A Robust Approach to Characterize the Human Ear: Application to Biometric Identification

Article *in* International Journal of Imaging and Robotics · August 2015

2 authors:

Amir Benzaoui
Université de Bouira

Boukrouche Abdelhani
Université 8 mai 1945 - Guelma

International Journal of Imaging and Robotics,
[Formerly known as the “International Journal of Imaging” (ISSN 0974-0627)]
Volume 15; Issue No. 1; Year 2015; Int. J. Imag. Robot.
ISSN 2231–525X;

A Robust Approach to Characterize the Human Ear: Application to Biometric Identification

Amir Benzaoui[1] and **Abdelhani Boukrouche**[1]

[1]Laboratory of Inverse Problems, Modeling, Information and Systems (PI: MIS),
Department of Electronics and Telecommunications,
University of 08 Mai 1945, P.O box 401 - 24000, Guelma (Algeria);
Email : amirbenzaoui@gmail.com abdelhanib@yahoo.fr

ABSTRACT

The Human ear is a new technology of biometrics which is not yet used in a real context or in commercial applications. For this purpose of biometric system, we present an improvement for ear recognition methods that use Elliptical Local Binary Pattern operator as a robust descriptor for characterizing the fine details of the two dimensional ear imaging. The improvements are focused on features extractions and dimensionalities reductions steps. The realized system is mainly appropriate for identification mode; it starts by decomposing the normalized ear image into several blocks with different resolutions. Next, the local textural descriptor is applied on each decomposed block. A problem of information redundancies is appeared due to the important size of the concatenated histograms of all blocks, which has been resolved by reducing of the histogram's dimensionalities and by selecting the pertinent information using Haar Wavelets. Finally, the system is evaluated on the IIT Delhi Database containing two dimensional ear images and we have obtained a success rate about 97% for 493 images from 125 persons and about 96% for 793 images from 221 persons.

Keywords: Biometrics, Ear Recognition, LBP, ELBP, Wavelets.

Mathematics Subject Classification: 74E25, 68T10, 30F45

Computing Classification System: G.4

1. INTRODUCTION

Biometric systems have an important role in the information and public security domains. 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 for recognizing the human identity, we can cite face [1], fingerprint [2], voice [3], iris [4], brain [5], palmprint [6], computer keyboards [7], or gait [8]. In the last few years, the use of the human ear in forensics and in biometric applications has become a quite interesting way; it is considered as a new class of biometrics which is not yet used in a real context or in commercial applications. The human ear is characterized by a rich and stable structure which provides important information to differentiate and recognize peoples; we can visualize 10 features and 37 sub-features from two dimensional ear imaging. The terminology of the human ear is presented in Fig.1; this terminology is made up of standard features. It includes an outer rim (helix) and ridges (antihelix) parallel to the helix, the concha (hollow part of ear), the lobe and the tragus [9, 10, 11]. The human ear has several advantages compared with others modalities; it has a rich structure, smaller object (small resolution), modality stable over the time, modality accepted by people, not affected by changes in age, facial expression, position and rotation, the acquisition of the ear imaging can be affected without participation of the subject and can be captured from distance [9, 12].

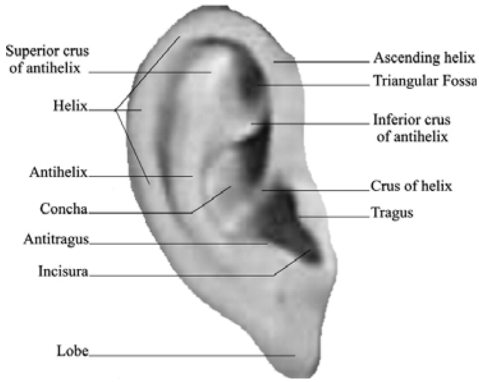

Figure 1. The terminology of the Human Ear [9, 10, 11].

