frequency bands, as shown in Fig.7. The extraction of the pertinent information is based on the concept of selecting a small number of approximated coefficients at a suitably chosen level and some detail coefficients can accurately represent regular signal components. Choosing a decomposition level for the

DWT usually depends on the type of signal being analyzed or some other suitable criterion such as entropy. The level of decomposition and the type of wavelet that gives the best results in our experiments are: level 2 of decomposition and *Haar Wavelets*.

E. Classification / Template Matching

Regarding the learning algorithm, several approaches have been proposed, including, among others, neural networks and their invariant, Support Vector Machines (SVM) / Support Vector Regressors (SVR), Random Forests (RF), and projection techniques such as Canonical Correlation Analysis (CCA)… From the wide variety of learning schemes presented in the literature, Support Vector Machines (SVM) and its derivations have recently obtained state-of-the-art results in challenging large databases.

In our system, matching is carried out after training the system, the aim of matching is to evaluate the similarity between the testing image and the stored templates; giving a new image ξ considered as a test example. First, we build its 1D representation (the template) using the feature extraction method proposed in this work. Next, we use the SVM to classify the image ξ to the corresponding class.

The type of the SVM used for classification is: SVM with RBF kernel, we select the most widely used kernel function, i.e., RBF (Radius Basis Function). The parameter in the RBF kernel function was empirically selected in this paper (γ = 0.0001).

IV. EXPERIMENTAL RESULTS

To evaluate the performances of the proposed descriptor, we have carried out several tests on ORL and AR databases; we randomly selected half of samples for training set and the remaining samples for testing set. In all our experiments, we considered the average recognition rates of several random permutations (50 permutation with ORL database and 100 permutations with AR database), and we compared the obtained results (identification and false alarm rates) with other methods using the same testing protocols.

A. ORL database

The ORL database contains 400 frontal images in different facial expression, conditions of illumination, hairstyles with or without beard, moustaches and glasses for 40 persons, 10 images for each person. Each sample is a 92×112 gray image, with tolerance for some tilting and rotation of up to 20° (Fig. 8).

Fig. 8. Some Images from ORL Database.

B. AR database

The AR database was collected at the Computer Vision Center in Barcelona, Spain in 1998 [19]. It contains images of 116 individuals (63 men / 53 women). The imaging and recording conditions (camera parameters, illumination setting, and camera distance) were carefully controlled and constantly recalibrated to ensure that settings are identical across subjects. The resulting RGB color images are 768×576 pixels in size. The subjects were recorded twice at a 2–week interval. During each session 13 conditions with varying facial expressions, illumination and occlusion were captured (Fig. 9).

Fig. 9. Some Images from AR Database.

C. Experiments

In the first test, we have applied the methods inspired from the LBP texture analysis in the two databases (classical LBP, extended LBP, and the one dimensional Local Binary Patterns 1DLBP). The performances of these methods are shown in Table 1.

TABLE I. COMPARATIVE RECOGNITION RESULTS OF THE ISPIRED LBP METHODS ON ORL AND AR DATABASES.

RR: Recognition Rate %. **FAR:** False Alarm Rate%.

	ORL RR %	AR RR %	Average RR %	Average FAR %
LBP (8,1)	85.2	87.5	86.4	3.62
LBP (8,2)	86	89.4	87.7	3.1
LBP (8,3)	86.4	89.7	88.1	2.9
LBP (16,2)	86.1	89.2	87.7	3.2
LBP (8,1) + PCA	91.4	93	92.2	1.92
LBP (8,3) + PCA	92	93.2	92.6	1.8
LBP (8,1) + DWT	92.5	93.5	93	2.5
LBP (8,3) + DWT	92.9	94.3	93.6	2
1DLBP	92	92.6	92.3	2.36
1DLBP + PCA	95.8	97.9	96.9	1.44
1DLBP + DWT	**96.1**	**98.3**	**97.2**	**1.1**

We see that LBP(8,3) provides a good results than LBP(8,2) and LBP(8,2) provides a good results than LBP(8,1), which means that the descriptors of the large scale sizes can capture more information and details than the descriptors of short sizes. Also, we can see that projected LBPs in one dimensional space (1DLBP), shows a good performances than LBPs in the two dimensional space with different extensions. The results of the experiments clearly showed that the projected Local Binary Pattern with dimensionalities reduction using DWT enhances the recognition performance in all configurations, presents a very good improvement and significant results in recognition rate, false alarm rate against other variants of LBPs and an improvements in comparison to our previous works [11, 12] where we have used the PCA in the dimensionalities reduction.

We also conducted tests comparing our method against recent and classical state-of-the-art approaches (like PCA, LDA, POEM and LGBP) using the similar protocols under more challenges and scenarios. The results, shown in Table 2, indicate clearly the effectiveness of our approach which outperforms all other methods.

TABLE II. COMPARISON TO OTHER STATE OF THE ART APPRAOCHES.

