Telecommunication Systems manuscript No.
(will be inserted by the editor)

An Efficient Ear Recognition Technique Invariant to Illumination and Pose

Surya Prakash · Phalguni Gupta

Received: date / Accepted: date

Abstract This paper presents an efficient ear recognition technique which derives benefits from the local features of the ear and attempt to handle the problems due to pose, poor contrast, change in illumination and lack of registration. It uses (1) three image enhancement techniques in parallel to neutralize the effect of poor contrast, noise and illumination, (2) a local feature extraction technique (SURF) on enhanced images to minimize the effect of pose variations and poor image registration. SURF feature extraction is carried out on enhanced images to obtain three sets of local features, one for each enhanced image. Three nearest neighbor classifiers are trained on these three sets of features. Matching scores generated by all three classifiers are fused for final decision. The technique has been evaluated on two public databases, namely IIT Kanpur ear database and University of Notre Dame ear database (Collections E). Experimental results confirm that the use of proposed fusion significantly improves the recognition accuracy.

Keywords Biometrics · Ear Recognition · Image Enhancement · Fusion

1 Introduction

Biometrics deals with the recognition of a human using his or her inherent biometric characteristics which may be of physiological or behavioural in nature. Few examples of physiological biometrics are face, ear, iris, fingerprint, hand geometry, vein patterns, palm print etc whereas behavioural biometrics include signature, voice, gait pattern, key-strokes etc. There exists a number of systems developed based on these biometric traits and tested in real world applications. Among the various physiological biometric traits, ear has received much attention in recent years as it has been found to be a reliable biometrics for human recognition [4].

Surya Prakash* · Phalguni Gupta
Department of Computer Science & Engineering,
Indian Institute of Technology Kanpur, Kanpur-208016, India.
Tel.: +91-512-259-7579, Fax: +91-512-259-7586
E-mail: *psurya@cse.iitk.ac.in

In [17], Iannarelli has proposed a manual ear based recognition system. This system has used twelve features of the ear which are the distances between specific ear features that are measured manually. It has used $10,000$ ear images to find the uniqueness criteria between any two ears. It has suggested that ears may be distinguishable based on limited number of characteristics and features. Analysis of the decidability index also indicates the uniqueness of an individual ear where the decidability index of ear is found to be an order of magnitude greater than that of face, but not as large as that of iris. The characteristics making ear biometrics much popular are given below.

1. Ear is remarkably consistent and does not change its shape under different expressions like face. Moreover, ear has uniform color distribution.
2. Changes in the ear shape happen only before the age of 8 years and after that of 70 years [17]. Shape of the ear is very much stable for the rest of the life.
3. Like face, handling background is a challenging issue and often it requires data to be captured under controlled environment. However, in case of ear, background is predictable as an ear always remains fixed at the middle of the side face.
4. Size of the ear is larger than fingerprint, iris, retina etc. and smaller than face, and hence ear can be acquired easily.
5. Ear is a good example of passive biometrics and does not need much cooperation from user. Ear data can be captured even without the knowledge of the users from the far distance.

A biometric based security system is expected to fulfill user's demand such as low error rates, high security levels, testing for liveliness of the subject, possibility of fake detection etc. Even though the recognition performance of biometric systems has been significantly improved in recent past, there is a need of further improvement of existing techniques. Most of the existing ear recognition techniques have failed to perform satisfactorily in presence of varying illumination, occlusion and poor image registration. This paper proposes an efficient ear based recognition technique which can handle some of these factors. In this proposed technique, an ear image is enhanced using three image enhancement techniques applied in parallel. SURF feature extractor is used on each enhanced image to extract local features. A multi-matcher system is trained to combine the information extracted from each enhanced image. The technique is found to be robust to illumination changes and works well even when ear images are not properly registered.

The use of multiple image enhancement techniques has made it possible to counteract the effect of illumination and poor contrast while SURF based local feature helps in matching the images which are not properly registered and suffer from pose variations. For a given ear image, three enhanced images are obtained which are used by SURF feature extractor to generate three sets of SURF features for an ear image. Three nearest neighbor classifiers are respectively trained on these three sets of features and finally the output of all the classifiers are fused to get the final result. Experimental results show an improvement in performance compared to existing techniques. This paper also observes the advantage of use of multiple enhancement algorithms.

The rest of the paper is organized in the following way. Section 2 reviews some of the well known techniques available for ear recognition. Section 3 discusses SURF feature extractor and various enhancement techniques used in the proposed technique. Next section presents the technique for ear recognition. Experimental results are analyzed in Section 5. The paper is concluded in the last section.

2 Literature Review

Most of the well known techniques for 2D ear recognition can be divided into three types: Appearance Based techniques, Force Field Transformation based techniques and Geometric techniques. Appearance based techniques use either global or local appearance of the ear image for recognition. Techniques based on Principal Component Analysis (PCA) [8], Independent Component Analysis (ICA [26]), intensity and color space [20,21] etc. fall under this category. PCA based technique is the extension of the use of PCA in face recognition. It exploits the training data to find out a set of orthogonal basis vectors representing the directions of maximum variance in the data with minimum reconstruction mean square error. Usually, it drops the first eigenvector assuming that it represents the illumination changes in the image. Zhang et. al. [26] have used ICA for ear recognition which performs better than PCA. However, authors have not dropped the first eigenvector while comparing the results. Major drawback of this type of techniques is that they are only usable when images are captured in control environment and properly registered. Nanni and Lumini [20] have proposed a multi-matcher based technique for ear recognition which exploits appearance based local properties of an ear. It considers overlapping sub-windows to extract local features using bank of Gabor filters. Further Laplacian Eigen Maps are used to reduce the dimensionality of the feature vectors. Ear is represented using the features obtained from a set of most discriminative sub-windows selected using Sequential Forward Floating Selection (SFFS) algorithm. Matching in this technique is performed by combining the outputs of several 1-nearest neighbor classifiers constructed on different sub-windows. Another technique based on fusion of color spaces is proposed by Nanni and Lumini [21] where few color spaces are selected using SFFS algorithm and Gabor features are extracted from them. Matching is carried out by combining the output of several nearest neighbor classifiers constructed on different color components.

Force field based techniques [15,13,16] transform an ear image into a force field and extract features using force field energy functionals discussed in [14]. To transform an image into force field, an image is considered as an array of mutually attracting particles that act as a source of Gaussian force field. Underlying the force field, there exists a scalar potential energy field which, in case of an ear, appears as a smooth surface that looks like a small mountain with a number of peaks joined by ridges. Force field based techniques consider these peaks and ridges as features for ear representation. The directional properties of the force field are utilized to identify the extrema of a small number of potential energy wells and associated potential channels.