An ear recognition system can be divided into three main steps: ear normalization, feature extraction and classification. In the normalization step, the ear image must be normalized to standard size and standard direction according to the long axis of the outer ear contour. The long axis was defined as the line crossing through the two points which have the longest distance on the ear contour. After normalization, the long axes of different ear images were normalized to the same length and the same direction (fig.2) [13, 14]. In feature extraction, we can recapitalize the existing methods in two principal classes, geometric and global classes. In geometric approaches, the ear analysis is given by an individual description of its parts and their relationships, like measure of distances, angles, or the triangulation between the closed edge resulting from the helix shape and the ear's lobule [15], Crus of Helix and the Lobe [15], or the Vornoi diagram of the ear triangulation [16]. However, the methods of this approach are not effective in the unconstrained cases, i.e., situation where occlusion, lighting or noise are uncontrolled. Recently, the scientists concentrate on global approaches, which are considered as robust approaches in the unconstrained cases compared to geometric approaches. Global approaches are based on pixel information; all pixels of the normalized ear image are treated as a single vector; the vector's size is the total number of the pixels. Principal Component Analysis (PCA) [17], Color Spaces fusion [18], 2D Gabor Filter [19], or Curvelet transformation [20] are the most popular methods used to change the space of representation, to change the data representation, to reduce the dimensionalities or to select only the useful information.

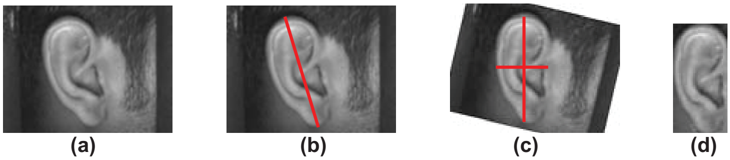

Fig. 2. Process of ear normalization. (a) Original image (b) The long axis of outer ear detection (c) Rotation (d) Normalized ear extraction [13, 14].

This paper presents improvements for automated ear recognition systems and comparison with last state of the art approaches. There are four steps in the proposed algorithm. First, a pre-processing is applied to improve the contrast and the quality of the images. Second, the ear image is decomposed into several blocks with different resolutions. Next, the textural descriptor ELBP is applied on all blocks of the decomposed image. Finally, the *Haar Wavelets* in the one dimensional space are used to reduce the dimensionalities of the concatenated histograms extracted from each block. K neighbor nearest (K-NN) classifier with Hamming distance are used to classify the testing images into the corresponding class. *IIT Delhi* database 1 and 2 are used to evaluate the performances of the proposed approach. The rest of this paper is organized as follows: in the next Section, we describe the classical LBP and ELBP. In the Section 3, the enrolment phase and ear characteristic extraction is presented. In the Section 4, we present the experimental results and performances evaluation applied on *IIT Delhi* database. For this purpose, *Hamming* distance is required to measure similarities between ear templates. Finally, a conclusion related to this work is given in Section 5.

2. EAR DESCRIPTION USING LOCAL BINARY PATTERN (LBP)

The original LBP operator introduced by Ojala *et al.* [21], which has been used for texture discrimination, has shown a powerful and effective results against the variations in rotation and illumination. The operator labels the pixels of an image by thresholding the 3×3 neighborhood of each pixel with the central value and considering the result as a binary code. Next, the histogram of the labels can be used as a textural descriptor (see Fig. 3).

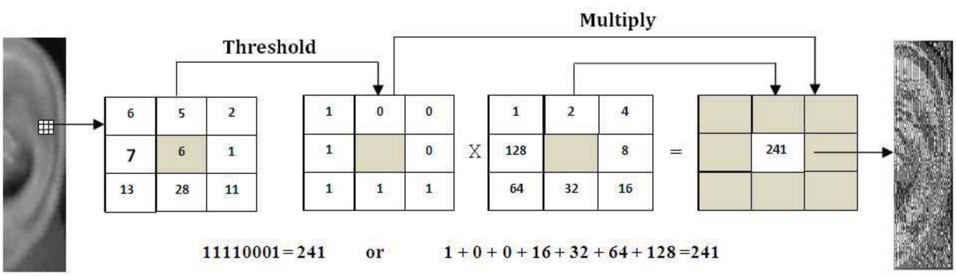

Figure 3. LBP Calculation performed into 3×3 neighborhood.

