Hakim Doghmane[1,*], Hocine Bourouba[1], Kamel Messaoudi[2] El-Bey Bournene[3]

J. Electrical Systems 15-4 (2019): 607-625

Regular paper

A Spatial Pyramidal Decomposition Method for ear representation using local dual cross patterns

Journal of Electrical Systems

In recent years, several scientific works are oriented to develop optimal ear representation, for ear recognition, which is discriminant, compact, and easy-to-implement to ensure the best performance in terms of accuracy, computation cost, and storage requirement. In this manner, this paper presents a novel ear representation based on texture analysis framework, which relies mainly on Dual Cross Pattern (DCP) descriptor and Spatial Pyramid Histogram (SPH) method. The features are extracted using DCP descriptor to capture the textural structure then, the SPH of horizontal ear decomposition is applied to obtain the local information. The feature vector representations of ear image are constructed by concatenating all normalized histograms calculated at each level of the SPH method. Experiments conducted on three ear databases (IIT-Delhi-1, IIT-Delhi-2 and USTB-1) confirm its performance compared to the recent existing methods.

Keywords: Ear recognition; LDCP; D-LDCP; SPH; WLDA; K-NN.

Article history: Received 1 June 2019, Accepted 14 September 2019

1. Introduction

Due to the growing need for security, biometric personal identification technologies have been motivated. Usually, the main goal of the biometric system is to identify / verify the subject of interest based on physical or behavioral characteristics [1]. Among the various traits used for human recognition, the human ear has been used as a means of identification in forensic medicine over several years. This is due to its non-intrusive collection, stable structure (does not change significantly during aging) and its shape does not vary with changes in facial expression [2]. Nevertheless, the quality of ear images is affected by some factors such as: the variation of the pose and hidden objects(the hair, the earrings, the earphones, etc.).

Three main steps are, generally, required for human ear-based technology: i) ear detection, ii) feature extraction and iii) classification step. The aim of the first step consists to localize and isolate the human ear from an input profile face image. Ear detection step is important because errors at this stage can undermine the utility of the system.

The second step consists to extract a salient set of features to represent the human ear. The feature extraction step reduces the segmented ear to a mathematical model called as a feature vector that summarizes the discriminative information present in the ear image. In the recent decades, many feature extraction techniques have been proposed. These techniques can be classified into three categories, which are: appearance based-techniques, local based-techniques and hybrid based-techniques.

* Corresponding author: H. Doghmane, Guelma university, Laboratory of Inverse Problems, Modeling Information and Systems (PIMIS), Algeria, E-mail: doghmane.hakim@univ-guelma.dz or doghmane_hakimaz@yahoo.fr

[1] Guelma university, Laboratory of Inverse Problems, Modeling Information and Systems (PIMIS), Algeria

[2] Mohemed Cherif Messaadia University, Souk-Ahras, Algeria

[3]LE2I Laboratory, Burgundy University, BP 47 870, Dijon, France

The first category exploits the appearance of the entire ear image. It contains several linear and nonlinear projection techniques, in which the ear image can be presented by the weighted linear combination of some corresponding eigenvectors. Moreover, the appearance based-techniques are based on pixel information; all pixels in the image are treated as a single vector. Indeed, the total number of these pixels represents the size of the vector. Most of the techniques in this category use another representation space (subspace) to reduce the number of pixels and eliminate redundancies. Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Independent Component Analysis (ICA) are the most popular techniques used to reduce the dimensions. Victor et al. [3] was the first research group that transferred the idea of using the subspaces (Eigen-spaces) of face recognition to ear recognition. They suggested that the performance of the ear, as a biometric modality, is less inferior than the performance of the face. This may be due to the fact that in their experiments, they considered that the left and right ears are symmetrical. Zhang and Mu [4] conducted studies on the effectiveness of holistic techniques. They showed, on the USTB-1 ear database, that independent component analysis (ICA) is more efficient compared to PCA. Furthermore, Xie and Mu [5] showed, on the USTB-3 ear database, that the Locally Linear Embedding (LLE) algorithm is more efficient than the PCA, if the input data contains pose variations. However, the most techniques of this category are sensitive to the variation of lighting and pose. Therefore, a specific preprocessing step should be established before [6].

The local based-techniques consist in extracting relevant features that are more discriminative. Among them: The Scale Invariant Feature Transform (SIFT) is considered a robust descriptor with some geometric transformations and scale under different lighting conditions. The SIFT descriptors is used to extract the discriminative information features for ear image in [7]. The Speeded Up Robust Features (SURF) transform is used to extract the robust features for ear image [8], which are invariant to scale variation and other geometrical transforms. The local texture descriptors have been studied for ear recognition system in [9]. The authors show that the performances obtained by the Binarized Statistical Image Features (BSIF) are more better than those obtained by Local Binary Pattern (LBP) and Local Phase Quantization (LPQ). Moreover, it has been shown that the human ear can be composed of two regions: rigid and semi-rigid [10]. As a result, using only the first region improves the recognition accuracy. Based on multi-scale analysis framework a novel feature representation called Multi-scale Binarized Statistical image features (MS-BSIF) is proposed, in which the feature extraction of the BSIF descriptors is extended to the multi-scale space [11].

The hybrid based-techniques consist of extracting local features by the local based-techniques and reduces its dimensionality using one of the previous appearance based-techniques. This category is based mainly on the combination of the two previous categories to improve the performance of recognition system.

In this paper, a novel feature representation for ear recognition is proposed. After the preprocessing step of the raw ear images, the local textural structure are extracted using Dual Cross Pattern (DCP) descriptor. Then the spatial pyramid histogram(SPH) method is used for horizontal decomposition to construct a set of local features that will be then normalized and concatenated into a large histogram called Local Dual Cross Pattern (LDCP) feature for each ear image. After that the LDCP features are projected into a

whitened linear discriminant (WLDA) [12] subspace to reduce its dimensionality and make these features more discriminant providing, therefore, the Discriminant Local Dual cross Pattern (D-LDCP) feature. Finally, the K-nearest neighbor (K-NN) classifier is used for classification stage to identify each test ear image.

The reminder of this paper is organized as follows: in Section 2, we describe the main steps of the proposed ear representation strategy. Then, Section 3 reports and discusses the results of ear identification experiments. Finally, Section 4 draws some conclusions.

2. Ear image representation based on the LDCP descriptor

Motivated by C. Ding et al. [13], this paper presents a new attempt to merge the Dual Cross Patterns and the Spatial pyramid histogram decomposition method for ear recognition. A novel ear representation method is proposed, which explores the textural structure invariant property in the spatial domain among different scales and orientations. An illustration of the proposed ear representation is depicted in Figure 1.

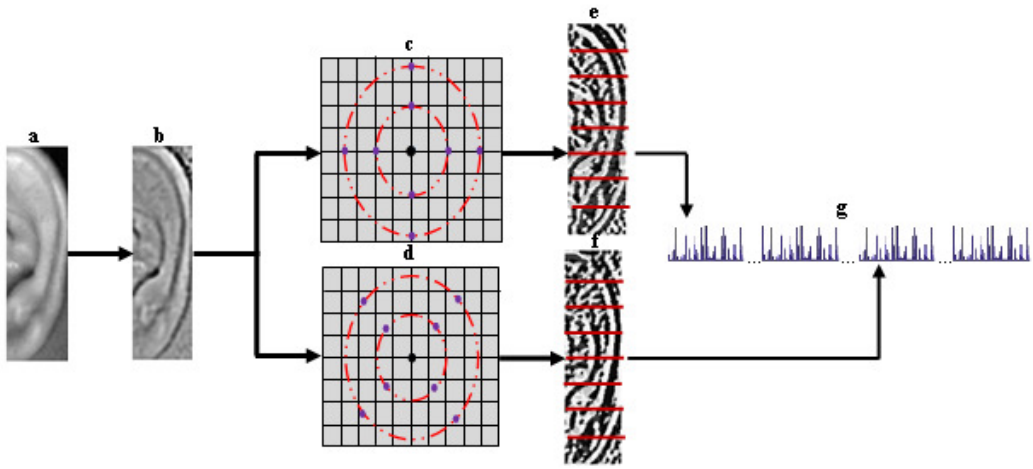

Figure 1: Ear image representation using local pyramid DCP scheme: (a) original ear image, (b) normalized image, (c) DCP_1sampling, (d) DCP_2 sampling, (e)-(f) SPH decomposition method applied to DCP_1 and DCP_2 descriptors respectively, (g) local DCP (LDCP) descriptor.