Burge and Burger [5,6] have proposed a technique for ear recognition using geometric information of the ear. The ear has been represented using a neighborhood graph obtained from a Voronoi diagram of the ear edge segments whereas template comparison has been performed using sub-graph matching. Choras [9,10] has used geometric properties of the ear to propose an ear recognition technique in which feature extraction is carried out in two steps. In the first step, global features are extracted whereas second step extracts local features. While matching, local features are only used when global features are found to be matching. In another geometry based technique proposed by Shailaja and Gupta [23], an ear is represented by two sets of features, global and local, obtained using outer and inner ear edges respectively. Two ears in this technique are declared similar if they are matched with respect to both the feature sets. The technique proposed in [7] has treated ear as a planar surface and has created a homography transform using SIFT [19] feature points to register ears accurately. It has

achieved robust results in presence of background clutter, viewing angle and occlusion. In [25], Yuan et al. have proposed a technique for human recognition with partially occluded ear images using neighborhood preserving embedding. Marsico et al. in [11] have proposed a fractal based technique to classify human ears. The technique adopts feature extraction locally, so that the system gets robust with respect to small changes in pose/illumination and partial occlusions.

3 Preliminaries

3.1 Speeded Up Robust Feature Transform

Speeded-Up Robust Features (SURF)[3, 2] is a rotation-invariant interest point detector and descriptor. It has been designed for extracting highly distinctive and invariant feature points from images. It is found to be more robust with view point, scale and illumination changes and occlusion. It identifies salient feature points in the image called key-points. It makes use of hessian matrix for key-point detection. For a given point $P(x, y)$ in an image I, the hessian matrix $H(P, \sigma)$ at scale σ is defined as:

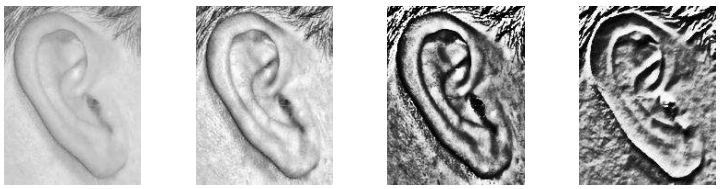
$$H(P,\sigma) = \begin{bmatrix} L_{xx}(P,\sigma)\ L_{xy}(P,\sigma) \\ L_{yx}(P,\sigma)\ L_{yy}(P,\sigma) \end{bmatrix}$$

where $L_{xx}(P, \sigma)$, $L_{xy}(P, \sigma)$, $L_{yx}(P, \sigma)$ and $L_{yy}(P, \sigma)$ are the convolution of the Gaussian second order derivatives $\frac{\partial^2}{\partial x^2}g(\sigma)$, $\frac{\partial^2}{\partial x \partial y}g(\sigma)$, $\frac{\partial^2}{\partial y \partial x}g(\sigma)$ and $\frac{\partial^2}{\partial y^2}g(\sigma)$ with the image I at point P respectively. To speed up the computation, second order Gaussian derivatives in Hessian matrix are approximated using box filters. To detect key-points at different scales, scale space representation of the image is obtained by convolving it with the box filters. The scale space is analysed by up-scaling the filter size rather than iteratively reducing the image size. In order to localize interest points in the image and over scales, non-maximum suppression in a $3 \times 3 \times 3$ neighborhood is implemented.

In order to generate key point descriptor vector, a circular region is considered around the detected key-points and Haar wavelet responses dx and dy in horizontal and vertical directions are computed. These responses are used to obtain the dominant orientation in the circular region. Feature vectors are measured relative to the dominant orientation resulting the generated vectors invariant to image rotation. Also a square region around each key-point is considered and it is aligned along the dominant orientation. The square region is divided into 4×4 sub-regions and Haar wavelet responses are computed for each sub-region. The sum of the wavelet responses in horizontal and vertical directions for each sub-region are used as feature values. In addition, the absolute values of responses are summed to obtain the information about the polarity of the image intensity changes. Thus, the feature vector V_i for i^{th} sub-region is given by

$$V_i = \{\Sigma dx, \Sigma dy, \Sigma|dx|, \Sigma|dy|\}$$

SURF feature vector of a key-point is obtained by concatenating feature vectors (V_i) from all sixteen sub-regions around the key-point resulting a vector of 64 elements. Extended version of SURF (known as SURF-128), which is more distinctive, adds couple of similar features. It uses the sums same as described above, but splits these values up further. The sum of d_x and of $|d_x|$ are computed separately for $d_y < 0$ and



(a) (b) (c) (d)

Fig. 1 Image enhancement examples: (a) Original image from UND-E dataset, output after applying (b) ADHist, (c) NLM and (d) SF enhancement techniques

$d_y \geq 0$. Similarly, the sum of d_y and of $|d_y|$ are found according to the sign of d_x, hence doubling the number of features. The proposed ear recognition technique uses SURF-128 (referred as only SURF in further discussion) for feature representation.

Matching in SURF is performed using nearest neighbor ratio matching. The best candidate match for a keypoint in an image is found by identifying its nearest neighbor in the keypoints from the test image where nearest neighbors are defined as the keypoints with minimum Euclidean distance from the given descriptor vector. The probability that a match is correct is determined by computing the ratio of distance from the closest neighbor to the distance of the second closest. All matches in which the distance ratio is greater than a threshold (τ) are rejected.

3.2 Image Enhancement

This subsection describes three image enhancement techniques, namely Adaptive Histogram Equalization, Steerable Gaussian Filter and Non-Local Means Filter that are used in the proposed technique for enhancing the ear images.

3.2.1 Adaptive Histogram Equalization

Adaptive histogram equalization (ADHist) [27] can be used to improve the contrast of an image. It divides an image into multiple non-overlapping tiles (regions) and performs histogram equalization for each one individually. This enhances the contrast of each tile. The neighboring tiles are combined together to get the entire enhanced image. ADHist uses bilinear interpolation to remove artificially induced boundaries while combining the tiles. Adaptive histogram equalization is capable of improving the local contrast of the image and bringing out more details in the image.

Let $I \in R^{a \times b}$ be the image of size $a \times b$ to be enhanced. It is divided into the tiles of size $\alpha \times \beta$, when $\alpha < a$ and $\beta < b$. These tiles are enhanced individually and stitched together to get the overall enhanced image. Selection of appropriate values for α and β greatly affects the enhancement performance. In this paper, these values are chosen empirically. Figure 1(b) shows the enhanced image obtained using ADHist.