For a given a pixel $g_c(x_c,y_c)$ from gray image, its texture LBP is calculated by comparing g_c with its neighbors pixels P on a circle of radius R (see Fig.4 for more details on circular neighborhood). The value of LBP (g_c) is obtained as [22, 23, 24]:

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

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

However, the major disadvantage of the original LBP operator resides in the size of the descriptor, a mask of 3×3 pixels cannot capture the structures of large scales which can considered as a dominants structures in the image [24]. Recently, the size of the operator has been extended by using a mask with different large sizes. Fig.4.a, Fig.4.b, and Fig.4.c show three examples of the extended LBP.

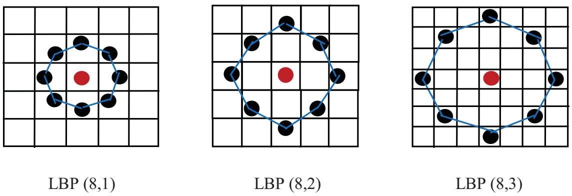

Figure 4. Neighborhood set for different (P,R). (a) The basic LBP operator (P,R)=(8,1) (b) LBP with circular neighborhood (8,2). (c) LBP with circular neighborhood (8,3).

A new model of the extended operators of LBP called: Elliptical Local Binary Patterns (ELBP) [25, 26] which presents significant improvements for encoding the micro features (the fine details) of the image compared with LBP. In ELBP, at each pixel $g_c(x_c,y_c)$, we consider its surrounding pixels that lie on an ellipse (see Fig.5) with (x_c,y_c) is the center. ELBP of (x_c,y_c) with P neighboring pixels at (R1, R2) distances is computed as:

$$ELBP^{P,R1,R2}(x_c, y_c) = \sum_{i=1}^{P} S(g_i^{P,R1,R2} - g_c)2^{i-1} \tag{3}$$

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

In details, the coordinates of the i^{th} neighboring pixel of (x_c,y_c) are calculated using the formulas:

$$angle\_step = 2 * \pi / P \tag{4}$$

$$x_i = x_c + R1 * cos\,((\,i - 1) * angle\_step) \tag{5}$$

$$y_i = y_c + R2 * sin\,((\,i - 1) * angle\_step) \tag{6}$$

Fig.6 shows the results of LBP and ELBP applications using different masks. However, the extended versions of LBP operators and the Elliptical LBPs present good results by capturing the local patterns and the micro features of the image but they are not performed for capturing the global characteristics which can be considered as dominants structures in the image.

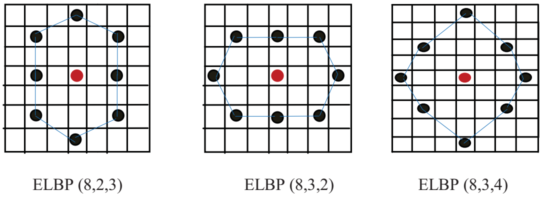

Figure 5. ELBP samples with different extention [25,26].

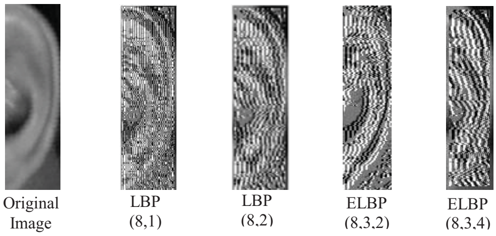

Figure 6. Results of LBP and ELBP application with different masks.

3. PROPOSED APPROACH

The proposed algorithm used to extract information for ear recognition is described in the following recapitulation; next we present each step in details. The process of features extraction is composed of five principal stages for each person:

- *Pre-processing of the all normalized ear images.*
- *Decomposition of each ear image into several blocks with different resolutions.*
- *Application of the textural descriptor ELBP for each decomposed block.*
- *Concatenation of the resulting histograms from each block in one global histogram.*
- *Dimensionalities reduction of each global histogram using Haar Wavelets.*

3.1. Pre-processing

The objective of the preprocessing is the modification of the source's image representation to facilitate the task of the following steps and to improve the performances of recognition. First, the ear image is converted into grayscale image. Next, every grayscale image is filtered by median filter to suppress noise. Lastly, the noise suppression image is then adjusted to improve the contrast of the image [27].