The following subsections detail the different steps of the proposed ear representation.

2.1. Preprocessing and feature extraction step

Ear preprocessing step is executed before ear feature extraction. It is applied on the original ear images using the median filter. Then, The filtered ear images are normalized to zero mean and unit standard deviation. This allows cancelling the effect of the illumination variations. The feature extraction scheme is obtained by constructing structural and textural information, using DCP descriptor. Furthermore, the local ear feature descriptors are obtained via the SPH method. The whitened LDA is used to generate the Discriminant LDCP (D-LDCP) descriptor to: i) reduce the high dimensionality of a large global feature histogram and ii) enhance the discriminating ability of the features.

2.1.1. DCP descriptor

The dual-cross pattern (DCP) proposed by C. Ding et al., [13] is a texture descriptor, which characterizes the spatial structure of the local image texture at multiple levels. The DCP operator is used to perform local sampling and pattern encoding in the most informative directions. The principle of Local sampling of the DCP descriptor is illustrated in Figure 2. The sampled points are encoded in two steps. Firstly, the textural information in horizontal, vertical and diagonal directions are encoded independently. Secondly, the encoded patterns are combined to form The DCP codes.

Around the central pixel, sixteen points are sampled. The sampled points A_0 to A_7 are uniformly spaced on an inner circle of radius R_{In}, while B_0 to B_7 are evenly distributed on the exterior circle with radius R_{Ex} [13].

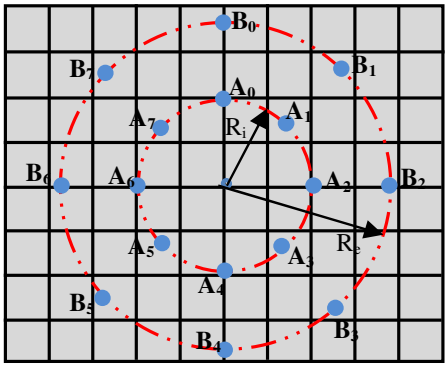

Figure 2: An illustration of DCP sampling.

For each of the eight directions, we assign a unique decimal number:

$$DCP_i = S\left(I_{A_i} - I_0\right)2 + S\left(I_{B_i} - I_{A_i}\right) \qquad for\ 0 \le i \le 7 \tag{1}$$

Where $S(x)$ is the sign function given by:

$$S(x) = \begin{cases} 1 & if\ x > 0 \\ 0 & otherwise \end{cases} \tag{2}$$

Two subsets are constructed from the eight patterns DCP_0-DCP_7. The first subset is defined by {DCP_0, DCP_2, DCP_4, DCP_6} (Figure 1.c) and the second subset is given by {DCP_1, DCP_3, DCP_5, DCP_7} (Figure 1.d). The two subsets described above are called cross encoders DCP_1 and DCP_2, respectively. Therefore, the codes produced by these two encoders at each pixel are represented as:

$$\begin{cases} DCP\_1 = \sum_{i=0}^{3} DCP_{2i}.4^i \\ DCP\_2 = \sum_{i=0}^{3} DCP_{2i+1}.4^i \end{cases} \tag{3}$$

So, the DCP descriptor for ear image is result of the concatenation of the two codes DCP_1 and DCP_2 [13].

$$DCP = \{DCP\_1, DCP\_2\} = \left\{ \sum_{i=0}^{3} DCP_{2i}.4^i, \sum_{i=0}^{3} DCP_{2i+1}.4^i \right\} \quad (4)$$

2.1.2. Spatial Pyramid Histogram (SPH) method

The Spatial Pyramid Histogram (SPH) approach introduced firstly by Lazebnik [14]. The process of building the spatial pyramid histogram with level L is indicated in Figure 3. Firstly, the histogram at level 0 from the entire DCP descriptor image is constructed. Then, the DCP descriptor image is divided into two equal sized regions at level 1 by using horizontal decomposition and the histogram is computed for each region. The process is repeated by recursively subdividing each region at level *l* and computing histograms in each region until the desired level L is reached. As a result, there are 2^l histograms at level *l* and that by summing this number over *l*=0,…, L a spatial pyramid histogram with L levels will have a total of $(2^{L+1}-1)$ histograms, then all these histograms are normalized in range [0 1] and concatenated together into a large vector.

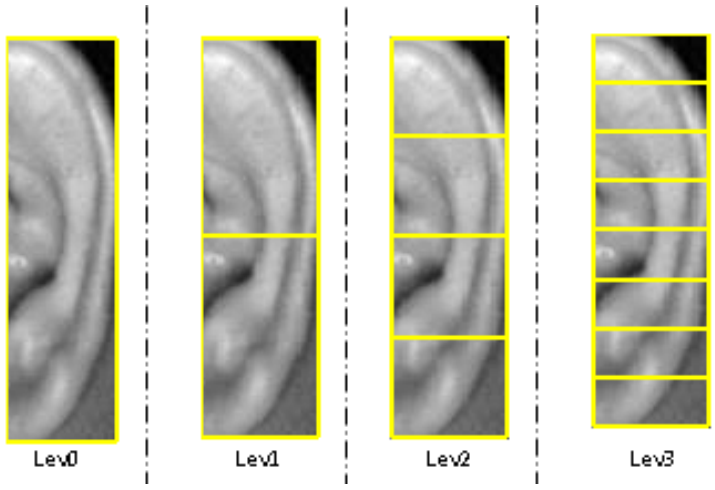

Figure 3: The principle of SPH for ear horizontal decomposition.

2.1.3. Whitened Linear Discriminant Analysis (WLDA)

The data analysis methods are used for reducing the initial space dimension while keeping only the more discriminant properties of the extracted features. Several approaches, linear or nonlinear projection, were used in the literature for ear recognition to reduce the dimensionality of features extraction, such as PCA and LDA. The WLDA technique proved its effectiveness compared to PCA and LDA for face recognition [12]. The based idea of WLDA is to apply a data whitening transformation before performing LDA algorithm. Thus, with the above transformation, we can whiten the random vector to have zero mean and the identity covariance matrix to decorrelate the features.

In our experiments, the dimension of features is reduced from $2x256x(2^{L+1}-1)$ to C−1 for each database (i.e., 124, 220, and 59 for IITD-1, IITD-1 and USTB-1 database, respectively), where C represents the class number.

2.2. Classification step

Let us see a test ear sample Y from one of the classes in the training set. At first, we compute the proposed representation. Then, the obtained feature vector of the input ear is matched with all the stored templates, and the most similar one is taken as the matching result. For the classification step, a nearest neighbor classifier is used with Chi-square distance d to calculate the distance between two histograms for LDCP descriptor vector.

