

# On Shape-Mediated Enrolment in Ear Biometrics

Banafshe Arbab-Zavar and Mark S. Nixon

Department of Electronics and Computer Science, University of Southampton,  
Southampton SO17 1BJ, UK

**Abstract.** Ears are a new biometric with major advantage in that they appear to maintain their shape with increased age. Any automatic biometric system needs enrolment to extract the target area from the background. In ear biometrics the inputs are often human head profile images. Furthermore ear biometrics is concerned with the effects of partial occlusion mostly caused by hair and earrings. We propose an ear enrolment algorithm based on finding the elliptical shape of the ear using a Hough Transform (HT) accruing tolerance to noise and occlusion. Robustness is improved further by enforcing some prior knowledge. We assess our enrolment on two face profile datasets; as well as synthetic occlusion.

## 1 Introduction

Ears have long been considered as a potential means of personal identification, yet it is only in the last 10 years that machine vision experts started to tackle the idea of using ears as a biometric. Empirical evidence supporting the ear's uniqueness was provided in studies by Iannarelli [1], conducted over 10,000 ears. Ears have appealing properties for personal identification; they have a rich structure that appears to be consistent with age from a few months after birth. Clearly, ears are not affected by facial expressions. Images of ears can be acquired without the subject's participation and ears are big enough to be captured from a distance. However there exists a big obstacle - the potential occlusion by hair and earrings, which is almost certain to happen in uncontrolled environments. Up-to-date surveys of ear biometrics have recently been provided by Hurley et al. [2] and Choras [3].

Non-invasive biometrics such as gait, face and ear are often acquired with standard equipment, and they include unnecessary background information. The input samples of an ear biometric system are often head profile images with eyes, mouth, nose, hair and etc. as background. Enrolment (or registration) in an automatic biometric system is the process of detecting and isolating the area of interest.

Yan et al. [4] developed an automatic ear biometric system which inputs 2D colour images and 3D range information of the human head profiles. Their enrolment uses a multistage process which uses both 2D and 3D data and curvature estimation to detect the ear pit which is then used to initialize an elliptical active contour to locate the ear outline and crop the 3D ear data. Chen et al. [5]

also use 2D and 3D face-profile images and the ears are detected by a two step alignment procedure, which aligns edge clusters with an ear model based on the helix and anti-helix of one ear image. Alvarez et al. [6] fits an ovoid model to the ear contour. However, their algorithm initializes with a manual estimate of the ear contour. These studies have not presented an evaluation of occlusion tolerance. Yan et al. merely suggest that their method is capable of handling small amounts of occlusion by hair.

The increasing variety of biometrics applications in everyday identification and authorization problems urges the development of easily applicable methods. Therefore rather than use specialised and elaborate 3D equipment, our focus is for generic deployment via planar images derived from video cameras. The ear is largely a planar shape and 3D must penetrate the intricate inner ear, which restricts deployment potential. We propose the use of the Hough Transform (HT), which can extract shapes with properties equivalent to template matching and is tolerant of noise and occlusion, to find the elliptical shape of the ears in 2D head-profile images. We also add a number of refinements and customizations to make this technique more suitable for application to ear images. We report the enrolment results on two standard databases: (*i*) a subset of XM2VTS [7] face-profile database with the total number of 252 images, (*ii*) the UND ear database [8] with 942 images.

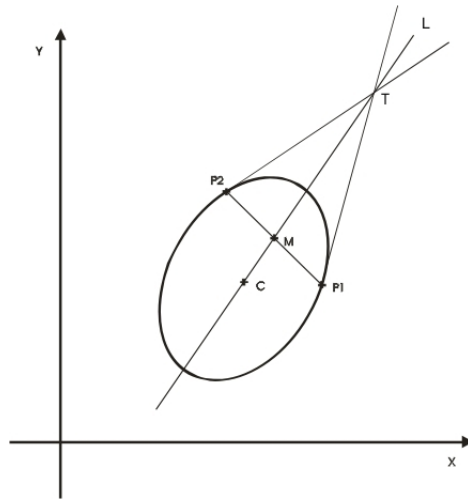
We shall describe our ear enrolment technique in section 2, describing the applied Hough Transform and our refinements which incorporate some prior knowledge to our method. In section 3 we enrol the ears of two different databases with 2D head-profile images and also evaluate our technique for partially occluded ears, followed by conclusions.