3.2. Image Decomposition

Most LBP operators characterize the ear texture distribution of each pixel with its neighborhood only. But, the difference between two ears cannot be demonstrated by the texture distribution of each pixel with its neighborhood only, but also by the relative connection with other pixels. With this intention, we have decomposed the original image into several sub-images (see Fig.7) to characterize better the details and the relationships between most pixels of the image. Next, the extracted histograms will be concatenated in one global histogram in the next stages. With this technique, we can obtain the fine details and the relative connections between most pixels.

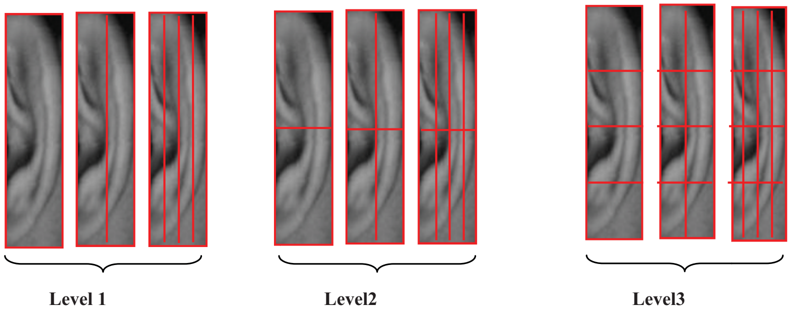

Figure 7. Image decomposition into different blocks.

3.3. Application of the Textural Descriptor and Histogram's Concatenation

The textural descriptor ELBP, presented in the previous section, is applied on all blocks of the decomposed image at different resolutions like presented in Fig.7. Next, the extracted histograms from each block will be concatenated in one global histogram (vector) representing an ear template. A problem of information redundancies is appeared due to the important size of each template.

3.4. Dimensionalities Reduction Using Haar Wavelets

To resolve the problem of information redundancies, we have used the *Discrete Wavelet Transform (DWT)* as a technique of data compression, to reduce the dimensionalities and to select only the useful information needed to model each person. Wavelet analysis is the breaking up of a signal into a set of scaled and translated versions of an original (or mother) wavelet. Taking the wavelet transform of a signal decomposes the original signal into wavelets coefficients at different scales and positions. These coefficients represent the signal in the wavelet domain and all data operations can be performed using just the; corresponding wavelet coefficients [28].

Wavelets work by decomposing the concatenated histograms (global histogram) into different resolutions or frequency bands, choosing the *Haar Wavelet* and computing the 1D Discrete Wavelet Transform (DWT). The extraction of pertinent information is based on the concept of selecting a small number of approximation 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. For the processing of ear's templates decomposition, in our application, up to scale 3 are adequate.

3.5. Template Matching

Now, for a new given image ξ considered as a testing example. We start by building its pertinent global histogram representation using the feature extraction method presented in this work. Next, we calculate the Hamming Distance (HD) to classify the testing image ξ in the nearest neighbor class corresponding to a degree of similarities that exceeds a fixed threshold. Hamming Distance formula is given as:

$$HD(X,Y) = \frac{1}{M}\sum_{i=0}^{M} X_i \ XOR \ Y_i \qquad (07)$$

X_i and Y_i are the two bit-wise codes to be compared.

M is the size of the vector representing the template.

4. EXPERIMENTAL RESULTS

To evaluate the performances of the realized system, we have carried out several tests on *IIT Delhi* database; we randomly selected two images from each person as training set and the remaining samples for testing set. In all our experiments, we considered the average recognition rates of several random permutations (50 permutations), and we compared the obtained results (identification and false alarm rates) with other methods using the same testing protocols. Our experiments are implemented with *Matlab* 2010a, *Windows* 7, HP Core 2 Duo, 3 Ghz CPU with 2 GB Ram.

