

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/320911734>

# Ear Recognition based on Embryological Inspirations

Conference Paper · October 2016

CITATIONS

0

READS

43

3 authors:



**Amir Benzaoui**

Université de Bouira

24 PUBLICATIONS 165 CITATIONS

[SEE PROFILE](#)



**Adjabi Insaf**

Université de Bouira

3 PUBLICATIONS 17 CITATIONS

[SEE PROFILE](#)



**Boukrouche Abdelhane**

Université 8 mai 1945 - Guelma

74 PUBLICATIONS 307 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



PAPR Reduction in Multicarrier Modulation (MCM) Systems [View project](#)



Ear Biometric [View project](#)

# Ear Recognition based on Embryological Inspirations

Amir Benzaoui <sup>a,b</sup>, Insaf Adjabi <sup>b</sup>, and Abdelhane Boukrouche <sup>b</sup>

<sup>a</sup> Department of Electrical Engineering, Faculty of Science and Applied Sciences  
University of Akli Muhand Oulhadj – BOUIRA, Algeria

<sup>b</sup> Laboratory of Inverse Problems, Modeling, Information, and Systems (PIMIS)  
Department of Electronics and Telecommunications, Faculty of Sciences and Technology  
University of May 08<sup>th</sup> 1945 – Guelma, Algeria  
[amirbenzaoui@gmail.com](mailto:amirbenzaoui@gmail.com)

**Abstract**— Morphological shape of the human ear presents a rich and stable information embedded on the curved 3D surface, which has invited lot attention from the forensic and engineer scientists in order to differentiate and recognize people. However, recognizing identity from morphological shape of the human ear in unconstrained environments, with insufficient and incomplete training data, dealing with strong person-specificity, and high within-range variance, can be very challenging. In continuation to our previous works, we propose to use some embryological inspirations of the human ear in order to find the autonomous components and also their locations on which we can achieve large inter-individual variations. This is especially beneficial to our research as it provides some advices on the possible changes that can be observed in the external structure of the ear. We experiment with two publicly available databases, which are IIT Delhi-1 and IIT Delhi-2, consisting of several ear benchmarks of different natures under varying conditions and imaging qualities. The experiments show excellent results beyond the state-of-the-art.

**Index Terms** — Biometrics, Identity Recognition, Ear, Embryology, Texture Analysis, BSIF.

## I. INTRODUCTION

Biometrics is the science which makes it possible to recognize the identity of a person on the basis of her physiological, chemical, or behavioral characteristics, such as: face, iris, fingerprint, odor, DNA, gait, or electronic signature...etc [01]. With the need for solid techniques of identity recognition in the critical applications, such as: secure access control, passing of international borders, and legal applications, biometrics is positioned as a viable technology that can be integrated into the large-scale identity's management systems. Biometric systems operate under the principle that most biological characteristics of the human beings are distinctive for each individual, can be acquired in a reliable manner using suitable sensors, and can be represented in a digital format [02]. Thus, these systems can be considered

as a pattern recognition engine and may be incorporated in various markets.

Several human characteristics were studied and tested in literature, these modalities can be further subdivided into different sub-categories according to their respective position in the human body, such as: (i) attributes of the hand region (fingerprint, hand geometry,...), (ii) attributes of the facial region (face and ear), (iii) attributes of the eye region (iris and retina), (iv) behavioral attributes (gait, electronic signature...), and (v) chemical-medical attributes (bone, odor, DNA,...) [03]. Although that DNA, iris, and fingerprint are considered among the highly reliable modalities, they are unfortunately based on the cooperation of the subject-sensor [04]. Within the framework of our research work, we concentrated by the facial attributes, namely: the face and the ear, since they are often effective and require no cooperation or prior knowledge of the user. For example, the identification of criminals and suspects in the airports and in other public sectors by surveillance videos is a typical scenario in which appears the importance of the facial attributes. In addition, this model of recognition is the most accepted by humans due to its natural character [05].

Research in the field of facial recognition is motivated not only by the fundamental challenges of this problem but also by the several numbers of practical applications where the human identification is necessary. Face recognition, as one of the pilot biometric technologies, has become increasingly significant due to the fast progress in technologies, such as: internet, smart phones, and digital cameras; as well as the augmentation in the requirement of security. Face recognition has several advantages compared to other modalities: it is natural, not-intrusive, and requires less cooperation [06]. However, the conditions of acquisition, such as: position of the face to the camera, lighting, resolution of the camera,

partial occlusion, as well as natural aging, facial expressions, disguises, and spoofing attacks cause several changes in the facial appearance; those challenges affect negatively on the performances of recognition in a real world conditions [07].