The Chi-square distance between two histograms H^1, and H^2 of length N, is defined as:

$$d_{\chi^2} = \sum_{i=1}^{N} \frac{\left(H^1(l) - H^2(l)\right)^2}{H^1(l) + H^2(l)} \tag{5}$$

In the case of D-LDCP descriptor vector, cosine similarity has been used. This similarity S between two features vectors x and y of length N is defined as:

$$S(x, y) = \frac{\sum_{i=0}^{N-1} x_i y_i}{\sqrt{\sum_{i=0}^{N-1} x_i^2} \sqrt{\sum_{i=0}^{N-1} y_i^2}} \tag{6}$$

3. Experimental results and discussions

The important parameters of the proposed approach are empirically proved to achieve the highest identification rate. The proposed approach depends on three important parameters which are: the radius of inner circle (R_{In}), the radius of exterior circle (R_{Ex}) used for DCP sampling, and the level (L) of horizontal decomposition of the SPH method. In order to determine the optimal combination parameter values that yield the best results, different values of radius R_{In} and R_{Ex} are used with different level decompositions (L) that change from 0 to 4. The Error Identification Rates (EIR) and Identification rates (IR) are used to evaluate the performance.

To demonstrate the effectiveness of the proposed method for the ear identification task, a thorough evaluation is performed. We applied it on three ear databases which are: Indian Institute of Technology Delhi (IIT Delhi) ear database [15] for the both versions (IIT Delhi-1 and IIT Delhi-2), and the first version of the USTB (USTB-1) database [16] .

3.1. Evaluation protocols

The following protocols are established for the three ear databases:

- Two ear images of each person are taken as a training set and the remaining ear images as a testing in the first experiment.
- One ear image of each subject is taken as a training set and the remaining ear images as a testing in the second experiment.

- In the third experiment, the proposed approach will be compared to other recent methods for ear recognition.
- The effect of the parameters R_{In}, R_{Ex} and L are evaluated.
- The K-NN classifier is used for classification stage.
- Three permutations are performed to provide the average recognition rate.

It should be noted that, a preprocessing step is performed in all experiments, by applying the median filter and then the filtered images are normalized to zero mean and unit standard deviation.

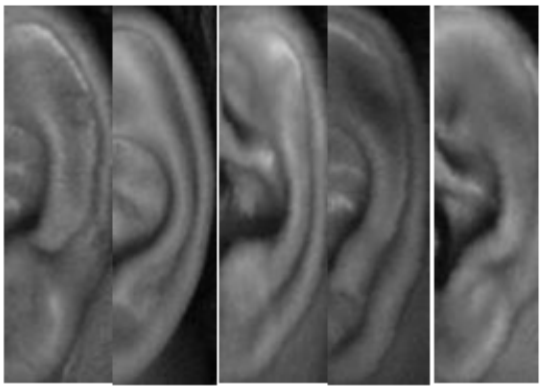

Figure 4: Some ear images of normalized II Delhi database.

3.2 IIT Delhi database

For the both versions of IIT Delhi (IIT Delhi-1 and IIT Delhi-2) database:

- The resolution of each ear image is 272x204 pixels.
- Each subject has at least three ear images.
- A normalized and cropped ear images of size 180 × 50 pixels are provided.
- Significant scale, translational and rotational variations.
- All subjects are in the age group 14–58 years.

3.2.1 IIT Delhi-1

The IIT Delhi-1 database contains 493 ear images acquired from 125 different subject.

Experiment #1

In this case, we will study the effect of parameters R_{In}, R_{Ex} and L on the Error Identification Rate (EIR). We choose R_{In} changed from 1 to 6 and R_{Ex} varies from R_{In}+1 to 7 in sequence to discover its impact on recognition performance of our method. Meanwhile, the decomposition level L is changed from zero to four.

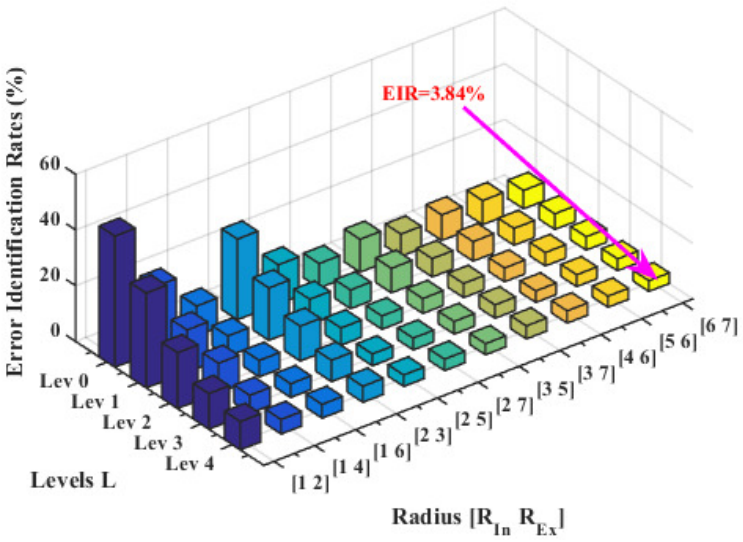

Figure 5: Error Identification Rates using LDCP descriptor.

We use two ear images of each subject as a training set and remaining ear images as a testing set, which is similar to [9, 10, 11]. Under such experimental settings, the training set contains 250 images, whereas the testing set contains 243 images.

From the results given by Figure 5, It can be seen that when the parameters [R_{In} R_{Ex}] are equal to [6 7], we obtain a good performance (EIR=3.84%) with level decomposition L=4. For a fixed value of R_{In}, the error identification rate (EIR) decreases when the radius R_{Ex} increases. Furthermore, the EIR is inversely proportional to the radius R_{Ex}. Moreover, the EIR varies inversely with the decomposition level L of the SPH method.

Table 1: Recognition rates using LDCP descriptor.

	[1, 2]	[1, 3]	[1, 4]	[1, 5]	[1, 6]	[1, 7]	[2, 3]	[2, 4]	[2, 5]	[2, 6]	[2, 7]	[3, 4]	[3, 5]	[3, 6]	[3, 7]	[4, 5]	[4, 6]	[4, 7]	[5, 6]	[5, 7]	[6, 7]
L0	52.95	65.98	76.68	84.64	89.44	91.08	71.06	79.84	86.01	90.12	91.22	84.64	88.20	90.81	92.04	89.03	90.81	91.91	91.36	93.14	93.42
L1	65.98	78.33	84.91	90.53	92.73	93.42	81.21	86.42	90.95	92.73	93.28	89.57	91.22	92.04	93.42	93.00	93.00	93.83	93.96	94.24	94.38
L2	79.97	85.87	89.44	93.55	94.10	95.06	87.52	92.73	93.83	94.38	95.20	92.73	94.65	94.92	94.92	95.06	95.06	95.20	95.20	95.34	95.20
L3	87.24	90.95	93.55	94.65	95.06	95.20	92.32	94.38	95.75	96.16	95.75	95.47	95.47	95.47	95.20	95.75	95.88	95.61	95.75	95.88	95.75
L4	89.85	93.55	94.79	94.92	95.06	95.20	94.24	94.92	95.88	95.88	95.88	96.16	96.02	95.47	95.34	95.88	95.61	95.61	96.02	95.88	**96.16**

In addition, Table 1 gives more results and details based on the LDCP descriptor, using the Identification Rates (IR) to evaluate the performance. If no decomposition is performed (L=0), the identification rates (IR) take small values for different combinations of R_{In} and R_{Ex}. The lowest value of IR is of 52.95% obtained for radius values [R_{In} R_{Ex}]=[1 2] and L=0.

In order to prove the effectiveness of the proposed and the discriminating ability of the proposed approach, the previous experiment is repeated using WLDA reduction. In this way, the cosine distance with K-NN classifier is used. Figure 6 and Table 2 show the results obtained with reduction based on EIR and IR performances, respectively.

Based on the results presented in Figure 6 (D-LDCP), it can be seen that the EIR is significantly reduced compared to the results obtained in Figure 5 (LDCP). The lowest EIR

goes from 3.84% (LDCP descriptor) to 2.61% (D-LDCP descriptor). Moreover, it can be seen that when the parameters [R_{In} R_{Ex}] are equal to [2 5], a lowest error (EIR=2.61%) is obtained with level decomposition L = 4.