4.1. IIT Delhi Database

The *IIT Delhi* ear image database consists of the ear image database collected from the students and staff at IIT Delhi, New Delhi, India. This database has been acquired in IIT Delhi campus during Oct 2006 - Jun 2007 using a simple imaging setup. All the images are acquired from a distance using simple imaging setup and the imaging is performed in the indoor environment. The currently available database of 493 images is acquired from the 121 different subjects (Database 1) and 793 images acquired from 793 different subjects (Database 2), each subject have at least three ear images. All the subjects in the database are in the age group 14-58 years. The resolution of these images is 272×204 pixels and all these images are available in jpeg format (Fig.8). In addition to the original images, this database also provides the automatically normalized and cropped ear images of size 50×180 pixels (Fig.9) [19, 29].

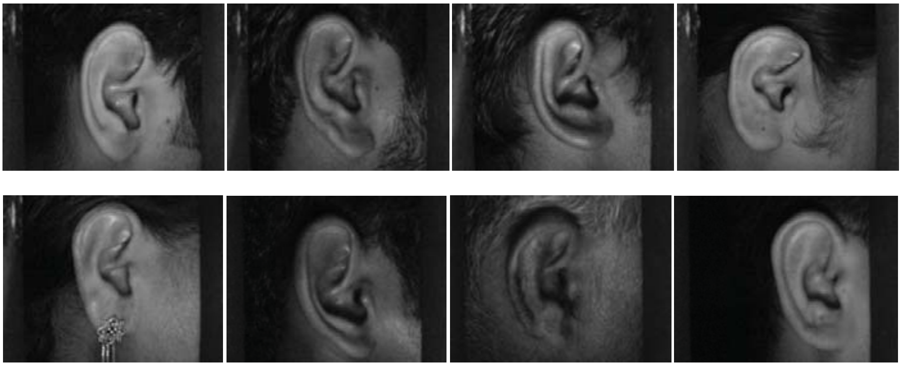

Figure 8. Some images from IIT Delhi database [19,29].

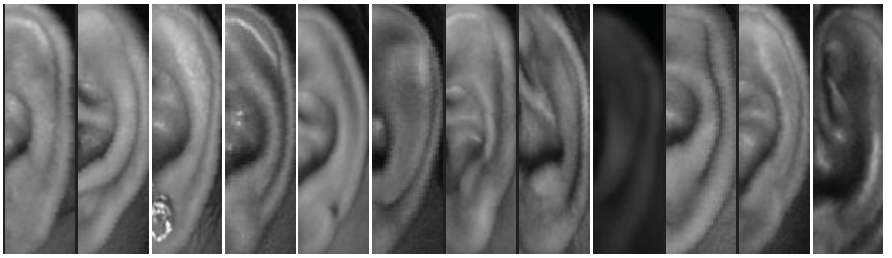

Figure 9. Some images from normalized IIT Delhi database [19,29].

In this evaluation, 2 images for each person are used as *training-set* and the remaining images of the same person (minimum 1image) are used as *testing-sets*. The proposed approach has been evaluated in all possible probabilities and the mean rate of all possible selection has taken in consideration. We have compared also the proposed approach with other state of art approaches using the similar protocols under more challenges and scenarios. The ear images are taken with *IIT Delhi* database which presents a very good variation, in lighting, noise and occlusion, to measure the performances of the proposed approach in the difficult situations. The results of the comparison are presented in Table1.

Table 1: Average Results and Comparison to other State of the Art Approaches using the same Distribution

	Database 1		***Database 2***	
	False Alarm Rate %	***Recognition Rate %***	***False Alarm Rate %***	***Recognition Rate %***
PCA without Decomposition	2,75	80,39	4	76
Classical LBP	1,56	86,3	3,7	83,1
Classical LBP +PCA	1,46	92,45	2,64	86,49
Classical LBP + DWT	1,46	92,71	2,11	90,14
LBP (8,2)	1,24	89,8	3,6	80,45
LBP (8,3)	1,45	86,14	3,32	81,14
LBP (8,2) + PCA	1,25	90,43	1,6	90
ELBP (8,2,3)	1.02	91,4	1.83	88,37
ELBP (8,3,2)	1,14	91,63	1,34	89,80
ELBP (8,3,4)	0,97	91,39	1,03	89,82
ELBP (8,3,2) + PCA	**0,26**	94,44	0,41	92,9
ELBP (8,3,2) + DWT	0,34	**96,4**	**0,39**	**95,88**

Table.1 shows clearly that ELBP presents very good results with *IIT Delhi* database 1 and 2, which indicates that this descriptor has important effectiveness against the variations in different factors like occlusion, lighting, rotation, and noise.