In the other hand, the human ear is considered as a new viable class of biometrics with some additional advantages compared to the face. Indeed, the ear is rich in terms of characteristics; it has a stable structure which does not change significantly during aging and its form does not vary by facial expressions, it can be also captured from a distance and without any cooperation of the user. It can be used as a complementary part to the face in a multi biometric system [08]. Because of these qualities, the interest of ear recognition systems has been developed significantly in the last recent years. But, sometimes it can be hidden by hair, headphones, scarf, or loops [09]. Although researchers have developed several techniques in last recent years which use 2D intensity ear images, unfortunately, their performances are considerably degraded with changes in the pose and the conditions of the image [10].

So the goal of our research work is simply to find a suitable descriptor or an effective approach to represent and to model the human ear in a real context, in other words, we search to improve the performances of recognition in the unconstrained conditions. The rest of the paper is organized as follows: in the next section, we describe the anatomy of the human ear and some embryological studies. In Section 3, our proposed approach is presented. In Section 4, we present our experimental results by applying the proposed approach on IIT Delhi-1 and IIT Delhi-2 databases. Finally, a conclusion related to this work is given in Section 5.

## II. EAR ANATOMY & EMBRYOLOGY

The human ear is characterized by a rich structure which provides important information to differentiate between people; we can visualize 10 features and 37 sub-features from 2D ear image. 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 (small prominence of cartilage) [11].

The ability to identify people by using the outer ear shape was discovered for the first time by the French criminologist Alphonse Bertillon [12] and confirmed by the US police officer Alfred Iannarelli [13]. This latter has proposed the first automatic system of ear identification, based only on seven characteristics. The detailed structure of the ear is not only unique, but also stable because the appearance of the ear does not change during the life of the human being. In addition, the

acquisition of ear pictures does not require human cooperation, it is considered by most researchers as a non-invasive modality.

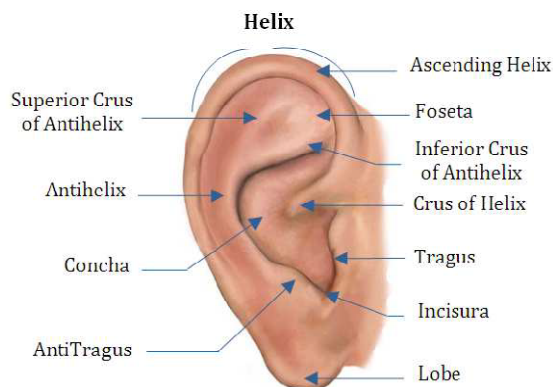


Fig. 1. Anatomy of the Human Ear.

The formation of the ear in the human embryo is usually taken as an individual development of its separate components. The identification of the components that make up the complex structure of the human ear is considered as a main concern in embryological studies of the ear. Indeed, the premise of local and independent structures in the auricular pavilion takes a great interest in the identification approaches. While embryologists assume and argue the exact identification of the hillock (component) which forms a specific element of the auricular pavilion, our concern is simply to identify the autonomous components and also their locations on which we can achieve large inter-individual variation. In addition, the study of malformations of the outer ear shape is considered as one of the main approaches to understanding their embryology. This is especially beneficial to our research as it also provides some advices on the possible changes that can be observed in the external structure of the ear.

The lobule was considered by most embryological researchers as the only component of the ear composed by grease rather than cartilage. The shape of the lobule can be changed from natural to an attached form, and in some ears, the lobule is almost nonexistent. In addition, it may be, in some cases, very discriminant. However, the lobule seems to be like the only part of the ear that continues the development and the deformation during aging of the human being.

Based on the anatomical and embryological inspirations of the ear, performed in this Section, in order to identify all the autonomous components and also their locations on which we can reach a large inter-individual variation, we suggested that the form the ear's lobe is considered as the component that changes significantly with aging, increasingly, the lobule is frequently found occluded or cluttered by earrings. So, we did not lend much attention to this unstable auricular component

and we considered this part as a semi-rigid region, the rest of the ear's regions was considered as a rigid region. Fig.2 shows an example of separation between the rigid and semi-rigid regions.