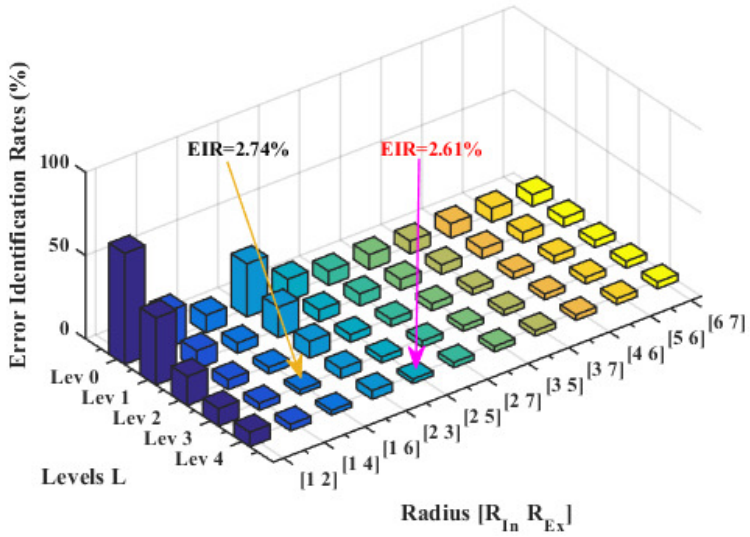

Figure 6: Error Identification Rates using D-LDCP descriptor.

Table 2: Recognition rates using D-LDCP descriptor.

	[1, 2]	[1, 3]	[1, 4]	[1, 5]	[1, 6]	[1, 7]	[2, 3]	[2, 4]	[2, 5]	[2, 6]	[2, 7]	[3, 4]	[3, 5]	[3, 6]	[3, 7]	[4, 5]	[4, 6]	[4, 7]	[5, 6]	[5, 7]	[6, 7]
L0	33.33	59.81	78.19	83.81	89.44	89.57	68.04	79.84	86.83	89.99	91.50	85.87	91.08	91.63	92.32	91.36	91.36	92.46	91.77	92.46	92.87
L1	59.81	78.05	89.03	91.77	93.83	93.55	81.76	88.48	92.87	93.96	92.73	92.87	93.42	94.10	94.24	94.24	94.10	93.96	93.83	94.24	94.51
L2	82.58	89.57	93.96	95.06	96.02	96.84	89.99	93.55	95.34	96.02	95.61	96.30	95.75	95.61	96.16	95.61	95.88	95.34	95.61	95.34	95.75
L3	89.57	93.28	95.75	96.16	97.26	97.26	93.96	95.34	96.43	96.84	96.30	96.98	96.71	96.71	96.57	96.71	96.57	96.43	96.02	95.75	95.88
L4	91.36	94.51	96.02	96.84	97.26	97.12	94.79	95.34	**97.39**	96.71	96.84	96.84	96.71	96.98	96.71	96.43	96.71	96.43	96.43	96.43	96.02

In addition, Table 2 gives more details on the different results in terms of recognition rates (IR). It takes a low value when no decomposition is performed (L=0). The lowest value of IR is 33.33% obtained for radius values [R_{In} R_{Ex}]=[1 2] and L=0.

Experiment #2

In this experiment, we use one ear image of each subject as a training set and remaining ear images as a testing set. Therefore, the training set contains 125 images, whereas the testing set contains 368 images.

From the results given by Figure 7, it can be seen that when the parameters [R_{In} R_{Ex}] are equal to [6 7], we obtain a good performance (EIR=7.88%) with level decomposition L = 4. For a fixed value of R_{In}, the error identification rate (EIR) decreases when the radius R_{Ex} increases. Moreover, the EIR varies inversely with the decomposition level L of the SPH method.

In addition, Table 3 gives more details on the different results in terms of identification rates (IR). It takes a low value when no decomposition is performed (L=0). The lowest value of IR is 41.85% obtained for radius values [R_{In} R_{Ex}]=[1 2] and L=0.

Experiment#3

In this sub-section a comparative study has been made between the proposed method and some recent existing work, to demonstrate its effectiveness. The same protocol is evaluated, using two ear images as the training set.

Table 4 summarizes the identification rates of the proposed method and the different algorithms recently proposed. Using the D-LDCP descriptor, the proposed approach gives better results compared to four algorithms described in Table 4, and offers the same performance as [9] in terms of accuracy. However, it gives a lower performance than that proposed in [19], [11](MS-BSIF), [11](DMS-BSIF), [17] and [18] with a difference of 2.21%, 0.69%, 0.42%, 0.38% and 0.21% , respectively. As a result, the proposed method achieves the best performance using the D-LDCP descriptor.

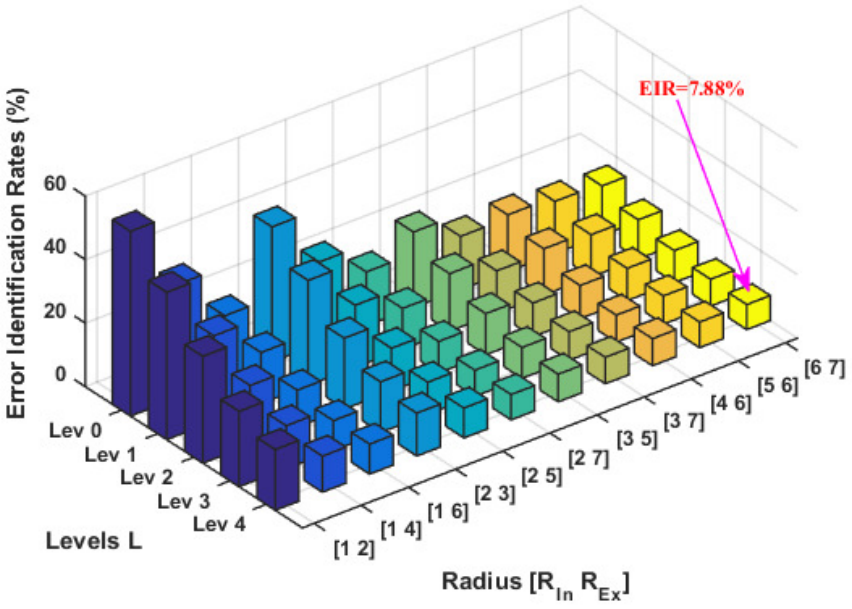

Figure 7: Error Identification Rates using LDCP descriptor.

Table 3: Recognition rates using LDCP descriptor.

	[1, 2]	[1, 3]	[1, 4]	[1, 5]	[1, 6]	[1, 7]	[2, 3]	[2, 4]	[2, 5]	[2, 6]	[2, 7]	[3, 4]	[3, 5]	[3, 6]	[3, 7]	[4, 5]	[4, 6]	[4, 7]	[5, 6]	[5, 7]	[6, 7]
L0	41.85	52.26	64.22	73.82	79.80	81.88	57.52	65.76	74.28	79.53	83.06	71.11	76.18	79.98	83.24	78.62	82.16	84.60	83.61	84.78	84.42
L1	53.62	63.68	73.01	80.62	83.97	85.60	66.94	74.00	80.43	84.51	87.14	77.54	81.88	84.24	86.68	92.88	85.42	86.96	86.68	87.05	87.68
L2	66.39	73.64	81.25	86.23	88.41	89.86	72.72	83.15	86.59	89.31	90.13	84.60	87.05	89.13	89.40	89.04	89.58	90.31	89.86	90.31	89.95
L3	76.27	82.43	86.50	88.77	90.58	91.21	84.06	87.50	89.86	91.49	91.76	88.77	90.31	90.76	91.03	90.94	91.49	91.49	91.12	91.49	91.58
L4	80.89	86.05	88.68	90.13	90.58	91.12	86.50	89.04	90.58	91.30	92.03	90.31	91.30	91.49	91.67	91.21	91.49	92.03	91.94	91.85	**92.12**

Table 4: Summary of related and recent work using two ear images in the training set.