## 2 Ear Enrolment

The enrolment stage prepares an acquired sample for matching and classification by separating the object of interest from the background. It is an important step for any automatic biometric recognition system. Recognizing the potential of the ear for being partially occluded by hair or earrings, we aim to develop an automatic ear biometric system which can handle occlusion. This system should be able to handle occlusion at all stages since for example it is impossible to recognize a half-occluded ear whose data has been discarded by enrolment as a result of incorrect detection. We propose the use of the Hough Transform, which can extract shapes with properties equivalent to template matching and is tolerant of noise and occlusion. As such, it appears an attractive choice for the initial location of the ears. Our enrolment finds the elliptical shapes in 2D face-profile images to locate the ear regions by using a Hough transform to gather votes for putative ellipse centres in an accumulator, which will go through a refinement process that will eliminate some of the erroneous votes; the location of the peak in this accumulator gives the coordinates of the best matching ellipse.

## 2.1 Reduced Hough Transform for Ellipses

Despite its advantages, known drawbacks of the HT include need for much memory and its high computational requirement. Hence, a reduced HT [9] appears attractive. In this the centre is determined by using known ellipse properties. In these properties, given two points whose tangents intersect, then the line between the intersection point and their midpoint will pass through the centre of the ellipse. Figure 1 illustrates this with two points P1 and P2.



**Fig. 1.** Tangents of the two points P1 and P2 intersect at the point T. The line L which passes through the intersection point T and the midpoint M, also passes through the ellipse centre C.

For application to ear images, first a smoothed edge detected image is derived using the Canny operator, and this is used to vote into an accumulator for the centre of the ellipse. The accumulator is constructed by voting for the lines which pass through the putative ellipse centre. When this process is completed, the locations of the peaks in the accumulator give the coordinates of the centre of the best matching ellipse.

The reduced HT for ellipses decreases the computational complexity by using the known geometrical properties of the ellipse to decompose the parameter space. The space which would otherwise be a 5D has been reduced to a 2D space, gathering evidence for putative ellipse centres whilst marginalizing on scale and orientation, which may cause ellipses of arbitrary scale and orientation to cast votes in a single cell of the accumulator. Hence, some erroneous peaks may appear in the accumulator. Thereby while this method is effective for easing the computation and memory requirements, it may impair the accuracy.

## 2.2 Eliminating Erroneous Votes

Applying the reduced Hough transform to a profile image of the human head, results in an accumulator which usually presents a number of peaks. To detect the peak corresponding to the ear, we use various geometrical cues to eliminate some of the votes or even some of the voting pair points which do not comply with the properties of a prospective ear rim ellipse.

**Vertical Ellipses.** Given a non-rotated head profile image, the ear usually appears as a vertical ellipse with a vertical axis approximately twice as large as the horizontal axis. This can be described as,

$$\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} = 1, \quad b = 2a \quad (1)$$

where equation (1) denotes a non-rotated ellipse where  $a$  and  $b$  are the ellipse size along each axis and  $(x_0, y_0)$  is the centre.

By differentiation,

$$\frac{\partial y}{\partial x} = -\frac{(x - x_0)b^2}{(y - y_0)a^2}, \quad \frac{b^2}{a^2} = 4. \quad (2)$$

The edge direction can be arranged to equal  $\frac{\partial y}{\partial x}$  and thereby a vote cast for a putative ellipse centre  $(x_0, y_0)$  by a point  $(x, y)$  is allowed in the accumulator if

$$\frac{\partial y}{\partial x} + \frac{(x - x_0)b^2}{(y - y_0)a^2} = t, \quad \frac{b^2}{a^2} = 4 \quad (3)$$

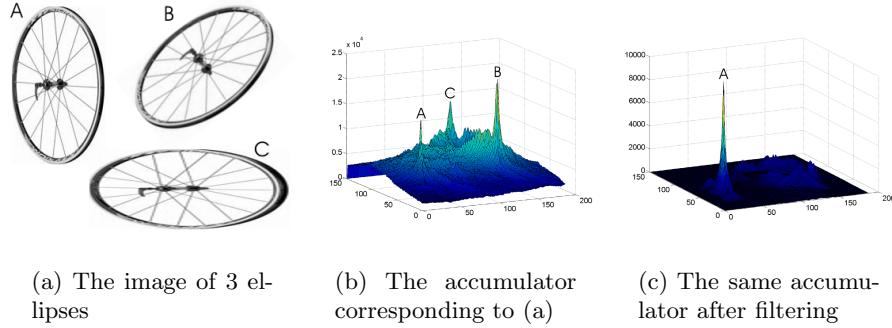
$$t \leq \text{threshold}.$$

The filter is set to ignore discrepancies less than a threshold value since: (i) some ears are slightly rotated, (ii) the exact value of  $\frac{b^2}{a^2}$  for each ear is not known, (iii) edge direction information cannot be accurately calculated.