3.2.2 Non-Local Means Filter

The non-local means (NLM) algorithm [24] is proposed for image enhancement by using image denoising. It considers pixel values from the entire image for the task of

noise reduction. The algorithm is based on the fact that for every small window of the image, several similar windows can be found in the image and all of these windows can be exploited to denoise the image. Let the noisy image be denoted by $I_n(p) \in R^{a \times b}$, where a and b are image dimensions and let $p = (x, y)$ stand for an arbitrary pixel location within the noisy image. The NLM algorithm constructs the denoised image $I_d(p)$ by computing each pixel value of $I_d(p)$ as a weighted average of pixels comprising $I_n(p)$, *i.e.*:

$$I_d(p) = \sum_{p \in I_n(p)} w(z,p) I_n(p)$$

where $w(z,p)$ represents the weighting function which measures the similarity between the local neighborhoods of the pixel at the spatial locations z and p. The weighting function used in this equation is defined as follows:

$$w(z,p) = \frac{1}{Z(z)} e^{-\frac{G_\sigma \| I_n(\Omega_p) - I_n(\Omega_z)) \|_2^2}{h^2}}$$

$$\text{where, } Z(z) = \sum_{p \in I_n(p)} e^{-\frac{G_\sigma \| I_n(\Omega_p) - I_n(\Omega_z)) \|_2^2}{h^2}}$$

where G_σ denotes a Gaussian kernel with standard deviation σ, Ω_p and Ω_z are the local neighborhoods of the pixels at the locations p and z respectively, h stands for the parameter that controls the decay of the exponential function and $Z(z)$ represents a normalizing factor. It can be observed that if the local neighborhoods of a given pair of pixel locations z and p display a high degree of similarity, the pixels at z and p can be assigned relatively large weights at the time of computing their denoised estimates.

A proper selection of the neighborhood size N and decay parameter h results in a smoothed image with preserved edges. Hence, it can be used to estimate the luminance of an input image and consequently, to compute the (logarithmic) reflectance. An example of the deployment of the NLM algorithm (for a 3×3 local neighborhood and $h = 50$) for estimation of the logarithmic reflectance is shown in Figure 1(c).

3.2.3 Steerable Filter

Steerable Filter (SF) [12] provides an efficient architecture to synthesize filters of arbitrary orientations from linear combinations of basis filters. This allows one to adaptively "steer" a filter to any orientation and to determine analytically the filter output as a function of orientation. These filters are normally used for early vision and image processing tasks such as angularly adaptive filtering, shape-from-shading, edge detection etc. However, they can also be used to produce illumination invariant representation of an image, such as the gradient image. For example, Gaussian function can be used as the basis filters to obtain steerable filters. To get illumination invariant representation of an image, steerable Gaussian derivatives can be applied at multiple scales and orientations to an image. Output image is computed by taking the weighted linear combination of the filtered images obtained after applying the Gaussian derivatives of various scales and orientation to the input image.

There are two critical parameters used to define steerable filters based on Gaussian function: one is σ to define the scale of the filter and another is θ to define the orientation of the filter. To define the Gaussian functions at multiple scales, a set of σ and θ values are required. Let these values be $\{\sigma_1, \sigma_2, .., \sigma_l\}$ and $\{\theta_1, \theta_2, .., \theta_n\}$. Each pair of (σ_i, θ_j)

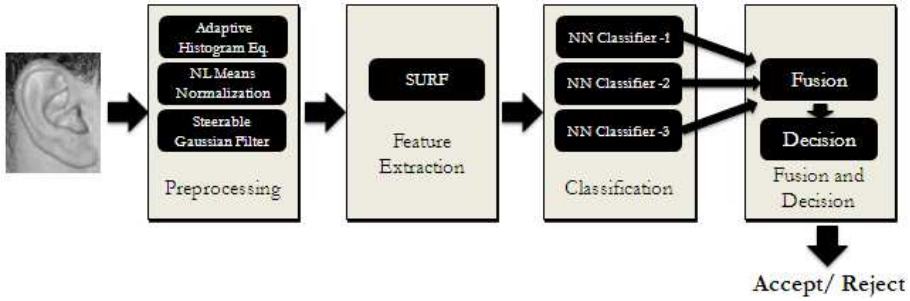


Fig. 2 Block diagram of the proposed ear recognition technique

defines one basis filter. Angular spacing of the filters is usually taken equal, hence n values of θ define n angles equally drawn from 0 to 180^0. The choice of values of σ and n depends on the size and content of the image respectively. An example of image enhancement of the image shown in Figure 1(a) using SF technique is given in Figure 1(d). It gives illumination normalized image.

4 Proposed Technique

The proposed ear recognition system follows three major steps: Image Enhancement, Feature Extraction and Classification and Fusion. Overview of the proposed system is shown in Figure 2. Various steps of the proposed system have been discussed in the following subsections.

4.1 Image Enhancement

This step involves three image enhancement techniques and is intended to enhance the contrast of the ear image and to normalize the effect of illumination and shadow. The purpose of enhancement is to get the correct SURF feature descriptor vectors for a feature points and to help in establishing the correct point correspondence between the feature points in two images. For example, a particular feature point in two different images of the same subject (which are differently illuminated) may get two different SURF descriptor vectors in the absence of enhancement. But when enhancement is applied, descriptor vectors for corresponding points in two images are found to be very similar.

The enhancement algorithms presented in Section 3.2 have been used in parallel on each input ear image to get enhanced image which are later used for feature extraction.

4.2 Feature Extraction

This step uses SURF technique for feature extraction which provides representation of an image in terms of a set of salient feature points, each point associated with a descriptor vector of 128 feature elements. SURF features are efficiently able to capture the properties of spatial localization, change in 3D viewpoint, orientation and scale

sensitivity. It provides highly distinctive features, in the sense that a single feature can be correctly matched with high probability against a large database of features from many images.

A technique for feature level fusion is proposed to obtain a fused representative template for a subject by combining the features of multiple training samples of the subject. If n samples of a subject are provided for training, a representative feature template for the subject is obtained by fusing the feature points together and by considering the redundant feature points only once. Let n biometric feature templates of a subject used for training be represented by $F_1, F_2, .., F_n$. A fused biometric feature template F_{fused} is obtained as follows:

$$F_{fused} = F_1 \cup F_2 \cup ... \cup F_n$$

where the cardinality of set F_{fused} provides the number of feature points present in the fused template. Fusion of the templates is done incrementally where first two feature templates F_1 and F_2 are fused to generate a new template T which is fused with feature template F_3. This procedure is continued until all the feature templates are fused together. While fusing two biometric templates F_i and F_{i+1}, SURF matching is applied between the templates to find out the redundant points. If a feature point in a template matches to a feature point in the another template, it is considered as common to both and is used only once in fusion.

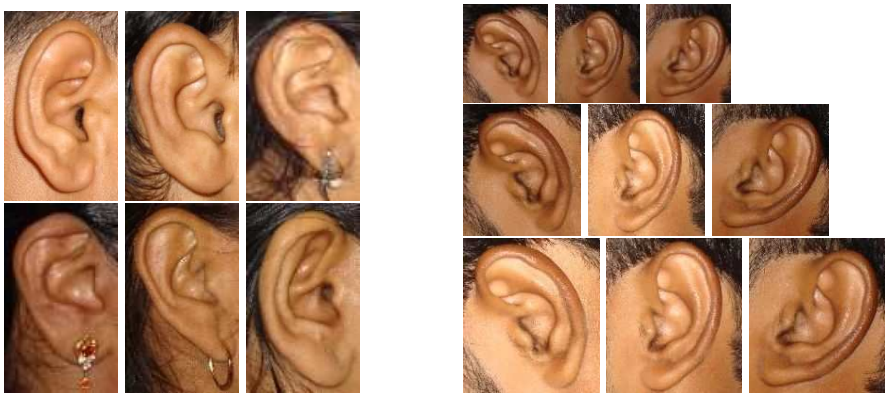
4.3 Classification and Fusion

Extracted features obtained from each enhanced image are used for classification to train a nearest neighbor classifier. The matching strategy in the nearest neighbor classifier is as follows. An interest point in the test image is compared to an interest point in the reference template by calculating the Euclidean distance between their descriptor vectors. Matching of two feature points is carried out using the nearest neighbor ratio matching strategy where a matching pair is detected if its distance is closer than τ times of the distance of the second nearest neighbor where τ is the matching threshold.