References	Feature extraction	Classifier	IITD-1
[9]	BSIF descriptor	K-NN	97.26
[10]	Improved BSIF descriptor	K-NN	97.39
[11]	DMS-BSIF	K-NN	**98.08**
	MS-BSIF	K-NN	**97.81**
[17]	Non linear curevelt features	K-NN	**97.77**
[18]	Local principal independent components	Inner product classifier	**97.60**
[19]	Geometric measurements	SVM	**99.60**
[20]	2-D quadrature filter	Hamming distance	96.53
[21]	Sparse representation of local gray level orientations	Sparse representation	97.07
[22]	Orthogonal Log-Gabor filter pair	K-NN	96.27
Our	D-LDCP	K-NN	97.39
	LDCP	K-NN	96.16

3.2.2 IIT Delhi-2 database

The IIT Delhi-2 database contains 793 ear images acquired from 221 different subject.

Experiment #1

This experiment is conducted on IIT Delhi-2 database to explore the effect on identification performance of parameters R_{In}, R_{Ex} and L to find the optimal values. Two ear images of each person are used as a training set and the remaining ear images as a testing set. Therefore, the training set contains 442 images, whereas the testing set contains 351 images.

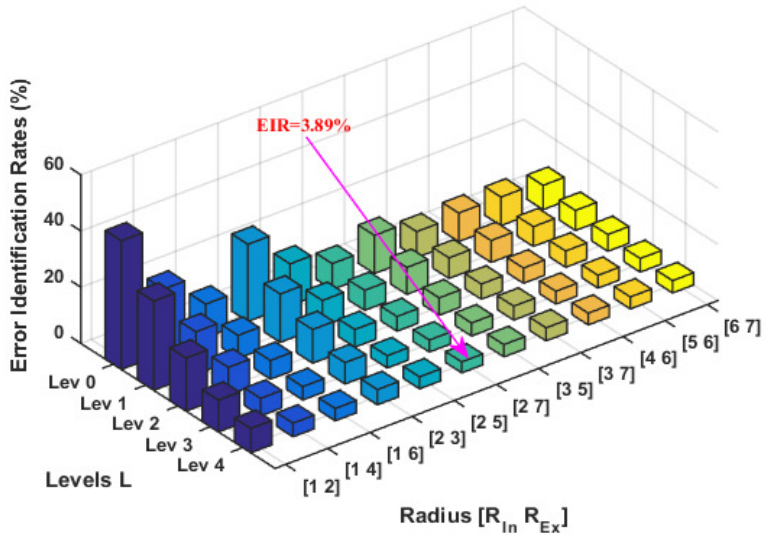

Figure 8: Error Identification Rates using LDCP descriptor.

From the results given by Figure 8, it can be seen that when the parameters [R_{In} R_{Ex}] are equal to [2 7], we obtain a good performance (EIR=3.89%) with level decomposition L=4.

For a fixed value of R_{In}, the error identification rate (EIR) decreases when the radius R_{Ex} increases. Moreover, the EIR is inversely proportional to the radius R_{Ex} and varies inversely with the decomposition level L of the SPH method.

Table 5: Recognition rates using LDCP descriptor

	[1, 2]	[1, 3]	[1, 4]	[1, 5]	[1, 6]	[1, 7]	[2, 3]	[2, 4]	[2, 5]	[2, 6]	[2, 7]	[3, 4]	[3, 5]	[3, 6]	[3, 7]	[4, 5]	[4, 6]	[4, 7]	[5, 6]	[5, 7]	[6, 7]
L0	54.23	66.76	77.11	83.48	88.13	89.36	72.55	79.87	85.19	89.17	90.79	83.48	86.42	89.55	90.79	88.03	89.84	90.69	89.84	90.98	91.36
L1	68.28	78.35	84.90	89.74	91.83	92.88	82.81	86.61	90.79	91.93	92.88	89.08	90.41	91.83	92.69	91.83	92.21	92.88	92.97	92.78	92.88
L2	81.10	86.23	90.12	93.35	93.35	94.68	88.03	92.40	93.73	94.02	94.78	92.59	93.83	94.30	94.59	94.11	94.59	94.59	94.21	94.40	94.02
L3	88.70	92.12	93.73	94.87	95.25	95.25	92.40	94.59	95.63	95.73	95.73	95.54	95.35	95.25	95.06	95.35	95.44	95.16	95.44	95.35	95.16
L4	91.36	94.30	95.16	95.25	95.63	95.54	94.59	95.25	95.73	96.11	96.11	95.82	95.63	95.44	95.35	95.82	95.82	95.44	95.73	95.73	95.63

In addition, Table 5 gives more details on the different results in terms of recognition rate (IR). It takes a low value when no decomposition is performed (L=0). The lowest value of IR is 54.23% obtained for radius values [R_{In} R_{Ex}]=[1 2] and L=0.

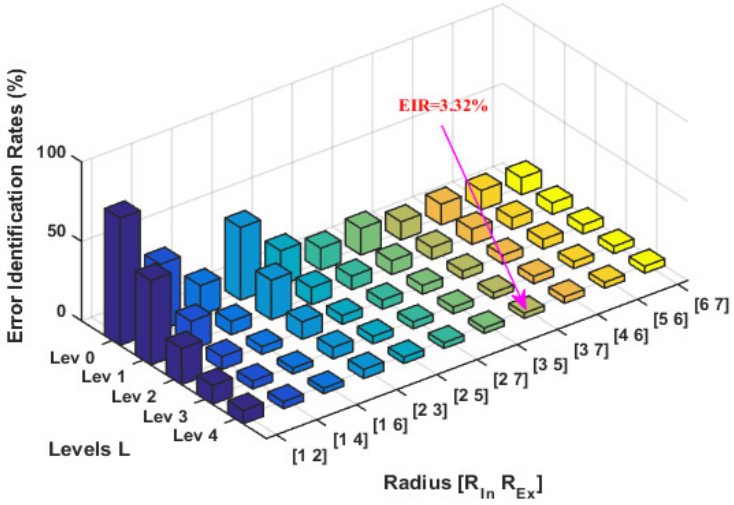

Figure 9: Error Identification Rates using D-LDCP descriptor.

From, the Figure 9 it can be seen that the use of D-LDCP feature vectors allows to reduces the EIR compared to results obtained using LDCP feature vectors (see Figure 8). In addition, it can be seen that when the parameters [R_{In} R_{Ex}] are equal to [3 7], a lowest error (EIR=3.32%) is obtained with level decomposition L = 4.

Table 6 gives more details on the different results in terms of recognition rate (IR). It takes a low value when no decomposition is performed (L=0). The lowest value of IR is 19.94% obtained for radius values [R_{In} R_{Ex}]=[1 2] and L=0.

Experiment #2

In this case, we use one ear image of each subject as a training set and remaining ear images as a testing set. Therefore, the training set contains 221 images, whereas the testing set contains 572 images.

From the results given by Figure.10, It can be seen that when the parameters [R_{In} R_{Ex}] are equal to [2 7], we obtain a good performance (EIR=9.62%) with level decomposition L=4.

For a fixed value of R_{In}, the error identification rate (EIR) decreases when the radius R_{Ex} increases. Furthermore, the EIR varies inversely with the decomposition level L of the SPH method.

