Shaped Wavelets for Curvilinear Structures for Ear Biometrics

Mina I. S. Ibrahim, Mark S. Nixon, and Sasan Mahmoodi

ISIS, School of Electronics and Computer Science, University of Southampton, UK
{mis07r, msn, sm3}@ecs.soton.ac.uk

Abstract. One of the most recent trends in biometrics is recognition by ear appearance in head profile images. Determining the region of interest which contains the ear is an important step in an ear biometric system. To this end, we propose a robust, simple and effective method for ear detection from profile images by employing a bank of curved and stretched Gabor wavelets, known as banana wavelets. A 100% detection rate is achieved here on a group of 252 profile images from XM2VTS database. The banana wavelets technique demonstrates better performances than Gabor wavelets technique. This indicates that the curved wavelets are advantageous here. Also the banana wavelet technique is applied to a new and more challenging database which highlights practical considerations of a more realistic deployment. This ear detection technique is fully automated, has encouraging performance and appears to be robust to degradation by noise.

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. Any automatic biometric system needs detection and partitioning process to extract the region of interest from the background.

Ear as a biometric identifier has attracted much attention in the computer vision and biometric communities in recent years. Ear, which is characterized by the appearance of the outer ear, lobes and bone structures 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. It contains a rich and stable structure that does not change significantly over time [1]. An ear can be remotely captured without any knowledge or consent of the person under examination. It also does not suffer from changes in 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.

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

The 2D approaches use the ear as a planar structure affixed to the head. Alternatively, 3D approaches can be used and this has so far been achieved with range scan data. Ear detection is the most important step in an ear recognition system, and the detection quality will therefore affect directly the performance of the whole recognition system.

Recent approaches mostly focus on ear recognition without a fully automated method for ear detection [1, 3]. However automated schemes have recently been proposed for ear detection prior to recognition. Some researchers have focused on 2D [4, 5, 6] and 3D [7, 8] ear detection. The two most sophisticated approaches in 2D ear detection are proposed in [4] and [5]. Islam et al. [4] modifies the cascaded AdaBoost approach to detect the ear from 2D profile images in a learning method by using a training data of ear images. They report good results on large size databases. However if the ear image is rotated with the respect to the training data or if its appearance is different from the ears in the training data, their method could fail to detect the ear, because the training data does not contain example test data in such cases. Forming a database of rotated ears will require much more storage than that required for a technique which is inherently immune to change in feature orientation. Arbab-Zavar et al. [5] propose an ear detection algorithm based on the elliptical shape of the ear by using a Hough transform. Their method is robust with occlusion; however their ear detection algorithm only works under some specific conditions applied to the images of the database to avoid errors caused by the presence of nose and/ or spectacles.

We contend that it is prudent to continue investigating approaches which consider the ear as a planar surface. This will allow for application in access control and surveillance, and for acquisition from documents. The ear plane is not aligned to that of the head, but it is chosen as such we can consider the ear to be on a flat surface. For ear detection, we shall need to locate the ear in a profile image automatically with an algorithm which is robust at the presence of noise. It therefore appears appropriate to investigate a technique which depends on the general structure of the ear.

We employ a bank of banana wavelets, which are generalized Gabor wavelets, to extract curvilinear structures. In addition to the frequency and orientation, banana wavelets are also characterized with properties associated with the bending and curvature of the filter. The ear is an image structure mainly contains features which are similar to those of banana wavelets. These features then appear well matched to the general structure of the ear which has many curvilinear structures, particularly in the region of the helix (the uppermost part of the ear) and the tragus (which are the lower parts).

This paper is structured as follows. Section 2 gives a brief background on banana wavelet filters. Section 3 describes our new technique to detect the ear. The extraction results are provided in section 4. Finally, the conclusions are presented in section 5.

2 Banana Wavelets

Banana wavelets are a generalization of Gabor wavelets and are localized filters derived from a mother wavelet [9], particularly suited to curvilinear structures.