Matching score between two ear images is obtained based on the number of matched feature points between two images. These matching scores are normalized using min-max normalization technique and are then fused using weighted sum rule. Final classification decision is taken by using the fused scores.

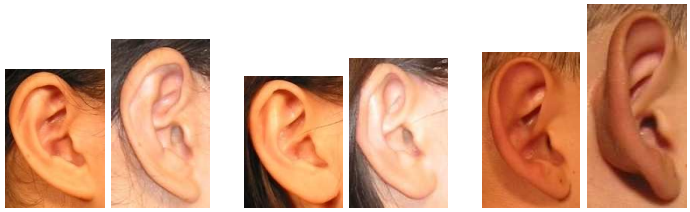
5 Experimental Results

The performance of a biometric system can be measured in terms of recognition accuracy, equal error rate (EER) and error under ROC curves (EUC). Recognition accuracy is used in verification system and is defined as $100 - \frac{(FAR+FRR)}{2}$ where FAR (False Acceptance Rate) indicates the rate at which an imposter is incorrectly accepted as genuine person and FRR (False Rejection Rate) is the rate at which a genuine person is incorrectly rejected as an imposter. EER is defined as the rate at which both FAR and FRR errors are equal. The performance of a verification system can also be evaluated using a receiver operator characteristic (ROC) curve, which graphically demonstrates the changes of GAR (Genuine Acceptance Rate, defined as 100 - FRR) with changes in FAR.



(a) Data Set 1 (a) Data Set 2

Fig. 3 Few sample images from IITK data sets



(a) Subject ID: 04202 (b) Subject ID: 04217 (c) Subject ID: 04295

Fig. 4 Few sample images from UND-E database

5.1 Databases

Experiments are conducted on two databases, namely IIT Kanpur database and University of Notre Dame database (Collections E) [1]. Table 1 provides the summary of these databases.

5.1.1 IIT Kanpur Database

IIT Kanpur (IITK) database is composed of two data sets. Data Set 1 contains 801 side face images collected from 190 subjects. Number of images acquired from an individual varies from 2 to 10. Figure 3(a) shows few sample images from Data Set 1. Data Set 2 consists of 801 side face images collected from 89 individuals. For each subject, 9 images are captured by considering three rotations and three scales for each rotation. Images of Data Set 2 consist of frontal view of the ears captured at three positions, first when person is looking straight, second when he/she is looking approximately 20^0 down and third when he/she is looking approximately 20^0 up. At all these positions, images are captured at 3 different scales by positioning the camera at a distance of approximately 1 meter and setting up the digital zoom of the camera at 1.7x, 2.6x and 3.3x. Figure 3(b) shows 9 images from Data Set 2 for an individual. The purpose of the use of multiple data sets is to show the robustness of the proposed approach. IITK

Table 1 Summary of the databases used in experimentation

Database	Number of Subjects	Total Samples	Description
IITK Data Set 1	190	801	2-10 samples per subject, frontal ear images
IITK Data Set 2	89	801	9 samples per subject, frontal ear images affected by scaling, rotation and poor registration
UND Dataset (Collection E)	114	464	3-9 samples per subject, images affected by illumination and pose variations, poor contrast and registration

Data Set 1 provides frontal ear images while IITK Data Set 2 provides challenging images which are affected by scaling and rotation.

5.1.2 University of Notre Dame Database

University of Notre Dame database, Collection E (UND-E) consists of 464 side face images collected from 114 subjects, 3 to 9 samples per subject. The images are collected on different days with different conditions of pose and illumination. Some of the sample ear images from UND-E database are shown in Figure 4. It can be noted that there is a huge intra-class variation present in these images due to pose variation and different imaging conditions.

5.1.3 Ear Extraction from the Background

IITK and UND-E databases contain side face images of human subjects. These ears are segmented from the side face images using ear segmentation technique discussed in [22]. Manual segmentation is performed for the images ($\approx 4\%$) where [22] is found to be deficient to segment the ears.

5.2 Parameters Tuning

Selection of appropriate values of parameters is critical for achieving the best performance of the proposed technique. Main parameters which have great impact on the performance are dimensions of the tiles in ADHist, values of σ and n in SF, values of h and N in NLM and value of τ in SURF Matching.

Since it is difficult to test the proposed technique for all possible values of these 6 parameters, the parameters are tuned for optimal values heuristically and the best performance is obtained. To achieve optimal values of parameters, a set of 25 subjects is randomly selected from each database and parameter tuning is performed only on this set. These parameters are used for testing the full database.

5.2.1 Dimensions of the Tile for ADHist

The proposed technique considers the tiles of square size, *i.e.* $\alpha = \beta$ in ADHist technique. Dimensions of the tiles are varied from 2×2 to 20×20 and for each value,

Table 2 Computation of optimal dimensions of the tile in ADHist for IITK database

(a) IITK Data Set 1

Tile Size	SURF Matching Threshold (τ)									
	0.3		0.4		0.5		0.6		0.7	
	EER (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)
2×2	3.49	2.58	3.60	1.90	5.47	1.99	8.37	2.85	12.40	5.78
4×4	3.51	2.01	3.50	1.25	5.07	1.33	7.89	2.31	12.21	5.41
6×6	3.48	1.95	3.50	1.17	4.19	1.10	7.25	2.01	12.23	4.84
8×8	3.54	3.03	**3.46**	**1.42**	5.23	1.58	8.44	2.52	12.54	5.36
10×10	4.56	4.10	4.15	2.61	5.73	2.23	8.53	3.01	12.25	5.70
12×12	8.49	8.62	6.11	5.55	6.64	3.84	9.87	4.55	14.05	7.11
14×14	9.91	9.98	5.93	5.23	7.05	4.31	10.31	4.87	14.39	7.46
16×16	9.64	9.70	6.40	5.51	7.52	4.48	10.61	4.80	15.11	8.09
18×18	8.82	8.77	6.68	5.36	7.02	3.99	10.46	4.84	15.12	8.17
20×20	10.63	10.70	7.34	6.27	7.37	4.50	10.56	5.01	15.52	8.23