Figure 2(a) shows three similar ellipses rotated and sized differently. The accumulator is depicted in figure 2(b) presenting three peaks corresponding to these three ellipses. Figure 2(c) shows the effects of filtering on this accumulator removing the peaks corresponding to non-vertical ellipses. Note that the noise level is also reduced as a result of this filtering.

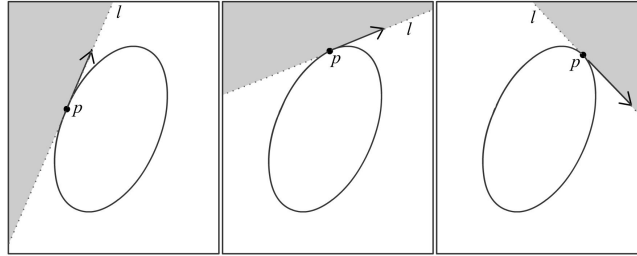
**Selective Pairing.** Given that a pair of edge pixels are within a size window which controls the size of the target ellipse, they will vote for putative ellipse centres along a specific line (see figure 1). Aiming to be more selective, we use the edge direction information to determine whether or not a pair of points can possibly be part of an ellipse. Choosing the pairs more carefully reduces the noise level in the accumulator and the peak corresponding to the ear becomes more pronounced.

The line  $l$  which passes through the edge point  $p$  with the same direction as the edge direction at  $p$  splits the image plane into two regions;  $l$  is tangent to any



**Fig. 2.** Removing the non-vertical ellipses. (a) An image of 3 ellipses. At the top left of this image a vertical ellipse with  $b = 2a$  is shown. At the top right the same ellipse with about  $45^\circ$  of rotation is depicted and on the bottom right there is a horizontal ellipse with  $a = 3b$ . (b) Three peaks are presented corresponding to the three ellipses. (c) After filtering only the peak corresponding to the vertical ellipse with  $b = 2a$  remains.

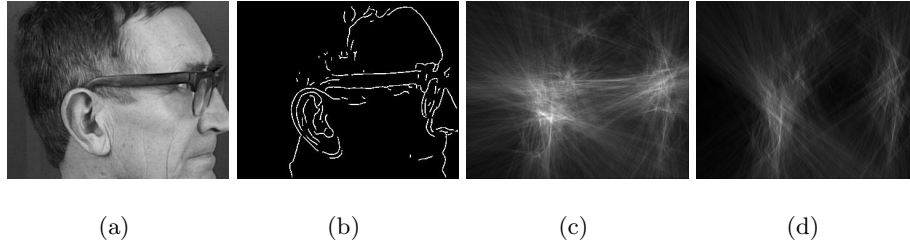
putative ellipse containing  $p$ . Examples of this are depicted in figure 3, where the gray regions are regions which do not contain ellipse points. Thus, only the points in the white regions can be considered as the pairing counterparts of the point  $p$ .



**Fig. 3.** Recognizing the ellipse free regions (the gray regions), to improve selectivity in edge points pairing

**Removing the Effects of Spectacles.** *Frames:* The lenses of spectacles may appear as ellipses when viewed from the side, this challenges the presumption that the ear being the most prominent ellipse shape in the head profile images. At times, the impairing effect of these ellipses is significant enough to lead the enrolment process to miss-localize. Accordingly,  $\frac{1}{3}$  of the front of the faces is cropped to eliminate these elliptical shapes.

*Handles:* Inspecting example accumulators generated by the reduced HT for ellipses, shows that the dominant orientation of the accumulators' vote lines in a region mimics the dominant edge direction of the corresponding region in the images. For instance the vote lines cast by ear rim edges are mostly vertical.



**Fig. 4.** The destructive effect of spectacle handles. The accumulator for image (a) is in (c) and it is filtered to eliminate the horizontal vote lines (d).

Given a region with numerous edge points and similar edge directions many pairs are recognized which vote in a limited space with similar vote line orientations. This leads to a ridge of votes, and possibly a peak. The case of spectacle handles brings about a similar scenario; these handles usually pose many points in the edge detected map (see figure 4(b)) which all have rather a horizontal edge direction. Figure 4 shows such an edge map and the corresponding accumulator. To remove the destructive effects of this type, bearing in mind that the ear ellipses are mostly vertical thus their dominant vote lines are also vertical, we cut out the vote lines which are close to being horizontal. Note that our datasets do not contain significantly rotated images, consistent with a subject looking forward.