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 is built from a rotated and curved complex wave function $F^{\mathbf{b}}(x,y)$ and a Gaussian $G^{\mathbf{b}}(x,y)$ function rotated and curved in the same way as $F^{\mathbf{b}}(x,y)$ [9]:

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

where $G^{\mathbf{b}}(x,y)=\exp\left(-\dfrac{f^2}{2}\cdot\dfrac{1}{\sigma_x^2}\cdot\left(x_c+c\cdot x_s^2\right)^2+\dfrac{1}{\sigma_y^2}\cdot\dfrac{1}{s^2}\cdot x_s^2\right)$, $x_c=x\cdot\cos\alpha+y\cdot\sin\alpha$,

$x_s=-x\cdot\sin\alpha+y\cdot\cos\alpha$, $F^{\mathbf{b}}(x,y)=\exp\!\left(i\cdot f\cdot\left(x_c+c\cdot x_s^2\right)\right)$, γ is a constant, chosen empirically, σ_x and σ_y are the scales of the Gaussian filter in x and y directions re-

spectively, and $DC^{\mathbf{b}}=e^{-\frac{\sigma_x}{2}}$ is the bias of the banana wavelets.

Any image can be represented by the banana wavelet transform allowing the description of both spatial frequency structure and spatial relations. The convolution of the image with complex banana filters with different frequencies, orientations, curvatures, and sizes, captures the local structure points of an object.

3 Ear Detection

We argue that any ear contains curvilinear structures such as helix, anti-helix and inter-tragic notch. The essence of our ear detection technique is to initially calculate the magnitude of the filter responses $AI(\mathbf{x_0},\mathbf{b})$ by convolving a banana wavelet $B^{\mathbf{b}}$ with an image I and then to find the positions where this magnitude has local maxima, i.e.:

$$AI(\mathbf{x}_0,\mathbf{b})=\arg\left(\max_{\mathbf{x}_0}\left(\left|FI(\mathbf{x}_0,\mathbf{b})\right|\right)\right) \tag{2}$$

where $FI(\mathbf{x}_0,\mathbf{b})=\left(B^{\mathbf{b}}*I\right)(\mathbf{x}_0)$ and $\mathbf{x}_0$ is the position of a pixel in an input image I. A banana wavelet $B^{\mathbf{b}}$ produces a strong response at pixel position $\mathbf{x}_0$ when the local structure of the image at that pixel position is similar to $B^{\mathbf{b}}$. An input ear image and the response magnitudes, which are calculated by convolving the input image with the filters depicted in Fig. 1, are shown in Fig. 2. In this figure, white pixels represent high values in the response magnitudes. Therefore there are local maxima (highlighted) at those positions where ear has similar curvature, orientation, and size to those of the corresponding banana wavelet.
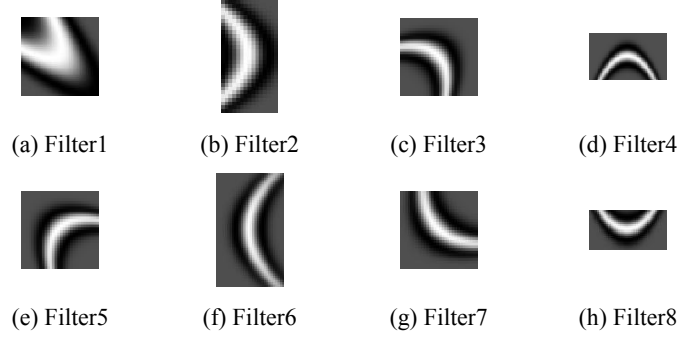
(a) Filter1   (b) Filter2   (c) Filter3   (d) Filter4

(e) Filter5   (f) Filter6   (g) Filter7   (h) Filter8
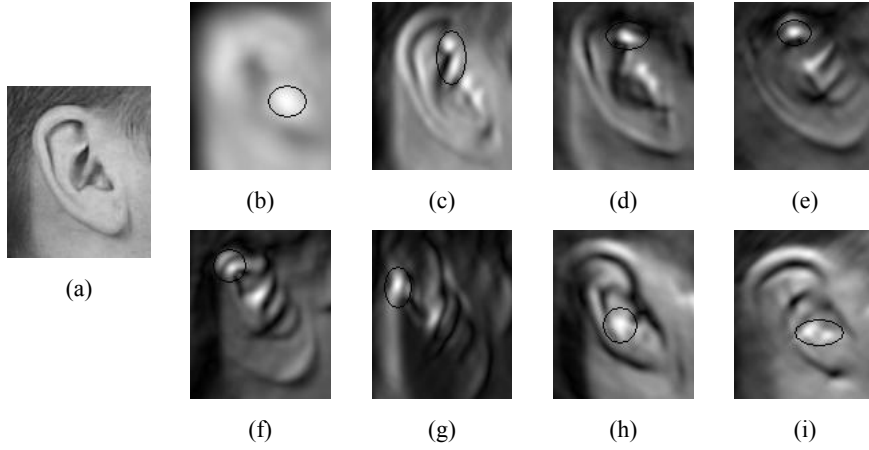
Fig. 1. (a)-(h) 8 filters used in this work