(b) IITK Data Set 2

Tile Size	SURF Matching Threshold (τ)									
	0.3		0.4		0.5		0.6		0.7	
	EER (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)
2×2	2.68	1.76	2.43	1.04	3.87	0.89	7.61	2.12	14.89	6.7
4×4	2.66	2.06	**2.25**	**1.03**	4.03	1.07	7.23	2.14	13.89	6.1
6×6	3.06	2.64	3.32	1.12	3.84	1.02	6.43	2.01	12.88	5.75
8×8	3.42	3.1	4.38	1.93	5.21	1.51	7.58	2.42	14.32	6.39
10×10	4.92	4.55	4.27	1.96	5.52	2.01	8.84	2.83	14.97	6.71
12×12	7.08	6.73	6.01	3.76	6.51	2.72	9.6	3.8	16.15	7.84
14×14	9.64	9.34	6.73	4.63	7.25	2.98	11.46	4.27	17.61	8.84
16×16	11.43	11.15	7.47	5.56	8.64	3.71	11.86	4.85	18.87	9.7
18×18	12.04	11.83	8.11	6.50	9.77	4.94	12.72	5.82	20.47	10.82
20×20	13.33	13.15	9.34	7.99	11.12	6.10	13.54	6.70	21.09	11.12

Table 3 Computation of optimal dimensions of the tile in ADHist for UND-E database

Tile Size	SURF Matching Threshold (τ)									
	0.3		0.4		0.5		0.6		0.7	
	EER (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)
2×2	11.75	6.73	9.97	3.93	8.83	3.53	9.47	3.63	11.44	4.4
4×4	15.14	8.37	10.42	4.71	10.11	3.12	9.57	3.22	11.5	4.08
6×6	13.27	7.05	9.62	3.9	9.47	3.71	9.78	3.53	10.73	3.94
8×8	13.2	7.93	9.34	4.21	8.68	3.04	8.14	2.48	9.63	3.26
10×10	13.88	9.33	9.66	4.19	8.1	2.71	8.46	2.60	10.00	3.38
12×12	12.81	8.48	10.48	4.82	8.20	2.53	7.24	2.18	8.03	2.89
14×14	13.00	9.24	10.78	4.87	8.31	3.16	8.06	2.74	8.15	2.36
16×16	12.98	9.55	10.51	3.99	8.07	2.88	**6.72**	**2.40**	8.39	2.41
18×18	17.91	10.69	13.05	5.16	10.14	3.26	8.2	2.67	8.67	2.65
20×20	13.2	10.01	10.32	4.02	7.96	2.57	7.33	2.12	7.26	2.22

Table 4 Computation of optimal values of h and N in NLM filters for IITK database

(a) IITK Data Set 1

h	N	SURF Matching Threshold (τ)									
		0.3		0.4		0.5		0.6		0.7	
		EER (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)
20	4	11.31	11.42	5.63	5.63	4.52	3.85	5.64	2.95	8.99	3.86
	5	16.60	16.79	9.78	9.90	6.31	6.20	6.54	4.90	9.34	4.53
	6	29.21	29.48	20.17	20.50	14.34	14.67	11.92	12.13	11.16	9.94
50	4	5.16	5.01	3.97	2.86	5.21	1.97	7.61	2.34	11.47	4.81
	5	5.33	5.28	4.31	3.32	4.30	1.92	6.51	2.18	10.98	4.51
	6	5.33	5.33	3.80	3.08	4.36	2.37	6.27	2.21	10.77	4.21
80	4	3.36	2.77	2.99	1.32	4.52	1.07	7.79	2.06	12.64	4.98
	5	3.46	2.93	3.25	1.51	4.80	1.07	7.69	2.05	12.32	4.99
	6	4.95	4.75	3.30	2.21	4.36	1.54	7.34	2.26	12.11	4.93
100	4	3.48	2.82	3.12	1.43	4.94	1.30	7.54	2.18	12.83	5.26
	5	3.40	2.80	3.03	1.44	4.77	1.25	7.93	2.10	12.76	5.15
	6	3.54	2.99	**2.90**	**1.17**	4.78	1.14	7.57	2.01	12.24	4.93

(b) IITK Data Set 2

h	N	SURF Matching Threshold (τ)									
		0.3		0.4		0.5		0.6		0.7	
		EER (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)
20	4	22.01	22.01	11.41	11.37	6.38	5.92	6.87	4.01	10.52	4.39
	5	28.05	28.05	17.41	17.40	10.42	10.35	7.64	6.75	11.40	5.79
	6	34.10	34.10	21.88	21.88	13.46	13.43	8.64	8.20	11.13	7.03
50	4	4.35	4.04	4.41	2.37	5.24	1.26	8.00	2.38	13.99	5.91
	5	5.02	4.84	3.90	2.24	4.76	1.41	7.29	1.98	12.74	5.31
	6	5.76	5.67	**3.48**	**2.32**	4.49	1.12	6.85	1.86	12.29	4.83
80	4	3.72	3.27	3.98	1.60	4.40	1.34	7.31	2.30	14.45	6.11
	5	3.88	3.46	4.08	1.67	4.42	0.94	7.37	2.12	14.23	6.01
	6	3.95	3.59	3.95	1.80	4.35	1.19	7.44	2.08	13.71	5.78
100	4	3.86	3.36	4.11	1.62	4.26	1.33	7.63	2.18	14.07	6.05
	5	3.89	3.44	3.71	1.54	4.64	1.29	7.35	2.14	14.01	6.07
	6	3.74	3.30	4.08	1.73	4.49	1.26	7.32	2.25	13.60	5.89

Table 5 Computation of optimal values of h and N in NLM filters for UND-E database

h	N	SURF Matching Threshold (τ)									
		0.3		0.4		0.5		0.6		0.7	
		EER (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)
20	4	20.85	20.84	11.07	10.82	9.11	6.65	8.02	3.1	7.12	2.27
	5	21.75	21.74	11.68	11.61	8.54	7.51	7.63	4.29	7.77	2.84
	6	27.05	27.05	13.78	13.75	9.00	8.58	8.67	4.40	7.77	2.72
50	4	10.55	10.07	9.85	5.14	7.31	2.63	5.80	1.83	5.97	1.76
	5	11.77	11.58	8.96	6.29	7.52	2.92	6.22	1.90	5.79	1.55
	6	13.42	13.34	8.62	7.21	6.59	3.7	5.79	1.81	**5.75**	**1.40**
80	4	9.47	7.44	8.22	3.30	7.07	2.54	5.87	1.62	5.89	1.64
	5	12.46	9.4	10.09	4.62	7.85	2.5	5.84	1.86	5.89	1.58
	6	12.81	10.13	9.63	4.75	7.27	2.53	5.85	1.48	5.83	1.33
100	4	12.71	7.86	10.87	4.46	8.49	2.88	5.89	1.78	6.21	1.68
	5	10.31	8.03	8.09	3.26	6.83	2.30	5.99	1.94	5.90	1.56
	6	10.91	9.25	7.31	3.87	6.13	2.35	5.85	1.47	5.85	1.50

Table 6 Computation of optimal values of σ and n in SF for IITK and UND-E databases

