The Effect of Time on Ear Biometrics

Mina I. S. Ibrahim, Mark S. Nixon and Sasan Mahmoodi
School of Electronics and Computer Science
University of Southampton, UK, SO17 1BJ
[mis07r|msn|sm3]@ecs.soton.ac.uk

Abstract

We present an experimental study to demonstrate the effect of the time difference in image acquisition for gallery and probe on the performance of ear recognition. This experimental research is the first study on the time effect on ear biometrics. For the purpose of recognition, we convolve banana wavelets with an ear image and then apply local binary pattern on the convolved image. The histograms of the produced image are then used as features to describe an ear. A histogram intersection technique is then applied on the histograms of two ears to measure the ear similarity for the recognition purposes. We also use analysis of variance (ANOVA) to select features to identify the best banana wavelets for the recognition process. The experimental results show that the recognition rate is only slightly reduced by time. The average recognition rate of 98.5% is achieved for an eleven month-difference between gallery and probe on an un-occluded ear dataset of 1491 images of ears selected from Southampton University ear database.

1. Introduction

Biometrics concerns the recognition of individuals based on a feature vector extracted from their anatomical and/or behavioral characteristic, and plays a vital role in security and surveillance systems. Finding efficient descriptors to represent these characteristics is essential for many pattern recognition tasks.

Using ears as a biometric identifier has attracted much attention in the computer vision and biometric communities in recent years. The ear, characterized by the appearance of bone structures and lobes is frequently used in biometric. Ear identification has some advantages over other biometric technologies for various reasons. An ear contains a large number of specific and unique features that assist in human identification. The hypothesis that ear structure does not change significantly over time is demonstrated in the earliest studies [1]. An ear can be remotely captured without any knowledge or consent of the person under examination. Ear recognition is not affected by changes similar to those associated with facial expression. These properties make ears very attractive as a biometric identifier. As a result, the ear biometric is suitable for security, surveillance, access control and monitoring applications.

There are some studies which show how the ear can be used for recognition, using 2D and 3D images [2, 3]. Iannarelli [1] performed two early studies suggesting ears are unique to individuals and supporting the use of an ear as a biometric modality.

One of the earliest works studying ear biometrics is that of Burge et. al [4], investigating the human ear as a biometric in the context of machine vision. They introduce a graph matching based algorithm for ear authentication by using the Voronoi diagram of the ear's edge segments to describe the ear. Moreno et al. [5] investigate the use of outer ear images for human identification. Force field transformation is used in [6] to reduce the dimensionality of the original pattern space. The algorithm consists of two stages: i) image to force field transformation, and ii) force field feature extraction. The most popular method used to extract the feature vectors of ear images in ear biometrics research is PCA [7, 8]. Arbab-Zavar et. al [9] propose the first model based approach for ear recognition, by the virtue of evidence from embryonic development of the human ear. In their work, the model based technique is combined with the Log-Gabor coefficients for recognition. Log-Gabor filter is used to extract the frequency content of the ear fluctuating surface representing the shapes of the Helix and the Anti-helix in an ear. Chen et. al [10] introduce a two-step Iterative Closest Point (ICP) algorithm to match 3D ears.

All of these techniques for ear recognition have mainly focused on recognition rate. The notion that ear is relatively time invariant is originally noted in [1]. However there is no experimental study to show the time invariance property associated with ear biometric. In this paper, the first experimental framework is proposed to investigate the effect of time on ear recognition.

A brief description of the recognition algorithm proposed in this paper to demonstrate the time invariance property of the ear biometrics is as follows: 1) For ear recognition, we shall need to locate the ear in a profile image manually to avoid the effects of ear detection process on the recognition. 2) It appears appropriate to consider a technique to describe an ear as a feature vector depending

978-1-4577-1359-0/11/$26.00 ©2011 IEEE

on its general structure. The idea of ear description technique was proposed by Zhang et al. in [11] by using Gabor wavelets. However instead of using Gabor wavelets, we employ here a bank of banana wavelets, which can be considered as generalized Gabor wavelets, to extract curvilinear structures in the ear. In addition to the frequency and orientation, banana wavelets are also characterized with properties associated with the bending and curvature of the filter. Ear images are characterized with structures similar to those of banana wavelets. Banana wavelets appear well matched to the curvilinear structure of ears, particularly in the region of the helix (the uppermost part of the ear) and the tragus (which are the lower parts). 3) ANOVA technique is used to identify the best banana wavelets for the recognition process.