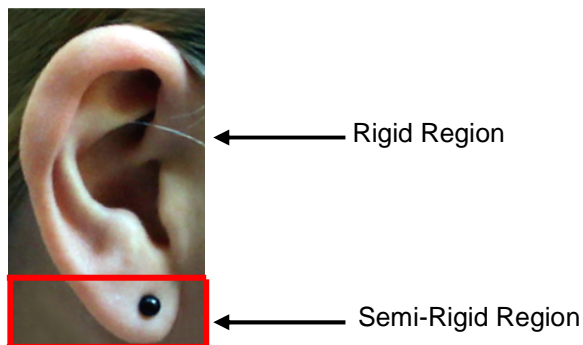


Fig. 2. Example of separation between Rigid and Semi-Rigid Regions.

### III. METHODOLOGY

The proposed biometric system requires two operational phases. The first is a training phase: it consists in recording the ear features of each individual in order to create his own biometric template; this is then stored in the database. The second is the test phase: it consists in recording the same features and comparing them to the biometric templates stored in the database. If the recorded data match a biometric template from the database, the individual in such case is considered identified. The proposed biometric system based on external shape of 2D ear imaging is described in the following modules:

#### A. Preprocessing

The objective of the pre-processing module is to prepare the source image representation in order to facilitate the task of the following steps and to improve the recognition performances. First, the ear imaging is converted into gray-scale image. Next, every gray-scale image is filtered by median filter to reduce noise. This filtered image is then adjusted to improve the image's contrast.

#### B. Feature Extraction

The choice of visual features to be extracted from normalized ear images and sent for classification plays an important role on the resulting recognition accuracy. In this paper, we have selected a number of recent local texture descriptors that are distinctive, do not require segmentation, and robust to occlusion, illumination variation, and weak lighting. In contrast to global image descriptors which compute features directly from the entire image, local descriptors representing the features in small local image patches have proven to be more effective in real-world conditions. In our previous work [14], we have tested and considered three recent local texture descriptors, namely LBP, LPQ, and BSIF.

**The Local Binary Patterns (LBP)** [15] is a texture analysis operator, introduced by *Ojala et al.* [16], is defined as a gray-scale invariant texture measure, and 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 are its discriminative power, computational simplicity, and tolerance against monotonic gray-scale changes. In the original LBP operator, the local patterns are extracted by thresholding the  $3 \times 3$  neighborhood of the eight neighbors of each pixel, from the original image, with the central value. All the neighbors are assigned the value 1 if they are greater than or equal to the current element and 0 otherwise, which represents a binary code of the central element. This binary code is converted into a decimal value by multiplying it with the given corresponding weights and is summed to obtain the LBP code for the central value. The histogram of these  $2^8 = 256$  different labels can then be used as a texture descriptor for further analysis. Recently, the LBP operator has been extended to use neighborhoods of different sizes in order to capture the large-scale structures which can be considered as the dominant patterns in the image.

**The Local Phase Quantization (LPQ)** [17] was proposed in order to tackle the relative sensitivity of LBP to blur, based on quantizing the Fourier transform phase in local neighborhoods. The phase can be shown to be a blur invariant property under certain commonly fulfilled conditions. In texture analysis, histograms of LPQ labels computed within local regions are used as a texture descriptor similarly to the LBP methodology. The LPQ descriptor has recently received wide interest in blur-invariant texture recognition. LPQ is insensitive to image blurring, and it has proven to be very efficient descriptor in pattern recognition from blurred as well as from sharp images.

**The Binarized Statistical Images Features (BSIF)** [18] was recently proposed for texture classification. Inspired by LBP and LPQ, the idea behind BSIF is to automatically learn a fixed set of filters from a small set of natural images, instead of using hand-crafted filters such as LBP and LPQ. BSIF applies learning, instead of manual tuning, to obtain statistically meaningful representation of the images, which enables efficient information encoding using simple element-wise quantization. Learning provides also an easy and flexible way to adjust the descriptor length and to adapt applications with unusual image characteristics. To characterize the texture properties within each image sub-region, the histograms of pixels' BSIF code values are then used. The value of each element (i.e. bit) in the BSIF binary code string is computed by binarizing the response of a linear filter with a threshold at zero. Each bit is associated with a different filter and the

desired length of the bit string determines the number of filters used. The set of filters is learnt from a training set of natural image patches by maximizing the statistical independence of the filter responses.