The comparison results from Table1 show that our ELBP (8,3,2) + DWT method outperforms other state of the art systems using the same distribution. These results also reveal that ELBP (8,3,2), in normalized ear modality, is more robust than LBP in the extraction of the fine details (micro features).

Another conclusion we can make from the Table.1 is that ELBP + DWT is much better than ELBP PCA; the association of the DWT as a robust technique in the dimensionalities reduction is very interesting to improve the performances of the realized approach.

5. CONCLUSION

In this work, we have successfully developed a simple approach for ear texture discrimination; this approach is primary based on Elliptical Local Binary Patterns. Each normalized ear image is decomposed on multi-blocks with different resolutions, and the textural descriptor ELBP will be applied on each decomposed block. Next, the extracted histograms from each block will be concatenated in one global histogram. Finally, DWT Wavelets is applied on each global histogram to reduce the dimensionalities and to extract the useful information. The experimental results applied on *IIT Delhi* database 1 and 2 have showed that the feature extraction approach based on ELBP and DWT has given a very significant improvement at the recognition rate, false alarm rate and a good effectiveness against several external factors like noise, illumination, and rotation.

REFERENCES

1. J.A. Dargham, A. Chekima, E. Gubin, and M. Hamdan, Hybrid Face Recognition System Based on Local Discriminant Analysis and Voting, *International Journal of Imaging and Robotics (IJIR)*. **12**(1) (2014) 106-116.

2. R. Cappelli, Fast and Accurate Fingerprint Indexing Based on Ridges Orientation and Frequency, *IEEE trans. on SMC-B*. **41**(6) (2011) 1511-1521.

3. H. Bourouba, C.A. Korba, and R. Djemili, Novel Approach in Speaker Identification using SVM and GMM, *Journal of Control Engineering and Applied Informatics (CEAI)*. **15**(3) (2013) 87-95.

4. C.H. Sulochana and S. Silvan, Iris Recognition Based on Directional Filtering and Classifier Fusion, *International Journal of Imaging and Robotics (IJIR)*. **2**(9) (2009) 106-116.

5. K. Aloui, A. Nait-Ali, and M.S. Naceur, A Novel Approach Based Brain Biometrics: Some Preliminary Results for Individual Identification, in *Proc. IEEE International Workshop on Computational Intelligence on Biometrics and Identity Management (CIBIM)*, (France, Paris, 2011) 91-95.

6. R. Cappelli, M. Ferrara, and D. Maio, A Fast and Accurate Palmprint Recognition System Based on Minutiae, *IEEE trans. on SMC-B*. **42**(3) (2012) 956-962.

7. D. Rudrapal, S. Das, and S. Debbarma, Improvisation of Biometrics Authentication and Identification through Keystrokes Pattern Analysis, in *Proc. 10th Springer Int. Conf. on Distributed Computing and Internet Technology (ICDCIT)*, (Bhubaneswar, India, 2014), 287-292

8. C.P. Lee, A.W.C. Tan, and S.C. Tan, Time-Sliced averaged motion history image for Gait Recognition. *Journal of Visual Communication and Image Representation (Elsevier).* **25**(5) (2014) 822-826.

9. B. Arbab-Zavar and M.S. Nixon, On guided model-based analysis for ear biometrics, *Computer Vision and Image Understanding (Elsevier).***115**(2011) 487-502.

10. A. Kumar and T.C. Tak-Shing, Robust ear Identification using Sparse Representation of Local Texture Descriptors, *Pattern Recognition (Elsevier)*. **46**(2013) 73-85.