Fig. 2. (a) Input image, and (b)-(i) after convolution with 8 banana filters.

A position of interest is selected by considering two conditions: i) the response magnitude has to represent a local maximum ($Q1$) and ii) its value should be greater than a certain threshold ($Q2$):

$$Q1: AI(\mathbf{x}_0,\mathbf{b}) \geq AI(\mathbf{x}_i,\mathbf{b}) \qquad\qquad \mathbf{x}_i \in N(\mathbf{x}_0) \tag{3}$$

$$Q2: AI(\mathbf{x}_0,\mathbf{b}) > \lambda \cdot T(\mathbf{x}_0) \tag{4}$$

where λ is a constant, $T(\mathbf{x}_0) = 0.25 \cdot (E(I) + E(I,\mathbf{x}_0))$, $E(I) = \sum_{\mathbf{x}\in I}\sum_{\mathbf{b}\in B} AI(\mathbf{x},\mathbf{b})$, $E(I,\mathbf{x}_0) = \sum_{\mathbf{x}\in N(\mathbf{x}_0)}\sum_{\mathbf{b}\in B} AI(\mathbf{x},\mathbf{b})$, and $N(\mathbf{x}_0)$ represents a square window with center $\mathbf{x}_0$ and length of side w.

In addition to the conditions (Q1) and (Q2), the spatial arrangement of the positions of local maxima should match a template representing the ear structure (Fig. 3 illu-

strates this template). Locations of the local structure points in this template are general for any ear. For example, the convolution of filter 4 with any ear, produces a strong response at the top middle part as illustrated in Fig. 2-e.
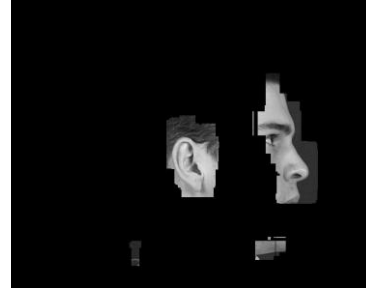

Fig. 3. The ear template   **Fig. 4.** Regions of interest

To reduce computational burden, the ear detection process starts with a coarse search by applying the banana wavelets to the whole image to extract the regions of interest containing curved lines in order to perform a finer search for the ear within these regions (Fig. 4). These regions are much smaller than the whole image and therefore the fine search performed by applying the banana wavelets to these regions, requires less computational demand.

In the fine search, the regions of interest are divided into a group of smaller neighborhoods and convolved with a bank of banana filters (8 filters in this paper are chosen, as shown in Fig. 1) to calculate positions corresponding to local maxima in each neighborhood. The neighborhood with maximum number of positions matching the ear template and meeting conditions (Q1) and (Q2) is considered as the neighborhood containing the ear. In the case that many overlapping neighborhood windows are detected, only one region is selected which contains maximum percentage of overlapped windows. Our technique is generic and applicable to any database. The parameters of the 8 filters are chosen by experiments (Table 1). R and C in the table are the number of rows and columns of the banana wavelet filters, respectively, which are used in the convolution process between banana wavelet filters and the image.

Table 1. Parameter Settings for the Banana Wavelets

	f	α	c	s	R	C
Filter 1	0.05	$\pi/4$	0.1	1	50	50
Filter 2	0.28	$\pi/2$	0.05	1	30	15
Filter 3	0.28	$3\pi/4$	0.05	1	30	30
Filter 4	0.28	π	0.05	1	30	50
Filter 5	0.28	$5\pi/4$	0.05	1	30	30
Filter 6	0.28	$3\pi/2$	0.02	1	50	30
Filter 7	0.28	$7\pi/4$	0.03	1	30	30
Filter 8	0.28	2π	0.05	1	20	40

4 Results

Our primary purpose is to evaluate success in ear detection. The efficiency of banana wavelet technique is tested using a database of 2D images selected from the XM2VTS face profile database [10]. Our database 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. [11] and Arbab-Zavar et al. [5].