This paper is structured as follows. Section 2 gives a brief background on banana wavelets, analysis of variance, and the local binary pattern techniques. Section 3 describes our proposed technique to recognize ears. The experimental results are presented in section 4. Finally, the conclusions are drawn in section 5.

2. Basic Techniques

2.1. Banana Wavelets

Banana wavelets are a generalization of Gabor wavelets which are localized filters derived from a mother wavelet [12], particularly suited to curvilinear structures (figure 1).

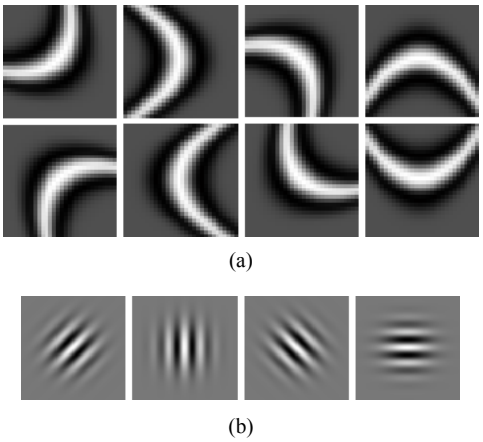
A banana wavelet $B^{\mathbf{b}}$ is parameterized by a vector $\mathbf{b}$ of four variables, i.e. $\mathbf{b}=(f,\alpha,c,s)$ where f, α, c and s are frequency, orientation, curvature, and size respectively. This filter consists of two components i) a rotated and curved complex wave function $F^{\mathbf{b}}(x,y)$ and ii) a Gaussian $G^{\mathbf{b}}(x,y)$ function rotated and curved in the same way as $F^{\mathbf{b}}(x,y)$ [12], i.e.:

$$B^{\mathbf{b}}(x,y)=\gamma\cdot G^{\mathbf{b}}(x,y)\cdot\left(F^{\mathbf{b}}(x,y)-DC^{\mathbf{b}}\right) \quad (1)$$

where $G^{\mathbf{b}}(x,y)=\exp\left(-\frac{f^2}{2}\cdot\left(\frac{\left(x_c+c\cdot x_s^2\right)^2}{\sigma_x^2}+\frac{x_s^2}{\sigma_y^2\cdot s^2}\right)\right)$

$F^{\mathbf{b}}(x,y)=\exp\left(i\cdot f\cdot\left(x_c+c\cdot x_s^2\right)\right)$, $x_c=x\cdot\cos\alpha+y\cdot\sin\alpha$, $x_s=-x\cdot\sin\alpha+y\cdot\cos\alpha$, γ is a constant, $DC^{\mathbf{b}}=e^{-\sigma_x/2}$ is the bias of the banana wavelets, and σ_x and σ_y are the scales of the Gaussian filter in x and y directions respectively.

Features can be extracted from any image by the banana wavelet transform to describe both spatial frequency structure and curvilinear structure present in the image. The convolution of the image with complex banana filters with various frequencies, orientations, curvatures, and sizes, captures the local structure of the image.



(a)

(b)

Figure 1: (a) Banana wavelet filters and (b) Gabor wavelet filters

2.2. Analysis of Variance (ANOVA)

ANOVA is a well established statistical method to provide a statistical analysis of data [13]. The importance of a feature can be realized by sorting its corresponding F-ratio in a descending order. F-ratio is calculated by comparing between classes' mean squares with the within classes' mean squares.