(a) IITK Data Set 1

σ	n	SURF Matching Threshold (τ)									
		0.3		0.4		0.5		0.6		0.7	
		EER (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)
{0.5, 1,	4	3.50	1.76	3.52	1.27	4.04	1.34	6.67	1.99	11.45	4.46
1.5,2,	6	3.48	2.95	3.60	2.07	5.23	2.02	7.17	2.59	12.44	5.16
2.5}	8	3.50	1.63	**3.46**	**0.98**	4.85	1.32	7.36	2.25	12.47	5.19
{0.1,1,	4	3.51	2.07	3.52	1.52	4.28	1.51	6.83	2.10	12.11	5.02
2,3	6	3.53	2.02	3.60	1.42	4.94	1.64	7.76	2.51	12.61	5.45
4}	8	3.56	1.84	3.58	1.25	5.02	1.46	7.31	2.46	12.97	5.75

(b) IITK Data Set 2

σ	n	SURF Matching Threshold (τ)									
		0.3		0.4		0.5		0.6		0.7	
		EER (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)
{0.5, 1	4	3.83	3.02	4.25	1.63	4.77	1.54	7.43	2.68	13.57	6.07
1.5, 2	6	4.11	3.23	3.86	1.54	5.01	1.49	7.34	2.68	13.77	6.17
2.5}	8	3.30	2.16	4.00	1.19	5.02	1.34	8.02	2.61	14.33	6.47
{0.1,1	4	4.26	3.42	3.74	1.35	4.51	1.48	7.96	2.86	14.60	6.68
2,3	6	3.60	2.76	3.28	1.15	4.77	1.34	7.99	2.64	15.06	6.83
4}	8	3.45	2.61	**3.28**	**1.11**	4.58	1.29	8.00	2.75	14.21	6.59

(c) UND-E Database

σ	n	SURF Matching Threshold (τ)									
		0.3		0.4		0.5		0.6		0.7	
		EER (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)	*EER* (%)	*EUC* (%)
{0.5, 1	4	12.92	9.69	10.64	5.05	9.28	3.62	8.92	3.17	8.18	2.72
1.5, 2	6	12.27	6.96	10.04	3.53	7.41	2.65	6.83	2.25	6.53	2.13
2.5}	8	12.41	7.37	9.15	3.5	7.1	1.79	**6.51**	**1.67**	6.61	1.72
{0.1,1	4	13.71	7.47	10.68	4.66	9.83	3.9	9.24	3.41	9.66	4.07
2,3	6	13.39	7.55	10.37	4.29	8.75	3.06	8.02	2.71	8.23	2.8
4}	8	14.45	7.62	10.45	4.6	8.01	2.73	8.02	2.48	7.62	2.63

EER and *EUC* of the system are computed when only ADHist is used for image enhancement. The tile size which corresponds to minimum *EER* is chosen as the optimal size. Also if two tile sizes give same *EER*, their corresponding *EUC* values are used to break the tie and the tile size for which less *EUC* is obtained, is considered as the optimum tile size. Experiments are conducted to find *EER* and *EUC* for IITK and UND-E databases which are shown in Table 2 and Table 3 respectively. It can be observed from the tables that the optimal values of tile size for IITK database Set 1 and Set 2 are 8×8 and 4×4 respectively while that for UND-E database is 16×16.

We have noticed that the changes in *EER* and *EUC* are gradual in Table 2 and Table 3 except a few exceptions. In Table 2, values of *EER* and *EUC* are gradually increased. But if one observes Table 3, one finds that for $\tau = 0.6$ and $\tau = 0.7$, *EER* and *EUC* are gradually decreased while for $\tau = 0.4$ and $\tau = 0.5$, *EER* and *EUC* are almost consistent. Also for $\tau = 0.3$, its behaviour is little abrupt because SURF matching at low threshold is not very stable. It can be noted that UND-E data set is having

illumination and contrast variations which cause the observations to be little abrupt. But it is not the case with IITK data sets.

From Table 2(a), it can be seen that error values are almost same for $\tau = 0.7$ and for different tile sizes lying between 2×2 and 10×10 or that between 12×12 and 14×14 or that between 16×16 and 20×20. Thus, little change in the tile size does not significantly change the error values.

5.2.2 Values of h and N in NLM Filters

In NLM filters, h is a scalar which controls the decay of the exponential function and N is a scalar defining the neighborhood size (*i.e.*, the size of the patches to be used in the NLM algorithm). To search the optimal values, h and N are changed between 20 to 100 and 4 to 6 respectively. For each combination of (h, N), image enhancement is performed and the enhanced image is used for recognition. The values of *EER* and *EUC* of the system are computed and (h, N) values which correspond to minimum *EER* is considered as optimal. *EUC* is used to break the tie in case of two or more (h, N) pairs give same *EER*. Experiments are conducted to find *EER* and *EUC* for IITK and UND-E databases which are shown in Table 4 and Table 5 respectively. It is observed that the optimal values of (h, N) for IITK database Set 1 and Set 2 are $(100, 6)$ and $(50, 6)$ respectively while for UND-E database it is $(50, 6)$.

5.2.3 Values of σ and n in SF

In steerable filters, σ defines a vector of length l where l is the number of filter scales and n is the angular resolution of filters. In our experiments, we have considered two sets of σ: $\{0.5, 1, 1.5, 2, 2.5\}$ and $\{0.1, 1, 2, 3, 4\}$ while value of n is taken as 4 (*i.e.*, $\theta = 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$), 6 (*i.e.*,$0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{5\pi}{6}$) and 8 (*i.e.*, $0, \frac{\pi}{8}, \frac{\pi}{4}, \frac{3\pi}{8}, \frac{\pi}{2}, \frac{5\pi}{8}, \frac{3\pi}{4}, \frac{7\pi}{8}$). For each combination of σ and n, image enhancement is performed and enhanced image is used for recognition using SURF features and nearest neighbor classifier. The values of *EER* and *EUC* of the system are computed and (σ, n) value which corresponds to minimum *EER* is considered as optimal. Experiments are conducted to find *EER* and *EUC* for IITK and UND-E databases which are shown in Table 6. It is observed from the table that the optimal values of parameters (σ, n) for SF are $(\{0.5, 1, 1.5, 2, 2.5\}, 8)$ and $(\{0.1, 1, 2, 3, 4\}, 8)$ for IITK database Set 1 and Set 2 respectively while $(\{0.5, 1, 1.5, 2, 2.5\}, 8)$ for UND-E database. Further, there are two values of (σ, n) pair (*i.e.*, $(\{0.1, 1, 2, 3, 4\}, 6)$ and $(\{0.1, 1, 2, 3, 4\}, 8)$) in Table 6(b) for which *EER* attains the minimum value. So to break the tie, *EUC* is used and $(\{0.1, 1, 2, 3, 4\}, 8)$ is chosen as the optimal parameter set as it has the minimum *EUC* value among the two.