The new technique correctly detects all the ears in the images in the database (the detection rate was 100%). Some results of detection using banana wavelets are shown in Fig. 5. The system is fully automatic and it does not require any manual interference for ear detection. As such the approach appears suitable for real time biometric applications. The parameters used in these results are $\sigma_x = 1$, $\sigma_y = 2\pi$, $\gamma = 0.08$, $\lambda = 1.8$, and $w = 7$.

Banana wavelet filters can capture the curved structures better than Gabor wavelet filters. To show this, Gabor wavelet technique is applied to the same subset of the XM2VTS face profile database with the same filter sizes, orientations and frequencies as those of banana wavelet filters (the same parameters in Table 1, except that curvature c was set to zero). The detection rate obtained by Gabor wavelets is 97.2%.

Banana wavelet technique is robust to degradation of images such as the motion blur, partly shown in Fig. 5-a. It is also accurate and robust to some rotations (note the large subject rotation in Fig. 5-b). In addition to that, the technique is robust at presence of earrings and/or glasses, as shown in Fig. 5-c.
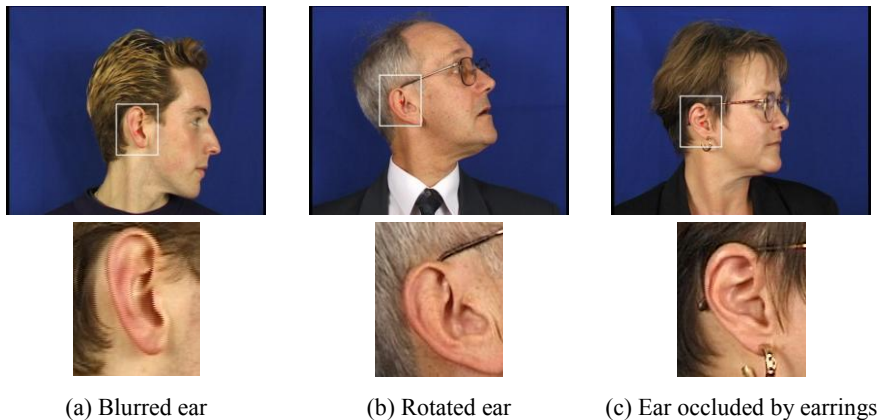

| (a) Blurred ear | (b) Rotated ear | (c) Ear occluded by earrings |

Fig. 5. Samples of ear detection using our technique

It also appears robust to noise. The accuracy of detection in presence of noise is more than 98% when the noise standard deviation σ is quite high. These results are illustrated in Fig. 6. Here, images are contaminated by additive zero mean Gaussian noise with various noise variances. Here, the technique is successful until $\sigma = 500$ in which case a region containing the eye is erroneously selected (see Fig. 6-f).
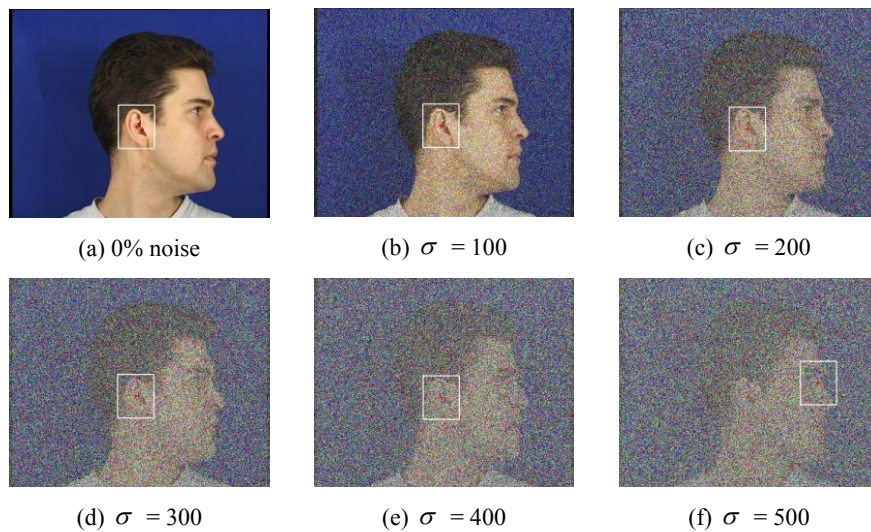

| (a) 0% noise | (b) $\sigma = 100$ | (c) $\sigma = 200$ |
| (d) $\sigma = 300$ | (e) $\sigma = 400$ | (f) $\sigma = 500$ |