2.3. Local Binary Pattern (LBP)

The LBP texture operator is currently one of the most popular techniques for texture description and it has become a popular approach in various applications. The original LBP [14] operator labels every pixel of an image by thresholding the 3 x 3 neighborhood of each pixel with the value of the center pixel and considering the result as a binary number. Then, the values of the pixels in the thresholded neighborhood are multiplied by the weights given to the corresponding pixels. Finally, the values of the eight pixels are summed to obtain the LBP number for the center pixel. The histogram of these 2^8 = 256 different labels can then be used as a texture descriptor. The advantage of using LBP texture operator is that it is invariant with illumination and shift.

3. The Proposed Recognition Technique

3.1. Ear Description

To describe the ear with a feature vector we apply the representation method proposed in [11], but instead of using Gabor wavelet filters, we use banana wavelet filters. The structure of the ear mainly contains features similar to

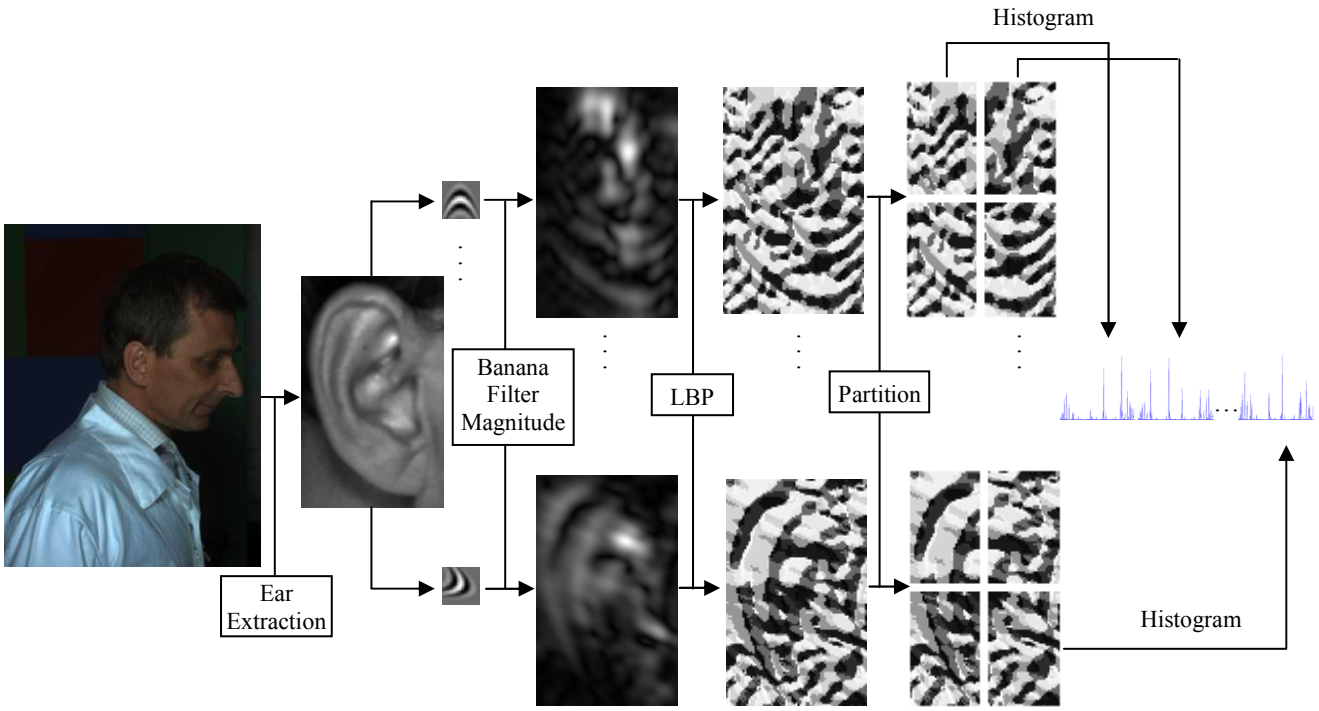


Figure 2: Ear description algorithm

those of banana wavelets. An ear image can therefore be described as a histogram sequence by the following procedure (figure 2):