11. A.D. Dinkar and S.S. Sambyal, Person Identification in Ethnic Indian Goans using Ear Biometrics and Neural Networks, *Forensic Science International (Elsevier)*. **223**(2012) 373.e1-373.e13.

12. L. Yuan and Z.C. Mu, Ear Recognition Based on Local Information Fusion, *Pattern Recognition Letters (Elsevier)*, **33**(2012) 182-190.

13. L. Yuan and Z.C. Mu, Ear Recognition Based 2D images, in *Proc. 1st IEEE Int. Conf. on Biometrics: Theory, Applications and Systems (BTAS)*, (USA, Washington, 2007).

14. Z. Xie and Z.C. Mu, Ear Recognition Using LLE and IDLLE Algorithm, in *Proc. 19th IEEE Int. Conf. on Pattern Recognition (ICPR)*, (USA, Florida, 2008).

15. M. Burge and W. Burger, Ear Biometrics in Computer Vision, in *Proc. 15th IEEE Int. Conf. on Pattern Recognition (ICPR)*, (Spain, Barcelona, 2000).

16. M. Burge and W. Burger, Ear Biometrics in Machine Vision, in *Proc. 21st Int. Workshop. of the Australian Association for Pattern Recognition*, (Australia, 1997).

17. B. Arbab-Zavar, M.S. Nixon and D.J. Hurley, On Model-Based Analysis of Ear Biometrics, in *Proc. 1st IEEE Int. Conf. on Biometrics: Theory, Applications and Systems (BTAS)*, (USA, Washington, 2007).

18. L. Nanni and A. Lumini, Fusion of Color Spaces for Ear Authentication*, Pattern Recognition (Elsevier)* **42**(2009) 1906-1913.

19. A. Kumar and C. Wu, Automated Human Identification Using Ear Imaging, *Pattern Recognition (Elsevier)*. **45**(2012) 956-968.

20. A. Basit and M. Shoaib, A human ear Recognition method using nonlinear curvelet feature subspace, *International Journal of Computer Mathematics*. (2013) To be published.

21. T. Ojala, M. Pietikainen and D. Harwood, A Comparative Study of Texture Measures with Classification based on Features Distribution, *Pattern Recognition (Elsevier)*. **29**(6) (1996) 51-59.

22. T. Ahonen, A. Hadid, and M. Pietikainen, Face Recognition with Local Binary Pattern, in *Proc. 8th Springer Int. European Conf. on* Computer Vision *(ECCV)*, (Czech, Prague, 2004), 469-481.

23. T. Ahonen, A. Hadid, and M. Pietikainen, Face Description with Local Binary Pattern: Application to Face Recognition, *IEEE trans. on PAMI*. **28**(12) (2006) 2037-2041.

24. M. Pietikainen, A. Hadid, G. Zhao and T. Ahonen (eds.), *Computer Vision Using Local Binary Patterns*, 1stedn. (Springer London, 2010).

25. H.T. Nguyen, N.S. Vu and A. Caplier, How far we can improve micro features based face recognition systems?, in *Proc. 3rd IEEE Int. Conf. on Image Processing, Theory, Tools and Applications (IPTA)*, (Turkey, Istanbul, 2012), 350-353.

26. H.T. Nguyen and A. Caplier, Elliptical Local Binary Pattern for Face Recognition, in *Proc. 12thSpringer Int. Workshops on Computer Vision (ACCV)*, (Korea, Daejeon, 2012), pp. 85-96.

27. A. Benzaoui, H. Bourouba and A. Boukrouche, System for Automatic Faces Detection, in *Proc. 3rd IEEE Int. Conf. on Image Processing, Theory, Tools and Applications (IPTA)*, (Turkey, Istanbul, 2012), 354-358.

28. S. Nasehi and H. Pourghassem, Online Mental Task Classification based on DWT-PCA Features and Probabilistic Neural Network, *International Journal of Imaging and Robotics (IJIR)*. **7**(1) (2012) 110-118.

29. IIT Delhi Ear Database, Version 1: http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Ear.htm