Fig. 6. Samples for the results at presence of noise

The results of the noise analysis are provided in Fig. 7. Here, for noise-free images the detection rate is 100%. As expected, the detection rate drops with increasing noise. The analysis of the database is shown in Fig. 7 where at $\sigma = 100$ the recognition rate is still above 98% but this drops to under 80% when $\sigma = 200$. This is actually quite a severe level of noise, as shown in Fig. 6-c. In much worse cases, i.e., beyond the level experienced in surveillance video footage, as shown in Fig. 6-f detection rate drops to about 10%. The graph also shows the performance of the Gabor wavelet. As the noise increases, the advantages associated with using curvature become masked by the noise and in cases of severe noise the Gabor wavelet is more successful than the banana wavelets.

We also apply banana wavelet technique to detect the ears from a new database [12] from which a selection of images is shown in Fig. 8. The advantage of this database is that, it has a lot of variations of ear orientation, size, color skin, and lighting condition, allowing investigation of the performance of our technique on a data acquired in a more realistic scenario (see Fig. 8). 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. As the acquisition is largely uncontrolled, subjects sometimes present the whole head without occlusion and other combination with partial or large occlusion, and partial or sometimes the head is even absent.
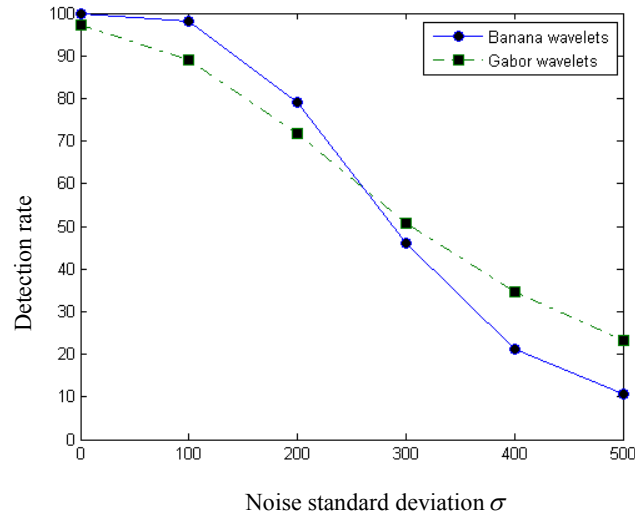
Fig. 7. Detection rate for banana wavelets technique and Gabor wavelets technique in the presence of noise

Table 2 shows the results of applying banana wavelet technique to the new database. The same parameters determined by analysis of the XM2VTS database are used in this analysis. As such, the approach is not tuned for this new database (and scenario) and it is likely that these results could be improved further. We do believe that the structure of the results will remain similar, in that some subjects' ears will remain concealed by hair, and that in a walk-through scenarios it is difficult to acquire images which consistently capture the whole ear.
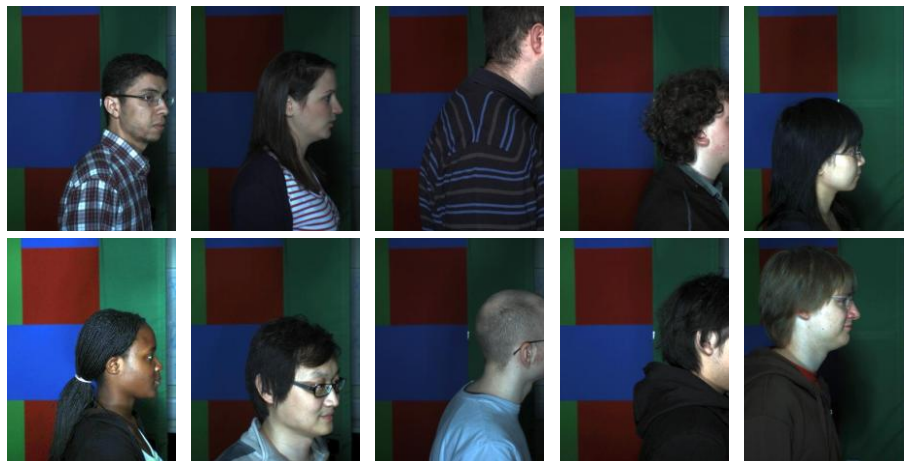
(a) Group A   (b) Group B   (c) Group C   (d) Group D   (e) Group E