- The ear is extracted manually from a head profile image. The size of extracted ear images is different. This paper is largely concerned with ear recognition rather than ear detection, however it is worth noting that the banana wavelet detection algorithm proposed in [15] can achieve a detection accuracy of 100% on XM2VTS face profile database and of 85% at Soton database. Here, the ear is extracted manually in order to minimize the effect of detection on the recognition performance.
- The magnitude of the banana filter responses $M^{\mathbf{b}}$ is calculated by taking the magnitude of convolving banana wavelet filters $B^{\mathbf{b}}$ for all possible value of $\mathbf{b}$ (80 filters are selected by using ANOVA, this will be discussed in section 4) with the extracted ear image I:

$$M^{\mathbf{b}} = \left| \left(B^{\mathbf{b}} * I \right)(x_0) \right| \qquad (2)$$

  where x_0 and $*$ are the location of a pixel in I, and the convolution operator, respectively.
- Local binary pattern $LBP^{\mathbf{b}}$ is calculated for each $M^{\mathbf{b}}$ which has invariant characteristics.
- Each $LBP^{\mathbf{b}}$ is divided into four non-overlapping regions which expose basic structures in the ear.
- Histogram $H_{r,n}$ calculated by applying filter n (n is from 1 to 80) on region r (r varies from 1 to 4) is used to produce the ear features.
- All of the calculated histograms are concatenated into a single histogram sequence HS to describe the given ear image as a feature vector:

$$HS = \left(H_{1,1}, H_{2,1}, H_{3,1}, H_{4,1}, H_{1,2}, ..., H_{4,80} \right) \qquad \mathbf{(3)}$$

3.2. Ear Comparison

Histogram intersection HI [16] is used as a similarity measure between two histograms. We normalize each histogram to produce a probability distribution function. The histogram intersection can then be calculated as:

$$HI\left(H^1, H^2\right) = \sum_{i=0}^{L} \min\left(\frac{h_i^1}{\sum_{j=0}^{L} h_j^1}, \frac{h_i^2}{\sum_{j=0}^{L} h_j^2} \right) \qquad (4)$$

where h^1 and h^2 are two histograms, and L is the number of bins in the histogram (in this case L equals 255). Once the ear is described as histogram sequence, the similarity S between two histogram sequences is measured by calculating HI between them:

$$S\left(HS^1, HS^2\right) = \sum_{r=1}^{4} \sum_{n=1}^{80} HI\left(H_{r,n}^1, H_{r,n}^2\right) \qquad (5)$$

To select the matched ear with the test one, the similarities between the histogram sequence of the test ear and all of the other histogram sequences are measured, and the ear with maximum S is selected as the correct one.

4. Results

4.1. Database Description

There are two databases used in our experiments. One database is used to select the best banana filters to describe the ear, and the other database is used to evaluate the effect of time on the performance of ear recognition:

1) XM2VTS database: First ear database is selected from the XM2VTS face profile database [17]. The selected dataset consists of 252 images from 63 individuals with four images per person collected during four different sessions over a period of five months to ensure the natural variation between the images of the same person. The images selected are those where the whole ear is visible in a 720×576 24-bit image. The ears in the database are not occluded by hair but there are few images with some occlusion by earrings. This is the same subset of the XM2VTS face profile database used by Hurley et al. [6] and Arbab-Zavar et. al [9]. Figure 3-a illustrates images from XM2VTS database.

2) SOTON ear database: Two datasets are selected from the SOTON ear database [18, 19]. The database is acquired as subjects walk past a camera triggered by a light beam signal, and other biometrics are acquired at the same time (face and gait biometrics). The subjects are between 20 and 55 years old. The advantage of this database is that, it has much variation in ear orientation, size, color skin, and lighting condition. This database therefore allows evaluation of the performance of our technique on a data acquired in a more realistic scenario. Dataset 1 [18] contains 548 face profile images from 137 subjects with four images per person (figure 3-b). Dataset 2 [19] is selected to enable analysis to be performed over time which is captured in five different sessions over a period of eleven months (figure 3-c). There are 25 subjects used in these sessions. The number of subjects in each session may vary. The number of the images for each subject may also change because we select only the images without ear occlusion. The number of subjects and images available in the SOTON database for each session is shown in table 1.