5.2.4 Value of τ for SURF Matching

Correctness of a match in SURF matching is determined by computing the ratio of distance from the closest neighbor to the distance of the second closest neighbor. All the matches in which the distance ratio is greater than τ are rejected. Experiments are performed on IITK and UND-E databases by changing the value of τ from 0.3 to 0.7 with an increment of 0.1. This range of values is used in each of the experiment conducted to determine the parameters of ADHist, NLM and SF filters. There

Table 7 Optimal parameters for the proposed technique

Enhancement Technique	Parameter	Databases		
		IITK Set 1	IITK Set 2	UND-E
ADHist	Tile Size	8×8	4×4	16×16
	τ	0.4	0.4	0.6
NLM	h	100	50	50
	N	6	6	6
	τ	0.4	0.4	0.7
SF	σ	{0.5, 1, 1.5, 2, 2.5}	{0.1, 1, 2, 3, 4}	{0.5, 1, 1.5, 2, 2.5}
	n	8	8	8
	τ	0.4	0.4	0.6

are 3 values of τ which are determined for each database, one for each enhancement technique. Values of all the parameters discussed above are summarized in Table 7.

5.3 Results

Table 8 and Table 9 give the values of recognition accuracy (with corresponding *FAR* and *FRR*), *EER*, *EUC* for IITK ear database for various combinations of enhancement techniques. It can be observed that the best results are obtained when all three image enhancement techniques are employed in the recognition process. *ROC* curves for Data Set 1 and Data Set 2 are shown in Figure 5(a) and Figure 5(b) respectively. The *ROC* curves obtained for the technique employing all three image enhancement techniques is found to be superior to others.

Accuracy obtained in Table 9 is always greater than that shown in Table 8 except for NLM. Greater accuracy in Table 9 is achieved due to the fact that in Data Set 2, all the subjects are having 9 samples while in Data Set 1, number of samples varies from 2 to 10 (almost 50% subjects have number of samples less than 4). This provides better training in Data Set 2 compared to Data Set 1 which leads to better accuracy.

Table 10 gives the values of various performance measures for UND-E database for various combinations of enhancement techniques. For this database also, it is noticed that the best results are obtained when all three enhancement techniques are employed in recognition process. From the table, it is observed that the best *EER* and *EUC* are much less than those reported in two well known ear recognition techniques [20] and [21]. Comparative performance of the proposed technique with the best known results for UND-E database is summarized in Table 11. Results obtained by the proposed technique are averaged over 30 experiments; hence it shows more stable performance compared to the results reported in [20] and [21] where they are averaged only for 10 and 20 experiments respectively. *ROC* curves for UND-E database are shown in Figure 6 where the *ROC* curve employing all three image enhancement techniques is found to be superior to others.

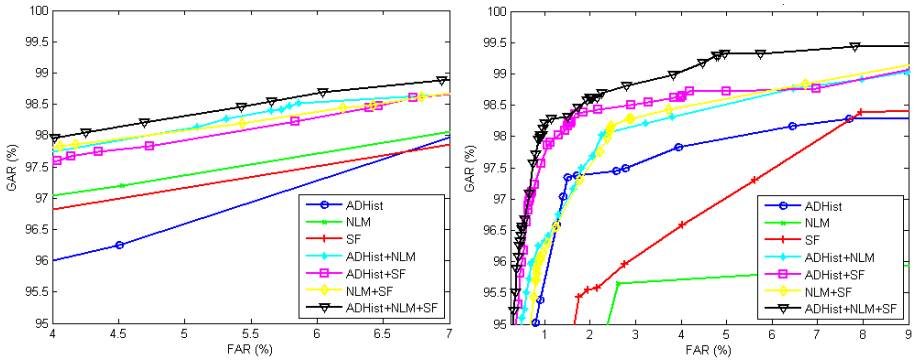
Score level fusion is performed by using weighted sum rule [18]. It is assumed that independent use of classifiers C_1, C_2 and C_3 for classification produces classification accuracies A_1, A_2 and A_3 respectively. In the proposed technique, these accuracies are used to weight the scores of individual classifiers for fusion. The modified fusion score

Table 8 Performance of the proposed technique on IITK Data Set 1 for various combinations of enhancement techniques

Fusion Scheme	Accuracy (*FAR*,*FRR*)	*EER*	*EUC*
ADHist	96.54(2.89,4.04)	3.46	1.42
NLM	97.10(3.07,2.72)	2.90	1.17
SF	96.68(2.92,3.72)	3.46	0.98
ADHist + NLM	97.25(2.83,2.67)	2.98	0.90
ADHist + SF	97.13(2.92,2.82)	3.09	0.80
NLM + SF	97.20(2.71,2.89)	2.94	0.83
ADHist + NLM + SF	**97.35(2.70,2.60)**	**2.88**	**0.75**

Table 9 Performance of the proposed technique on IITK Data Set 2 for various combinations of enhancement techniques

Fusion Scheme	Accuracy (*FAR*,*FRR*)	*EER*	*EUC*
ADHist	97.94(1.42,2.70)	2.25	1.03
NLM	96.55(2.10,4.79)	3.48	2.32
SF	96.85(1.70,4.61)	3.28	1.11
ADHist + NLM	98.17(1.49,2.17)	2.11	0.58
ADHist + SF	98.62(1.07,1.69)	1.68	0.40
NLM + SF	98.07(1.83,2.02)	2.26	0.48
ADHist + NLM + SF	**98.79(0.88,1.54)**	**1.59**	**0.36**



(a) Date Set 1 (b) Date Set 2

Fig. 5 *ROC* curves for IITK data sets showing the performance for various combinations of enhancement techniques

is given as follows:

$$S = \frac{A_1 \times S_1 + A_2 \times S_2 + A_3 \times S_3}{A_1 + A_2 + A_3}$$

where S_1, S_2, S_3 are the individual scores produced by classifiers C_1, C_2 and C_3 respectively. *ROC* curves, shown in Figure 5 and Figure 6, are drawn for the system which use the weighted sum rule for fusion of matching scores obtained through three classifiers.

Table 10 Performance of the proposed technique on UND-E database for various combinations of enhancement techniques

Fusion Scheme	Accuracy (*FAR*,*FRR*)	*EER*	*EUC*
ADHist	93.64 (5.18,7.54)	6.72	2.40
NLM	95.25 (2.31,7.19)	5.75	1.40
SF	94.17 (3.31,8.36)	6.51	1.67
ADHist + NLM	96.13 (2.97,4.77)	4.40	1.34
ADHist + SF	95.41 (4.01,5.18)	5.06	1.49
NLM + SF	96.31 (2.85,4.53)	4.22	**1.13**
ADHist + NLM + SF	**96.75 (2.58,3.92)**	**3.80**	1.16

Table 11 Comparison of performance of the proposed technique with the latest reported results for UND-E database

Technique	Accuracy (*FAR*,*FRR*)	*EER*	*EUC*
Proposed in [20]	-	4.20	3.00[a]
Proposed in [21]	-	-	1.50
Proposed Technique	**96.75 (2.58,3.92)**	**3.80**	**1.13**

[a] reported in [21] for the technique proposed in [20]