Table 6: Recognition rates using D-LDCP descriptor

	[1, 2]	[1, 3]	[1, 4]	[1, 5]	[1, 6]	[1, 7]	[2, 3]	[2, 4]	[2, 5]	[2, 6]	[2, 7]	[3, 4]	[3, 5]	[3, 6]	[3, 7]	[4, 5]	[4, 6]	[4, 7]	[5, 6]	[5, 7]	[6, 7]
L0	19.94	38.27	58.50	72.08	81.39	83.29	54.32	68.38	78.44	83.76	86.23	77.49	83.29	86.80	88.60	85.57	86.99	87.37	87.84	88.60	89.36
L1	46.91	67.24	82.05	88.41	91.55	91.74	75.12	83.29	89.27	91.17	91.55	89.17	91.26	91.83	92.40	91.83	90.31	91.64	92.69	92.78	92.78
L2	77.30	86.42	91.93	94.40	95.06	94.97	88.98	90.88	94.49	93.92	94.21	94.11	94.49	94.68	94.87	94.30	94.40	94.40	93.92	95.06	94.30
L3	88.32	92.59	94.87	95.73	96.30	96.11	92.78	93.26	95.25	95.63	95.54	95.35	95.82	95.63	96.11	95.54	95.35	95.25	95.63	96.20	95.44
L4	91.64	93.92	95.63	96.01	96.58	96.39	94.21	94.21	96.01	96.30	96.39	95.63	96.58	96.11	**96.68**	96.01	95.92	95.92	96.30	**96.68**	95.63

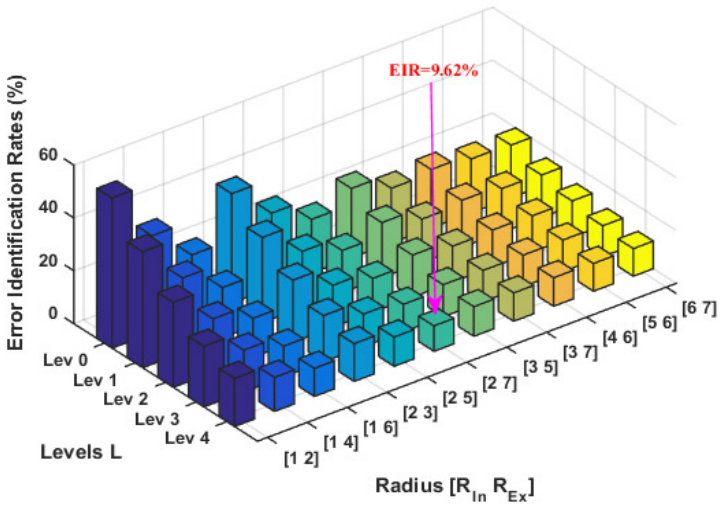

Figure 10: Error Identification Rates using LDCP descriptor.

Table 7: Recognition rates using LDCP descriptor

	[1, 2]	[1, 3]	[1, 4]	[1, 5]	[1, 6]	[1, 7]	[2, 3]	[2, 4]	[2, 5]	[2, 6]	[2, 7]	[3, 4]	[3, 5]	[3, 6]	[3, 7]	[4, 5]	[4, 6]	[4, 7]	[5, 6]	[5, 7]	[6, 7]
L0	43.24	53.26	63.17	71.10	76.11	78.15	58.62	65.15	71.68	76.11	79.72	68.82	73.72	76.40	79.31	75.06	77.91	79.72	79.60	80.36	80.07
L1	56.12	63.52	71.74	78.32	81.06	83.04	67.77	73.08	78.73	81.53	84.09	75.41	79.20	81.70	83.57	80.30	82.17	83.45	83.45	83.28	83.97
L2	67.89	73.37	79.78	83.62	85.26	86.95	76.52	81.35	84.15	86.01	87.24	82.23	84.09	86.07	86.42	85.61	86.01	86.60	86.31	83.83	86.36
L3	77.10	81.76	84.21	86.48	88.40	89.04	82.52	85.66	87.65	88.93	89.63	86.36	87.88	88.00	88.17	88.17	88.11	88.34	88.05	88.40	88.34
L4	81.82	85.20	87.12	88.40	89.39	89.80	85.72	87.59	88.93	89.86	90.38	88.00	88.81	89.28	89.28	88.93	89.22	89.39	89.51	89.57	89.57

Table 7 gives more details on the different results in terms of identification rates (IR). It takes a low value when no decomposition is performed (L=0). The lowest value of IR is 43.24% obtained for radius values $[R_{In}\ R_{Ex}]=[1\ 2]$ and L=0.

Experiment #3

In this experiment, the proposed method is compared with recent existing methods for the IIT-Delhi-2 ear database. Furthermore, the same protocol of the comparative methods is used to carry out the corresponding experiment.

Table 4 summarizes the identification rates of the proposed method and the different algorithms recently proposed. Using the D-LDCP descriptor, the proposed approach gives better results compared to four algorithms described in Table 4, and offers the same performance as [9] in terms of accuracy. However, it gives a lower performance than that proposed in [19], [11](MS-BSIF), [11](DMS-BSIF), [17], [18] with a difference of 2.21%, 0.69%, 0.42%, 0.38% and 0.21% , respectively. As a result, the proposed method achieves the best performance using the D-LDCP descriptor.

Table 8 summarizes the identification rates of eight comparative methods and the proposed method. It can be seen that, the identification rate given by D-LDCP descriptor is better than that proposed in [17], [20] and [22] with a difference of 0.46%, 0.60% and 0.75%, respectively.

3.3. USTB-1

It contains 185 images of 60 subjects, with at least three images of each subject. It comes with automatically standardized and cropped ear images of 150 x 80 pixels.

Experiment #1

In this experiment, we will study the effect of parameters R_{In}, R_{Ex} and L on the error recognition rate. We choose R_{In} changed from 1 to 6 and R_{Ex} varies from $R_{In}+1$ to 7. Meanwhile, the parameter level L is changed from 0 to 3.

Table 8: Summary of related and recent work using two ear images in the training set.

References	**Feature extraction**	**Classifier**	**IITD-2**
[9]	BSIF descriptor	K-NN	97.34
[10]	Improved BSIF descriptor	K-NN	97.63
[11]	DMS-BSIF	K-NN	97.72
	MS-BSIF	K-NN	97.53
[17]	Non linear curevelt features	K-NN	96.22
[18]	Local principal independent components	Inner product classifier	97.20
[20]	2-D quadrature filter	Hamming distance	96.08
[21]	Sparse representation of local gray level orientations	Sparse representation	97.73
[22]	Orthogonal Log-Gabor filter pair	K-NN	95.93
Our	D-LDCP	K-NN	96.68
	LDCP	K-NN	96.11

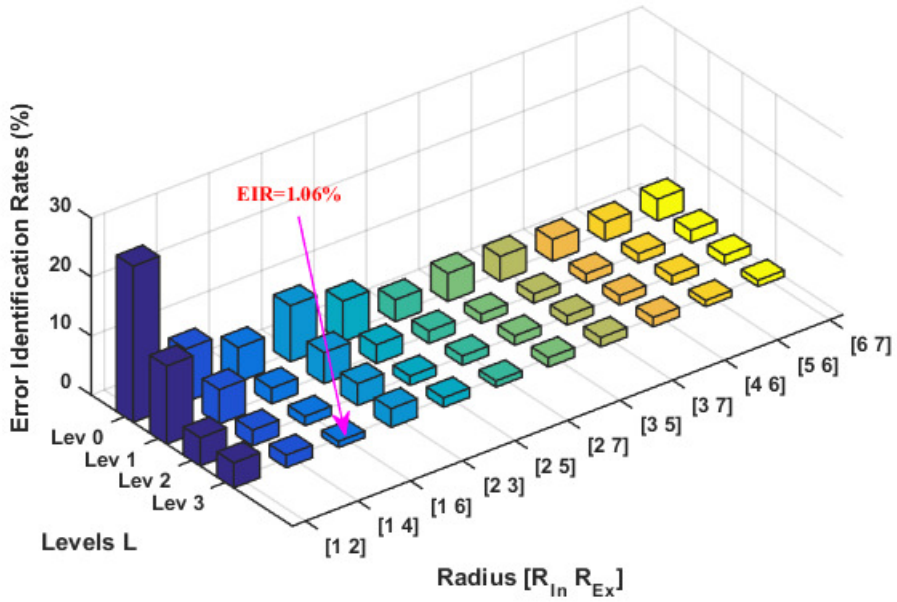

Figure 11: Error Identification Rates using LDCP descriptor.