Fig.3 shows an example of normalized ear image and their corresponding LBP, LPQ, and BSIF representations, respectively.

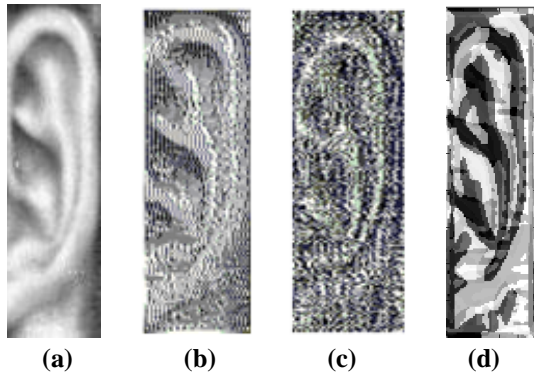


Fig. 3. (a) Example of normalized ear images and their corresponding (b) LBP, (c) LPQ, and (d) BSIF representations.

### C. Classification

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.

## IV. EXPERIMENTAL ANALYSIS

### A. Dataset

The IIT Delhi database [19] consists of ear images collected from students and staff at IIT Delhi, New Delhi (India). It was acquired on the IIT Delhi campus from October 2006 to June 2007 using a simple imaging setup. All the images were acquired from a distance in an indoor environment. The currently available database has two versions: the first version contains 493 images of 125 different subjects (IIT Delhi-1) and the second version contains 793 images of 221 different subjects (IIT Delhi-2). Each subject in the database has at least three ear images. The subjects are in the age range of 14 to 58 years. The resolution of the images is  $272 \times 204$  pixels. In addition to the original images, this database also comes with automatically normalized and cropped ear images of  $50 \times 180$  pixels in size.

### B. Settings

In the first protocol (“experiments using two images per person as training set”), two ear images from each person in the database were used as the “training set” and the remaining ear images of the same person (i.e., from one to four images) were used as the “test set.” Since most of the samples in the two databases have three images, we conducted three permutations and report the average rates (This is the most widely used protocol in literature). In the second protocol (“experiments using single image per person as training set”), only one image from each person was used as the “training set” and the remaining images of the same person (from two to five images) were used as the “test set.” We also did three permutations and report the average rates.

### C. Experiments

In our experiments, we tested the several parameters of the local texture descriptor i.e., the LBP was tested under several scale sizes (P,R), the LPQ descriptor was implemented with different radius, and the BSIF descriptor was implemented with several filter parameters (filter size and bit string:  $l$  and  $n$ ). Two types of SVM classifier were compared and used for classification: SVM with linear kernel and SVM with RBF kernel, we selected the two most widely used kernel functions, i.e., linear and RBF. The parameter in the RBF kernel function was empirically selected in this paper ( $\gamma = 0.0001$ ). Table 1 compares the rank-1 recognition results using two images in the training set and the two classifiers cited above applied to the IIT Delhi-1 and IIT Delhi-2 databases. Table 2 compares the rank-1 recognition results using single image per person in the training set and the two classifiers applied to the two databases.

The best results of the LPQ descriptor were given with a radius = 3 and the best results of the BSIF descriptor were obtained with a window size of  $15 \times 15$  pixels and 12 bits. As can be seen from Tables 1 and 2, the best recognition performances on both datasets are obtained with the Linear Kernel of the SVM classifier. The LBP descriptor seems to work better with large scale sizes than with smaller scale sizes. The results obtained indicate that LPQ outperforms LBP under different settings and BSIF gives the best performances. In addition, surprising and very interesting results appeared with the BSIF descriptor using only single image as a reference in the training set ( $> 90\%$ ).