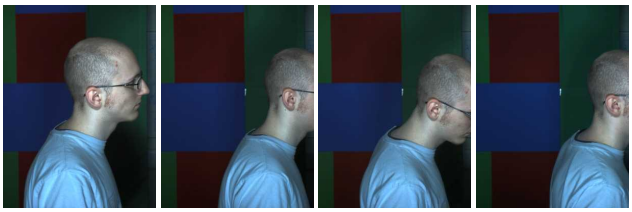
4.2. Filters Selection

ANOVA based feature selection algorithm is applied to select the best filters from a bank of banana filters used in the recognition process. To select the best filters, first a bank of banana filters (720 filters) is convolved with XM2VTS database images and the magnitudes of the filters' responses are computed (XM2VTS database is used for filter selection as it is constrained and has fewer variations). Then, LBP algorithm is applied on each magnitude image to produce a feature image which is then divided into four parts (quadrants). The histogram is calculated for each part and the four histograms are concatenated for each filter.

Month	Number of subjects	Number of images
0	18	337
1	18	350
4	14	234
9	12	231
11	15	339

Table 1. Number of Subjects and Images in each Session

(a) XM2VTS face profile database



(b) SOTON ear dataset 1

(c) SOTON ear dataset 2

Figure 3: Images from XM2VTS and SOTON databases

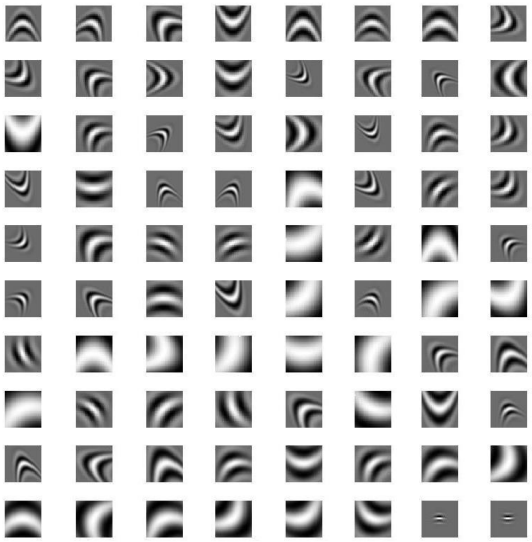


Figure 4: The selected banana filters

After that, ANOVA is applied to compute the F-ratio for each filter. Finally, the filters are ranked in descending order according to their F-ratio and the first 80 filters are selected as the best filters. Figure 4 illustrates the selected filters. Having selecting the best filters, our recognition technique is then applied to XM2VTS face profile database to obtain a recognition rate of 98%. A recognition rate of 99.3% is also achieved by applying the algorithm proposed here to SOTON ear dataset 1. We also investigated use of uniform LBP but recognition was not as good as for basic LBP.

4.3. The Effect of Time

The effect of time on the performance of ear recognition is investigated by applying our proposed recognition technique to SOTON ear dataset 2. The average recognition rate achieved over eleven months is 98.5%. Table 2 shows the combinations of probe and gallery which are used in this experiment. The recognition rate for each combination is shown in figure 5. The recognition rate for the time difference of 0, 2, 4, 5, 10 and 11 months is nearly 100%. But the recognition rate drops by a value between 2% and 5% for the time difference corresponding to 1, 3, 7, 8 and 9 months due to the higher values of rotation (figure 6-a). Our recognition approach is relatively immune to slight rotation (figure 6-b).

4.4. Verification Performance

The verification performance of the proposed technique is measured in terms of the receiver operating characteristic (ROC) curve and decidability measure. The ROC curve describes a relation between the false accept rate (FAR) and the false reject rate (FRR) at different thresholds. Figure 7 shows the ROC curve of the proposed technique on XM2VTS face profile database and SOTON ear dataset 1. The equal error rate (EER) of the experiments is 7.4% for XM2VTS database and 5.2% for SOTON ear dataset 1. The EER is the location on the ROC curve (figure 7) where the FAR and the FRR are equal. The decidability value provides an indication of mean separation with respect to standard deviation. The measured decidability [20] is 3.27 for XM2VTS database and 3.5 for SOTON ear dataset 1. For a database, the higher the decidability measure, the better the algorithm is performance.