Two ear images of each subject are used as a training set and the remaining ear images as a testing set. Therefore, the training set contains 120 images, whereas the testing set contains 65 images.

From the results given by Figure 11, It can be seen that when the parameters [R_{In} R_{Ex}] are equal to [1 6], we obtain a good performance (EIR=1.06%) with level decomposition L=3. For a fixed value of R_{In}, the error identification rate (EIR) decreases when the radius R_{Ex} increases. Moreover, the EIR is inversely proportional to the radius R_{Ex}. It should be noted that, the EIR increases as R_{Ex} increases and at the same time, R_{In} changes to a higher value. Moreover, the EIR varies inversely with the decomposition level L of the SPH method.

In addition, Table 9 gives more details on the different results in terms of recognition rate (IR). It takes a low value when no decomposition is performed (L=0). The lowest value of IR is 73.54% obtained for radius values [R_{In} R_{Ex}]=[1 2] and L=0.

From, the Figure 12 it can be seen that the use of D-LDCP feature vectors allows to reduces the EIR compared to results obtained using LDCP feature vectors (see Figure 11). In addition, it can be seen that when the parameters [R_{In} R_{Ex}] are equal to [3 7], a lowest error (EIR=1.06%) is obtained with level decomposition L = 1.

Table 9: Recognition rates using LDCP descriptor

	[1, 2]	[1, 3]	[1, 4]	[1, 5]	[1, 6]	[1, 7]	[2, 3]	[2, 4]	[2, 5]	[2, 6]	[2, 7]	[3, 4]	[3, 5]	[3, 6]	[3, 7]	[4, 5]	[4, 6]	[4, 7]	[5, 6]	[5, 7]	[6, 7]
L0	73.54	83.60	91.01	92.06	94.18	95.24	90.48	89.95	93.12	96.30	96.30	91.01	95.24	94.71	95.77	94.71	96.30	96.83	96.83	96.83	96.30
L1	86.77	92.59	94.18	97.88	97.35	97.35	94.71	95.24	96.83	97.88	97.88	95.77	98.41	98.41	98.41	98.41	98.41	98.41	98.41	98.41	97.88
L2	95.24	96.30	97.35	98.41	98.41	98.41	96.30	96.83	98.41	98.41	98.41	97.35	98.41	98.41	98.41	**98.94**	98.41	98.41	98.41	98.41	98.41
L3	95.77	96.83	97.88	98.41	**98.94**	**98.94**	96.83	97.88	98.41	98.41	**98.94**	97.88	98.41	98.41	98.41	**98.94**	98.41	**98.94**	**98.94**	**98.94**	**98.94**

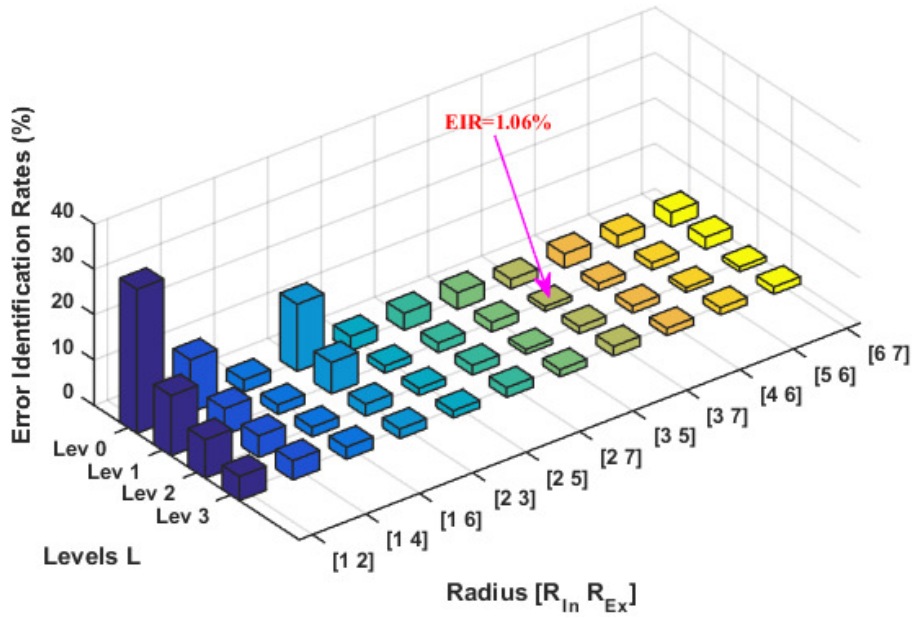

Figure 12: Error Identification Rates using D-LDCP descriptor

Table 10: Recognition rates using LDCP descriptor

	[1, 2]	[1, 3]	[1, 4]	[1, 5]	[1, 6]	[1, 7]	[2, 3]	[2, 4]	[2, 5]	[2, 6]	[2, 7]	[3, 4]	[3, 5]	[3, 6]	[3, 7]	[4, 5]	[4, 6]	[4, 7]	[5, 6]	[5, 7]	[6, 7]
L0	68.25	79.37	88.36	94.18	97.35	96.30	85.19	93.65	96.83	97.35	96.30	93.12	96.30	97.88	97.88	97.35	96.83	97.35	97.35	97.35	96.83
L1	86.77	88.36	94.18	94.18	97.88	95.77	93.12	96.30	98.41	97.88	97.88	96.83	97.88	98.41	**98.94**	97.88	98.41	97.88	98.41	97.88	97.35
L2	91.53	94.18	95.24	96.30	97.88	97.88	97.35	96.83	98.41	98.41	97.88	96.83	**98.94**	97.88	98.41	98.41	98.41	97.88	**98.94**	98.41	**98.94**
L3	94.71	95.24	95.77	96.30	97.35	97.88	97.88	97.35	98.41	97.35	97.88	97.88	98.41	98.41	97.88	**98.94**	98.41	97.88	98.41	98.41	98.41

Table 10 gives more details on the different results in terms of identification rate (IR). It takes a low value when no decomposition is performed (L=0). The lowest value of IR is 68.25% obtained for radius values $[R_{In}\ R_{Ex}]=[1\ 2]$ and L=0.

Experiment #2

In this experiment, one ear image of each person is used as a training set and the remaining ear images as a testing set. Therefore, the training set contains 60 images, whereas the testing set contains 125 images.

From the results given by Figure 13, It can be seen that when the parameters $[R_{In}\ R_{Ex}]$ are equal to [3 7], we obtain a good performance (EIR=3.52%) with level decomposition L=3. For a fixed value of R_{In}, the error identification rate (EIR) decreases when the radius R_{Ex} increases. Moreover, the EIR varies inversely with the decomposition level L of the SPH method.