**Table 1.** Best of Rank-1 Recognition Rates in the 1<sup>st</sup> Scenario

Descriptor	IIT Delhi-1		IIT Delhi-2	
	Linear kernel	RBF kernel	Linear kernel	RBF kernel
LBP (8,1)	77.50	73.94	74.74	71.51
LBP (8,2)	80.38	67.82	79.2	67.88
LBP (8,3)	81.62	62.55	80.53	64.96
LBP (16,2)	71.68	66.26	78.16	68.00
LPQ	79.84	79.79	79.77	79.87
BSIF	<b>96.68</b>	<b>96.68</b>	<b>97.31</b>	<b>97.31</b>

**Table 2.** Best of Rank-1 Recognition Rates in the 2<sup>nd</sup> Scenario

Descriptor	IIT Delhi-1		IIT Delhi-2	
	Linear kernel	RBF kernel	Linear kernel	RBF kernel
LBP (8,1)	63.32	63.32	57.28	57.23
LBP (8,2)	67.12	67.12	62.94	62.94
LBP (8,3)	68.21	68.21	64.63	64.63
LBP (16,2)	67.03	67.03	63.40	63.46
LPQ	73.82	73.82	71.5	71.5
BSIF	<b>91.26</b>	<b>91.26</b>	<b>90.29</b>	<b>90.29</b>

The previous results were obtained using the entire ear images, in other words, these results were obtained by the combination of the rigid and the semi-rigid regions. In order to justify our suggestion presented in Section 2: (*The form the ear's lobe is considered as the component that changes significantly with aging, increasingly, the lobule is frequently found occluded or cluttered by earrings. So, we did not lend much attention to this unstable auricular component and we considered this part as a semi-rigid region, the rest of the ear's regions was considered as a rigid region*), we have tested the performances of our approach by using only the components of the rigid region and eliminating the components of the semi-rigid region. Also, we have compared the new results with our previous results presented in Tables 1-2 (combination of rigid and semi-rigid regions). Table 3 compares the rank-1 recognition results using the rigid region and the combination between the rigid and semi-rigid regions.

**Table 3.** Best of Rank-1 Recognition Rates using the rigid region and the combination between the rigid and semi-rigid regions

	Two images per person		Single image per person	
	IIT Delhi-1	IIT Delhi-2	IIT Delhi-1	IIT Delhi-2
<b>Rigid and Semi-Rigid Regions</b>	96.68	97.31	91.26	90.29
<b>Rigid Region Only</b>	98	97.49	95.34	94.22

We can see from Table 3 that the new rates of classification obtained by the use of the rigid information only have been improved proportionally in comparison the results obtained by the use of the entire image (rigid + semi rigid information). So, with our embryological suggestion, we have presented some improvements on ear recognition by the illumination of the lobule component, this latter was considered as the not stable component which changes significantly with aging and some additional factors.

#### D. Comparison

For a comprehensive analysis, we also compared the results obtained against those of state-of-the-art automatic ear recognition. Table 4 shows and compares the average rank-1 recognition rate of the proposed method with some well known and recent feature extraction approaches under the same conditions and protocol of evaluation (the most widely used protocol in literature is "two images per person as training set"), to ear recognition using 2-D imaging. As can be seen, our proposed approach based on the BSIF descriptor and the introduce of the embryological inspirations show very competitive performance, outperforming other approaches under the same conditions.

**Table 4.** Rank-1 Comparison to Some Recent Works

Method	Delhi-1	Delhi-2
Force Field Transform [20]	74.93	66.67
Orthogonal log-Gabor filter pair [10]	96.27	95.93
2D quadratures filter [21]	96.53	96.08
Local principal independent components [22]	97.60	97.20
Non Linear Curvelet Transform [23]	97.77	96.22
<b>Our Approach</b>	<b>98</b>	<b>97.49</b>

#### V. CONCLUSION

In our investigation, we have successfully implemented a feature extraction approach for automated 2D ear description and recognition. We have introduced the use of the local texture descriptors, which are inspired from the statistics of natural images and produce binary codes. In contrast to global approaches which compute features directly from the entire image, local descriptors representing the features in small local image patches have proven to be more effective in real-world conditions. In addition, we have proposed to use some embryological inspirations of the human ear in order to find the autonomous components and also their locations on which we can achieve large inter-individual variation. Because, this is especially beneficial to our research as it provides some advices on the possible changes that can be observed in the external structure of the ear. A series of experimental

evaluations on IIT Delhi-1 and IIT Delhi-2 databases show that this implemented approach of feature extraction, based on BSIF descriptor and the embryological inspirations, given very significant improvements in the recognition performances.