Probe (month)	Gallery (month)	Time Difference (months)
0	0	0
0	1	1
9	11	2
1	4	3
0	4	4
4	9	5
4	11	7
1	9	8
0	9	9
1	11	10
0	11	11

Table 2. The Combinations of Probe and Gallery

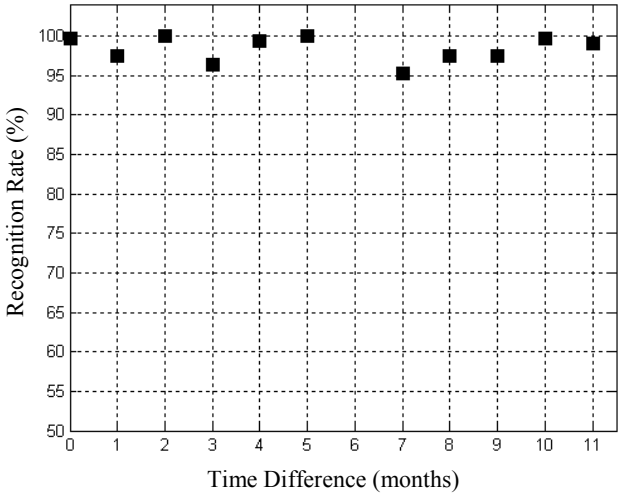
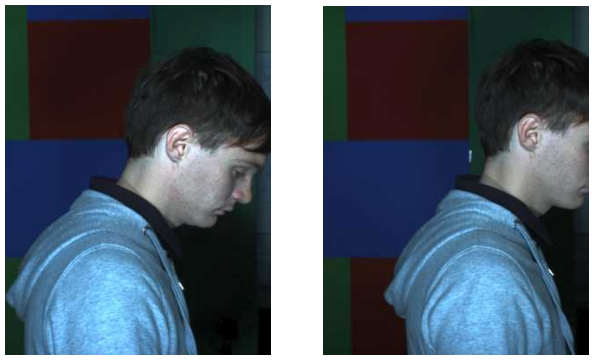


Figure 5: Recognition rate over time

(a) Highly rotated ears for the same person



(b) Slightly rotated ears for the same person

Figure 6: Examples of highly and slightly rotated ears

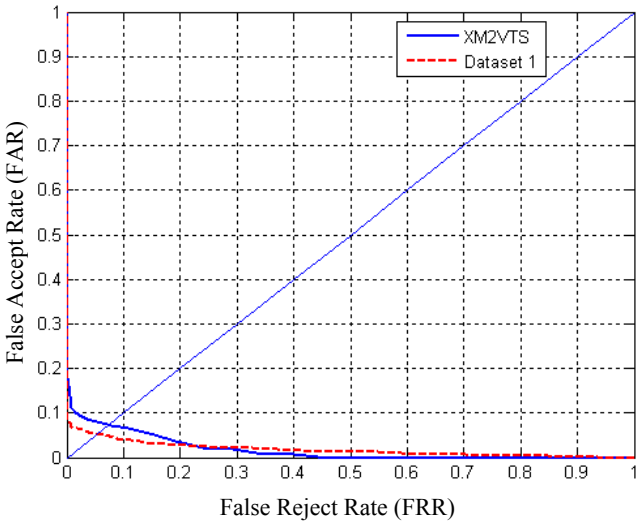


Figure 7: The ROC curve for XM2VTS and SOTON databases

5. Conclusions

The main contribution of this paper is to test the effect of time on the performance of the ear recognition. This is the first biometric experimental study to investigate the time effect on ear biometrics. The recognition approach is based on the description of the ear image as histogram sequences. Our method to describe the ear depends on the local binary pattern algorithm. It is therefore robust to local image transformation due to variations of lighting and shifting. One of major strength of our proposed method is that it appears to be general for various databases. The general structure of the ear has many curvilinear structures which match the structures of the banana filters. We apply a feature selection algorithm by using ANOVA based feature selection method to select the best filters from a bank of banana filters.