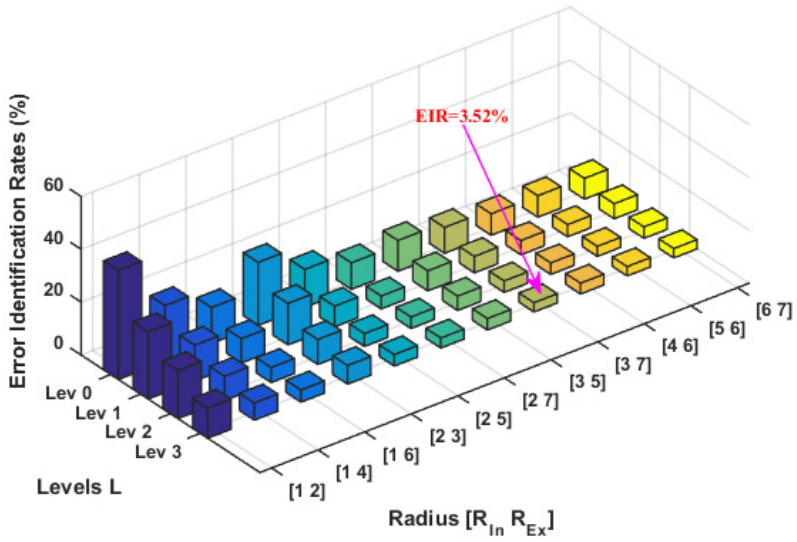

Figure 13: Error Identification Rates using LDCP descriptor

Table 11 gives more details on the different results in terms of identification rate (IR). It takes a low value when no decomposition is performed (L=0). The lowest value of IR is 58.81% obtained for radius values [R_{In} R_{Ex}]=[1 2] and L=0.

Table 11: Recognition rates using LDCP descriptor

	[1, 2]	[1, 3]	[1, 4]	[1, 5]	[1, 6]	[1, 7]	[2, 3]	[2, 4]	[2, 5]	[2, 6]	[2, 7]	[3, 4]	[3, 5]	[3, 6]	[3, 7]	[4, 5]	[4, 6]	[4, 7]	[5, 6]	[5, 7]	[6, 7]
L0	58.81	69.11	78.86	82.11	86.45	88.35	76.15	80.49	85.64	89.43	89.97	83.47	88.35	88.62	90.24	89.70	91.87	91.60	91.87	92.68	92.14
L1	74.53	81.03	86.45	90.79	90.79	92.68	84.28	89.16	92.14	92.95	95.12	89.97	92.41	93.50	93.77	93.77	94.58	94.85	95.39	95.12	94.31
L2	81.84	87.53	90.79	93.22	94.31	95.12	90.51	92.14	95.12	95.12	95.66	92.95	94.58	95.12	95.66	95.66	95.66	95.66	96.21	96.21	95.66
L3	88.62	90.79	93.50	95.39	95.39	95.66	92.95	94.58	95.66	95.66	96.21	94.85	95.66	95.93	96.48	96.21	95.93	**96.75**	96.48	96.21	96.48

Experiment #3

In this subsection, the proposed method is compared with recent existing methods for the USTB-1 ear database. Moreover, the same protocol of the comparative methods is used to carry out the corresponding experiment. Table 12 summarizes the identification rates of three comparative methods and the proposed method. It can be seen that, the identification rate given by D-LDCP feature is the highest compared to the BSIF and MS-BSIF descriptors.

Table 12: Recognition rates using LDCP descriptor

References	Feature extraction	Classifier	USTB-1
[9]	BSIF descriptor	K-NN	98.46
[10]	Improved BSIF descriptor	K-NN	98.97
[11]	DMS-BSIF	K-NN	**99.74**
	MS-BSIF	K-NN	98.41
Our	D-LDCP	K-NN	98.94
	LDCP	K-NN	98.94

4. Conclusion

A novel ear representation for ear recognition has been proposed in this paper. This representation use DCP descriptor, SPH decomposition and LDA subspace method whose purpose is to extract the discriminating features of ear image. Extensive experiments show that, compared to the existing methods in the literature, the proposed method effectively increases the accuracy of ear identification with a remarkably lower computational cost. In addition, the D-LDCP feature representation achieves rank-1 identification rates of 97.39%, 96.68% and 98.94% for IIT Delhi-1, IIT Delhi-2 and USTB-1 databases, respectively. It confirms its performances than the state-of-the-art in terms of recognition accuracy. In the future, we will focus on evaluating the performances of the proposed approach on verification case.

Acknowledgment

The authors are thankful to the anonymous reviewers and editor for their valuable suggestions, which helped us to improve the quality of the manuscript..

References

[1] D. Maltoni *et al.*, Handbook of Finger print Recognition, 1st ed., Springer- Verlag, New York (2003).
[2] A. V. Iannarelli, Ear identification, Forensic Identification Series, Paramount Publishing Company, Fremont, California, USA, 1989.
[3] B. Victor, K. Bowyer, and S. Sarkar, An evaluation of face and ear biometrics, in Proceedings of the 16th International Conference on Pattern Recognition (ICPR). Vol.01, pp.429-432, 2002.
[4] H. Zhang and Z. Mu, Compound structure classifier system for ear recognition, in Proceedings of the IEEE International Conference on Automation and Logistics (ICAL). pp.2306-2309, Qingdao (China), 2008.
[5] Z. Xie and Z. Mu, Ear recognition using LLE and IDLLE algorithm, in Proceedings of the 19th IEEE International Conference on Pattern Recognition (ICPR). pp.1-4, Tampa (USA), 2008.
[6] Ž. Emeršič, V. Štruc and P. Peer, Ear recognition: more than a survey, Neurocomputing Journal, Vol. 255, pp.26–39, Elsevier, 2017.
[7] K. Dewi and T. Yahagi, Ear photo recognition using scale invariant keypoints, in Proceedings of the Computational Intelligence, pp.253–258, 2006.
[8] S. Prakash, P. Grupta, An efficient ear recognition technique invariant to illumination and pose, Telecommun, Syst. 52 (3), pp. 1435-1448, 2013.
[9] A. Benzaoui, A. Hadid and A. Boukrouche, Ear biometric recognition using local texture descriptors, Journal of electronic imaging vol.23 (5), pp. 0530081-05300812, 2014.
[10] A. Benzaoui, I. Adjabi and A. Boukrouche, Experiments and improvements of ear recognition based on local texture descriptors, Journal of Optical engineering vol.56 (4), pp. 0431091-04310913, 2017.
[11] H. Doghmane, A. Boukrouche and L. Boubchir, A novel discriminant multi-scale representation for ear recognition. International Journal of Biometrics, pp 50-66, 2019.
[12] V. D. M. Nhat, S. Y Lee and H. Y. Youn, Whitened LDA for face recognition, CIVR'07, Amsterdam, the Netherlands, 2007.
[13] C. Ding, J. Choi, D. Tao, and L. S. Davis, Multi-directional multi-level dual-cross patterns for robust face recognition, IEEE transactions on pattern analysis and machine intelligence, 38(3): 518–531, 2016.
[14] S. Lazebnik, C. Schmid and J. Ponce, Beyond bags of features: spatial pyramid matching for recognizing natural scene categories, in Proc. of the 2006 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Vol. 2, pp. 2169–2178, 2006.
[15] A. Kumar, IIT Delhi Ear Image Database version1.0., New Delhi, India, http://www4.comp. polyu.edu. hk /~csajaykr/IITD/Database_Ear.htm, 2007.
[16] Z. Mu Z, USTB Ear Image Database, Beijing, China, http://www1.ustb.edu.cn/resb/en/index.htm, 2009.
[17] A. Basit and M. Shoaib, A human ear recognition method using non-linear curvelet feature subspace, *Int. J. Comput. Math* **91**(3), pp.616-624, 2014.
[18] Mamta and H. Madasu, Robust ear based authentication using Local Principal Independent Components, *Expert Syst. Appl*. **40**(16), pp. 6478- 6490, 2013.
[19] I. Omara et al., A novel geometric feature extraction method for ear recognition, *Expert Syst. Appl.* 65, pp. 127-135, 2016.

[20] T. S. Chan and A. Kumar, Reliable ear identification using 2-D quadrature filters, *Pattern Recognit. Lett*. **33**(14), pp. 1870-1881, 2012.
[21] A. Kumar and T. S. T. Chan, Robust ear identification using sparse representation of local texture descriptors, *Pattern Recognition* **46**(1), pp. 73-85, 2013.
[22] A. Kumar and C. Wu, Automated human identification using ear imaging, *Pattern Recognit*. **45** (3), pp. 956-968, 2012.