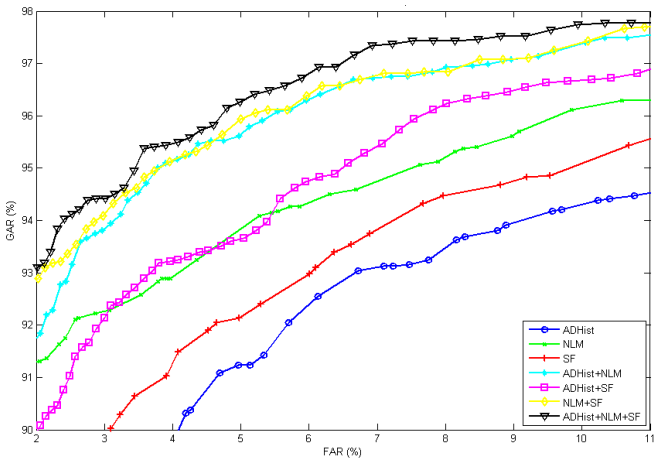


Fig. 6 *ROC* curves for UND-E database for combinations of various enhancement techniques

6 Conclusions

The available ear recognition techniques perform poor in presence of varying illumination, poor contrast, view point changes and non-registered images. This paper has attempted to overcome these challenges and has presented a novel technique for ear based human recognition. This technique uses three different image enhancement techniques in parallel to overcome the effect of illumination and contrast and extracts local features from the enhanced images using SURF. Use of SURF based local features

helps in dealing with the problem of pose variation and poor image registration. Three nearest neighbor classifiers are employed which are trained on the features obtained from three different enhanced images respectively. Fusion at score level is carried out to combine the scores generated from the three classifiers and decision is taken based on the fused score. The proposed technique has been evaluated on two ear databases, namely IIT Kanpur ear database and University of Notre Dame ear database (Collection E). IIT Kanpur ear database includes images of various rotations, sizes and shapes while University of Notre Dame database consists of ear images with variable illumination, pose changes and poor contrast. Experimental results show that the proposed technique provides a considerable improvement in terms of performance over existing techniques.

Acknowledgements Authors would like to thank the anonymous reviewers for their critical and constructive comments and suggestions which have helped very much in improving the readability of this paper.

References

1. University of Notre Dame Ear Database, Collection E. `http://www.nd.edu/~cvrl/CVRL/DataSets.html`
2. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features (SURF). Computer Vision and Image Understanding **110**(3), 346–359 (2008)
3. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: Speeded up robust features. In: Proc. of 9th European Conference on Computer Vision (ECCV' 06), pp. 404–417 (2006)
4. Bhanu, B., Chen, H.: Human Ear Recognition by Computer. Springer (2008)
5. Burge, M., Burger, W.: Ear biometrics for machine vision. In: Proc. of 21st Workshop of the Austrian Association for Pattern Recognition (WAAPR' 97). Hallstatt (1997)
6. Burge, M., Burger, W.: Ear biometrics in computer vision. In: Proc. of Int'l Conference on Pattern Recognition (ICPR' 00), vol. 02, pp. 822–826 (2000)
7. Bustard, J., Nixon, M.: Robust 2d ear registration and recognition based on sift point matching. In: Proc. of Int'l Conference on Biometrics: Theory, Applications and Systems (BTAS' 08), pp. 1–6 (2008)
8. Chang, K., Bowyer, K.W., Sarkar, S., Victor, B.: Comparison and combination of ear and face images in appearance-based biometrics. IEEE Transactions on Pattern Analysis and Machine Intelligence **25**(9), 1160–1165 (2003)
9. Choras, M.: Ear biometrics based on geometrical feature extraction. Electronic letters on computer vision and image analysis **5**(3), 84–95 (2005)
10. Choras, M.: Further developments in geometrical algorithms for ear biometrics. In: Proc. of 4th Intl Conference on Articulated Motion and Deformable Objects (AMDO' 06), pp. 58–67 (2006)
11. De Marsico, M., Michele, N., Riccio, D.: Hero: Human ear recognition against occlusions. In: Proc. of Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 178–183 (2010)
12. Freeman, W.T., Adelson, E.H.: The design and use of steerable filters. IEEE Transactions on Pattern Analysis and Machine Intelligence **13**(9), 891–906 (1991)
13. Hurley, D., Nixon, M., Carter, J.: A new force field transform for ear and face recognition. In: Proc. of Int'l Conference on Image Processing (ICIP' 00), vol. 1, pp. 25–28 (2000)
14. Hurley, D., Nixon, M., Carter, J.: Force field energy functionals for image feature extraction. Image and Vision Computing **20**(5-6), 311–317 (2002)
15. Hurley, D.J., Nixon, M.S., Carter, J.N.: Automatic ear recognition by force field transformations. In: Proc. of IEE Colloquium: Visual Biometrics, pp. 7/1–7/5 (2000)
16. Hurley, D.J., Nixon, M.S., Carter, J.N.: Force field feature extraction for ear biometrics. Computer Vision and Image Understanding **98**(3), 491–512 (2005)
17. Iannarelli, A.: Ear Identification. Paramont Publishing (1989)
18. Jayaraman, U., Prakash, S., Gupta, P.: Indexing multimodal biometric databases using Kd-tree with feature level fusion. In: Proc. of 4th Int'l Conference on Information Systems Security (ICISS' 08), LNCS 5352, pp. 221–234. Hyderabad, India (2008)

19. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision **60**(2), 91–110 (2004)
20. Nanni, L., Lumini, A.: A multi-matcher for ear authentication. Pattern Recognition Letters **28**(16), 2219–2226 (2007)
21. Nanni, L., Lumini, A.: Fusion of color spaces for ear authentication. Pattern Recognition **42**(9), 1906–1913 (2009)
22. Prakash, S., Jayaraman, U., Gupta, P.: Connected component based technique for automatic ear detection. In: Proc. of 16th IEEE Int'l Conference on Image Processing (ICIP' 09), Cairo, Egypt, pp. 2741–2744 (2009)
23. Shailaja, D., Gupta, P.: A simple geometric approach for ear recognition. In: Proc. Int'l Conference on Information Technology, pp. 164–167 (2006)
24. Štruc, V., Pavešić, N.: Illumination invariant face recognition by non-local smoothing. In: Proc. of Joint COST 2101 and 2102 Int'l Conference on Biometric ID Management and Multimodal Communication (BioID MultiComm' 09), LNCS 5707, pp. 1–8 (2009)
25. Yuan, L., hua Wang, Z., chun Mu, Z.: Ear recognition under partial occlusion based on neighborhood preserving embedding
26. Zhang, H., Mu, Z., Qu, W., Liu, L., Zhang, C.: A novel approach for ear recognition based on ICA and RBF network. In: Proc. of 4th Int'l Conference on Machine Learning and Cybernetics (CMLC' 05), pp. 4511–4515 (2005)
27. Zuiderveld, K.: Graphics Gems IV, chap. Contrast limited adaptive histogram equalization, pp. 474–485. Academic Press Professional, Inc. (1994)