The experiments show that when our proposed recognition technique is applied to XM2VTS face profile database and SOTON ear dataset 1, a recognition rate of 98% and 99.3% are achieved; respectively. The average recognition rate over eleven months is also 98.5% for the SOTON ear dataset 2. The experiments show that the recognition rate is not affected considerably over eleven months. We can therefore conclude that ear can be used in various applications as a time invariant biometric. Clearly a database gathered over a longer timeframe could extend the results achieved here, but an 11 month timeframe is all that is currently available.

References

[1] A. Iannarelli, *Ear Identification*, Paramount Publishing Company, Freemont, California, 1989

[2] D. J. Hurley, B. Arbab-Zavar, and M. S. Nixon: The ear as a biometric. In Jain, A., Flynn, P., Ross, A., editors, *Handbook of Biometrics*. 2008

[3] B. Bhanu, and H. Chen: Human Ear Recognition by Computer. Springer 2008

[4] M. Burge, and W. Burger, Ear biometrics, in: A. Jain, R. Bolle, S. Pankanti (Eds.), *Biometrics: Personal ID in Networked Society*, Kluwer 1998, pp. 273-286.

[5] B. Moreno and A. Sanchez, On the Use of Outer Ear Images for Personal Identification in Security Applications, *IEEE 33rd Intl. Conf. on Security Technology,* 1999, pp. 469-476.

[6] D. J. Hurley, M. S. Nixon, and J. N. Carter, Force field feature extraction for ear biometrics, *CVIU*, 2005, vol. 98, pp. 491-512.

[7] B. Victor, K.W. Bowyer, and S. Sarkar, An evaluation of face and ear biometrics, Proc. *ICPR 2002*, pp. 429-432.

[8] K. Chang, K.W. Bowyer, S. Sarkar, and B. Victor, Comparison and combination of ear and face images in appearance-based biometrics, *IEEE Trans. PAMI*, 2003, vol. 25, no. 9, pp. 1160-1165.

[9] B. Arbab-Zavar, M. S. Nixon, On Guided Model-Based Analysis for Ear Biometrics, *CVIU,* 2011, vol. 115, pp. 487-502.

[10] H. Chen and B. Bhanu, Contour matching for 3-D ear recognition, Proc. *IEEE Workshop on Applications of Computer Vision*, Colorado, 2005.

[11] W. C. Zhang, S. G. Shan, W. Gao, et al, "Local Gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition," in *Proc. ICCV,* pp. 786-791, 2005.

[12] N. Krüger, M. Pötzsch, and G. Peters. Principles of Cortical Processing Applied to and Motivated by Artificial Object Recognition*, in Information Theory and the Brain,* Cambridge University Press, 2000.

[13] G. M. Clarke and D. Cooke, *A Basic Course in Statistics,* chapter 22, pp. 520-546. Arnold, 1998.

[14] T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions", *Pattern Recog.*, **29**, no. 1, pp.51-59, 1996.

[15] M. I. Ibrahim, M. S. Nixon, and S. Mahmoodi, "Shaped Wavelets for Curvilinear Structures for Ear Biometrics." In: *ISVC,* Nov. 2010, Las Vegas, USA.

[16] M. Swain and D. Ballard. Color indexing. *International Journal of Computer Vision*, 1991, vol. 7, no 1, pp. 11-32.

[17] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, XM2VTSDB: The Extended M2VTS Database, *Proc. AVBPA'99* ,Washington D.C., 1999

[18] D. Matovski, M. Nixon, S. Mahmoodi, and J. Carter, The Effect of Time on the Performance of Gait Biometrics. In: *IEEE BTAS 2010,* Sept., Washington DC, USA, 2010

[19] S. Samangooei, J. Bustard, M. S. Nixon, and J. N. Carter, On Acquisition and Analysis of a Dataset Comprising of Gait, Ear and Semantic Data. In: B. Bhanu, V. Govindaraju, *Multibiometrics for Human Identification*. CUP (2011)

[20] J. Daugman. Biometric decision landscapes. Technical Report TR482, University of Cambridge, Computer Laboratory, 2000.