	ORL RR %	AR RR %	Average RR %	Average FAR %
PCA	85	81.2	83.1	5.12
LDA	82.4	84.7	83.6	5.36
POEM	90.9	89.4	90.2	2.6
LGBP	94.1	96.2	95.2	1.87
LBP (8,3) + DWT	92.9	94.3	93.6	2
1DLBP + DWT	**96.1**	**98.3**	**97.2**	**1.1**

The facial images are taken with ORL database which is considered as a stable database in the unconstrained cases, and the AR database which presents a very good variation, in lighting, occlusion and facial expression, to measure the performances of the approach in the difficult situations. The comparison results presented in Table2 shows that our proposed approach presents very good results with AR database which indicates that this approach has an important effectiveness against to the variations in different factors like: partial occlusion and illuminations variation.

V. CONCLUSION

In our investigation, we have successfully developed a new feature extraction approach for automated 2D face description and recognition. We have introduced the use of the 1DLBP as a local texture descriptor, which is inspired from the classical LBP and produces a binary code. In contrast to the earlier approach (LBP) where the descriptor is applied directly on the original image, the 1DLBP is applied on projected vectors, in 1D signal, from the original image. A series of experimental evaluations on *ORL and AR* databases show that this proposed approach of feature extraction, based on 1DLBP descriptor, given very significant improvements at the recognition rates, superiority in comparison to the state of the art, and a good effectiveness in the unconstrained cases.

REFERENCES

[1] J.A. Unar, W.C. Seng, and A. Abbasi, "A Review of biometric technology along with trends and prospects," Pattern Recognition (Elsevier), vol.47, no.8, pp.2673-2688, 2014.

[2] J. Ashbourn, editor "Biometrics in the new world: The cloud, Mobile technology and Pervasive identity," Springer Publishing, 2014.

[3] A.J. O'Toole, D.A. Roark, and H. Abdi, "Recognizing moving faces: A psychological and neural synthesis," Trends in Cognitive Science, Vol.06, No.06, pp.261-266, 2002.

[4] R.S. Kathavarayan and M. Karuppasamy, "Identification of Untrained Facial Image in Combined Global and Local Preserving Feature Space," International Journal of Biometrics and Bioinformatics, Vol.04, No.01, pp.1-12, 2010.

[5] M. Turk and A. Pentland, "Eigenfaces for recognition," Journal of Cognitive Neuroscience, Vol.03, No.01, pp.71-86, 1991.

[6] P.N. Belhumeur, J.P Hespanha and D.J. Kriegman, "Eigenfaces vs Fisherfaces: Recognition using class specific linear projection," IEEE Transactions on PAMI, Vol. 19, No.07, pp.711-720, 1997.

[7] A. Hyvarinen and E. Oja, "Independent component analysis: Algorithms and applications," Neural Networks, Vol. 13, No. 4-5, pp. 411–430, 2000.

[8] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Description with Local Binary Pattern: Application to Face Recognition," IEEE Transactions on PAMI, Vol.28, No.12, pp.2037-2041, 2006.

[9] H.V. Nguyen, L. Bai, and L. Shen, "Local Gabor Binary Pattern whitened PCA: A novel approach for face recognition from single image per person," Springer: Proceeding of the 3rd International Conference on Biometrics, Vol.5558, pp.269-278, 2009.

[10] N.S Vu and A. Caplier. "Face recognition with patterns of oriented edge magnitudes," Springer: Proceeding of the 11th European Conference on Computer Vision, Vol.6311, No.01, pp.313-326, 2010.

[11] A. Benzaoui and A. Boukrouche, "1DLBP and PCA for Face Recognition", Proceeding of the 11th International IEEE Symposium on Programming and Systems, pp.07-11, 2013.

[12] A. Benzaoui and A. Boukrouche, "Face Recognition using 1DLBP Texture Analysis", Proceeding of the 5th International Conference on Future Computational Technologies and Applications, pp.14-19, 2013.

[13] T. Ojala, M. Pietikainen and D. Harwood, "A comparative study of texture measures with classification based on featured distribution," Elsevier: Pattern Recognition, Vol.29, No.01, pp.51-59, 1996.

[14] A. Benzaoui, A. Hadid and A. Boukrouche. "Ear Biometric Recognition using Local Texture Descriptors," Journal of Electronic Imaging (SPIE), Vol.23, No.05, pp.053008, 2014.

[15] L. Houam, A. Hafiane, R. Jennane, A. Boukrouche and E. Lespessailes. "Trabecular Bone Anisotropy Characterisation Using 1D Local Binary Patterns," Springer: Proceeding of the 12th Conference on Advanced Concepts for Intelligent Vision Systems, Vol.6474, No.01, pp.105-113, 2010.

[16] A. Benzaoui and A. Boukrouche, "Face Analysis, Description, and Recognition using Improved Local Binary Patterns in One Dimensional Space", Journal of Control Engineering and Applied Informatics (CEAI), Vol. 16, No.04, pp.52-60, 2014.

[17] R. C. Gonzalez and R. E. Woods, "Digital Image Processing (2nd eds.)", Prentice Hall, Upper Saddle River, NJ, 2002.

[18] K. S. Thyagarajan, "Still Image and Video Compression with Matlab (1st eds.)", A John Wiley & Sons, Inc., Publication, 2011.

[19] A.M. Martinez and R. Benavente. "The AR face database", CVC Technical Report, 24, 1998.