Fig. 8. Samples from the new database: (a)-(e) show the groups according to Table 2

Table 2. Results of applying banana wavelet filters to the new database

	Number of samples	Success ear detection
Group A: whole head, no occlusion	885	85%
Group B: whole head, small occlusion	308	65.3%
Group C: partial head, no occlusion	653	79%
Group D: partial head, small occlusion	208	44.7%
Group E: large occlusion	383	17.8%

5 Conclusions

This paper demonstrates how banana wavelets can be used to find the ear from head profile images for biometric purposes. The complexity of the task has been reflected in the fact that ear images can vary in appearance under different viewing and illumination conditions. The experiments show that the system is effective for ear detection, which the proposed technique correctly detects all the test images selected from the XM2VTS database. The technique accrues advantages of noise tolerance and relative immunity to noise. It does not depend on a controlled lighting conditions or skin color; it therefore appears suitable for general applications. The technique proposed here is applied to a more complex database which the acquisition of this database is largely uncontrolled. The result of the ear detection for the new database is good enough according to the uncontrolled conditions, and shows the expected performance in occlusion. The performance of the banana wavelets technique is compared with that of Gabor wavelets technique which shows that banana wavelets can capture the curved structures better than Gabor wavelets. Finally, the technique proposed here is fully automated and does not require any help to detect the ear. The success of our technique relies on the fact that the selected curvilinear structures are general for any ear. We look forward to using this new approach as a primer for recognition purposes.

References

1. Hurley, D. J., Arbab-Zavar, B., Nixon, M. S.: The ear as a biometric. In Jain, A., Flynn, P., Ross, A., editors, Handbook of Biometrics. (2008)
2. Iannarelli, A.: Ear Identification. Paramount Publishing Company, Freemont, California (1989)
3. Bhanu, B., Chen, H.: Human Ear Recognition by Computer. Springer (2008)
4. Islam, S. M. S., Bennamoun, M., Davies, R.: Fast and Fully Automatic Ear Detection Using Cascaded AdaBoost. In: Proc. of IEEE Workshop on Application of Computer Vision (WACV'08), USA, pp. 1-6 (Jan 2008)
5. Arbab-Zavar, B., Nixon, M. S.: On shape mediated enrolment in ear biometrics. In: International Symposium on Visual Computing (ISVC'07), Nevada, California (Nov 2007)
6. Bustard, J. D., Nixon, M. S.: Robust 2D Ear Registration and Recognition Based on SIFT Point Matching. In: BTAS (2008)

7. Yan, P. Bowyer, K. W.: Biometric recognition using 3d ear shape. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(8), 1297–1308 (2007)
8. Chen, H., Bhanu, B.: Human ear recognition in 3d. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(4), 718–737 (2007)
9. Krüger, N., Pötzsch, M., Peters, G.: Principles of Cortical Processing Applied to and Motivated by Artificial Object Recognition. In: Information Theory and the Brain. Cambridge University Press (2000)
10. Messer, K., Matas, J., Kittler, J., Luettin, J., Maitre,G.: XM2VTSDB: The Extended M2VTS Database, Proc. AVBPA'99, Washington D.C. (1999)
11. Hurley, D. J., Nixon, M. S., Carter, J. N.: Force field feature extraction for ear biometrics. Computer Vision and Image Understanding, 98, pp. 491-512 (2005)
12. Samangooei, S., Bustard, J., Nixon, M. S. and Carter, J. N.N.: On Acquisition and Analysis of a Dataset Comprising of Gait, Ear and Semantic Data. In: Bhanu, B., Govindaraju, V.: Multibiometrics for Human Identification. CUP (2010) (in press)