#### REFERENCES

- [01] V.B. Nikolaos, N.P. Konstantinos, and E.M. Tzanakou: "Biometrics: Theory, Methods, and Applications". Eds., Wiley Online Library, 2009.
- [02] D.N. Kisku, P. Gupta, and J.K. Sing: "Advances in Biometrics for Secure Authentication and Recognition". Eds, CRC Press Taylor & Francis Group, 2013.
- [03] S. Chauhan, A.S. Arora, and A. Kaul: "A Survey of Emerging Biometric Modalities". *Procedia Computer Science* (Elsevier). Vol.02, pp.213-218, 2010.
- [04] I. Nigam, M. Vatsa, and R. Singh: "Ocular Biometrics: A Survey of Modalities and Fusion Approaches". *Information Fusion* (Elsevier). Vol.2, pp.01-35, 2015.
- [05] A. Benzaoui, H. Bourouba, and A. Boukrouche: "System for Automatic Faces Detection". In *Proceedings of the 3<sup>rd</sup> International IEEE Conference on Image Preprocessing, Theory, Tools, and Applications (IPTA)*, pp.354-358, 2012.
- [06] 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.
- [07] A. Benzaoui and A. Boukrouche: "1DLBP and PCA for Face Recognition". In *Proceedings of the 11<sup>th</sup> International IEEE Symposium on Programming and Systems (ISPS)*, pp.07-11, 2013.
- [08] J. Zhou, S. Cadavid, and M.A. Mottaleb: "Exploiting Color SIFT Features for 2d Ear Recognition". In *Proceedings of the 11<sup>th</sup> International IEEE Conference on Image Processing (ICIP)*, pp.553-556, 2011.
- [09] A. Abaza and T. Boulrai: "On Ear-Based Human Identification in the Mid-Wave Infrared Spectrum". *Image and Vision Computing* (Elsevier). Vol.31, pp.640-648, 2013.
- [10] A. Kumar and C. Wu: "Automated Human Identification using Ear Imaging". *Pattern Recognition* (Elsevier). Vol.45, pp.956-968, 2012.
- [11] S. Prakash and P. Gupta: "Ear Biometrics in 2D and 3D". Eds., *Augmented Vision and Reality Series* (Springer), 2015.
- [12] A. Bertillon: "La photographie Judiciaire, avec un appendice sur la classification anthropométrique". Technical Report, Gauthier-Villars, Paris, 1980.
- [13] A. Iannarelli: "Ear Identification". *Forensic Identification Series*, Paramount Publishing Company, Fremont, California, 1989.
- [14] A. Benzaoui, A. Hadid, and A. Boukrouche: "Ear biometric recognition using local texture descriptors". *Journal of Electronic Imaging (SPIE-JEI)*. Vol.23, No.05, 2014.
- [15] T. Ahonen, A. Hadid, and M. Pietikainen: "Face description with local binary patterns: application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*. Vol.28, No.12, pp.2037-2041, 2006.
- [16] T. Ojala, M. Pietikainen, and T. Maenpa: "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*. Vol.24, No.7, pp.971-987, 2002.
- [17] V. Ojansivu and J. Heikkil: "Blur insensitive texture classification using local phase quantization". In *Proceedings of the 3<sup>rd</sup> International Conference on Image and Signal Processing (ICSIP)*, Springer-Verlag, pp.236-243, 2008.
- [18] J. Kannala and E. Rahtu: "BSIF: binarized statistical image features". In *Proceedings of the International Conference on Pattern Recognition (ICPR)*, pp.1363-1366, 2012.
- [19] A. Kumar: "IIT Delhi ear database" New Delhi, India, 2007: [http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database\\_Ear.htm](http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Ear.htm).
- [20] D. Hurley, M. Nixon, and J. Carter: "Force field energy functional for image feature extraction". *Image Vision and Computing Journal* (Elsevier). Vol.20, pp.429-432, 2002.
- [21] T.S. Chan and A. Kumar: "Reliable ear identification using 2-D quadrature filters". *Pattern Recognition Letters* (Elsevier). Vol.33, pp.1870-1881, 2012.
- [22] Mamta and H. Madasu: "Robust ear based authentication using local principal independent components". *Expert Systems and Applications* (Elsevier). Vol.40, pp.6478-6490, 2013.
- [23] A. Basit and M. Shoaib: "A human ear recognition method using nonlinear curvelet feature subspace," *International Journal of Computer Mathematics* (Taylor & Francis). Vol.91, pp.616-624, 2014.