### 2.3 Averaging

After these pruning stages, the ear centre is usually presented by a wide peak which has less height than expected, due to the fact that ear is not a perfect ellipse and the edge direction information is not as accurate as desired. Occasionally, single peaks also appear in the accumulator, which are mostly caused by curls in the hair. To remove the single peaks and find the centre of the widest peak which represents the ear, an  $30 \times 30$  averaging template is applied to the accumulator array.

## 3 Results

We initially developed our enrolment algorithm for a database of 2D images selected from the XM2VTS [7] face-profile database. Our database consists of 252 images from 63 individuals who are those whose ear is not obscured by hair; 4 images per individual taken in 4 different sessions over a period of five months ensure natural variation between the images of the same subject. We also used our algorithm to enrol the ears from the UND ear biometric database [8], which consists of 942 image pairs from 302 subjects. Each image pair is one 3D scan and one colour image of the profile of the human head which are obtained nearly simultaneously. We only use the 2D images for enrolment. Example images from both databases are shown in figure 5. The XM2VTS subjects are imaged against



**Fig. 5.** Top row: examples of XM2VTS images. Bottom row: examples of UND images.



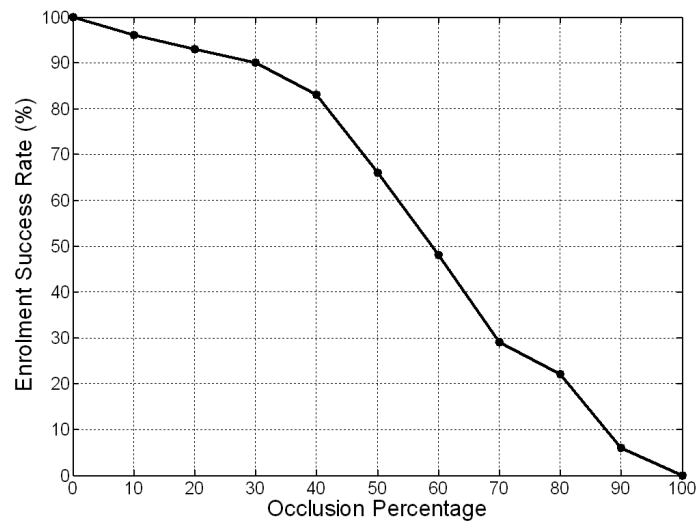
**Fig. 6.** The complete set of 252 enrolled ears

a blue background were the head is easily detected; diffused lighting was used to illuminate the head. On the other hand the UND images have a more natural indoor background and illumination. In this section we will present the results achieved on both databases, and we shall also evaluate the capability of our enrolment technique for partially occluded ears.

Figure 6 shows the enrolled ears of the XM2VTS derived, 252-image database. As shown in this figure, the enrolment process succeeds to detect the ear region



**Fig. 7.** An increasingly occluded ear enrolled successfully. The detected ellipse centres are shown by the X marks.



**Fig. 8.** Enrolment of occluded ears

in all the 252 profile images. A  $150 \times 120$  pixels image patch containing the ear is then cropped. We then use our algorithm to enrol the UND data. Given that the UND images include much more background information, we use both skin detection in colour images and the Canny edge operator to detect the face region of the images. The method proposed by Hsu et al. [10] was used for skin detection. Enrolment as was described in section 2 was then performed on the face region of the images. For UND data our algorithm offers a 91% enrolment success rate. The UND database does not offer as good quality for 2D images as presented in XM2VTS partly because of the indoor lighting conditions and partly because of using a rotating filter for capturing the 2D colour images which suffers from the slight movements of the subject. Thereby the edge points of the ear rim are less accurately detected which has an impairing effect on an evidence gathering algorithm such as HT which feeds on edge points.



### 3.1 Occlusion Test

The Hough Transform works with the edge points which vote in an accumulator. Given that adequate voting points are visible the structure will be detected in spite of occlusion of some of these points. For this test we synthetically occlude the ear region of the face-profile images in the XM2VTS derived database. Figure 7 shows that although the ear was increasingly occluded from the top it has still been enrolled correctly.

Figure 8 depicts the enrolment success rate of our 252-image database against increasing synthetic occlusion similar to the images in figure 7. As shown in this figure, 83% of the ears are enrolled successfully despite of 40% occlusion from the top, and at 50% occlusion 66% of the ears are still being correctly enrolled. Our algorithm shows promising results when used to enrol partially occluded ears.

## 4 Conclusions

The elliptical shape of the ear in profile images of the human head makes it possible to detect the ear by a HT, which was joined with refinements and additional constraints imposed by skin colour, rotation and relative sizes of the ear ellipse. The resulting method achieves error-free enrolment on our 252-image database. Using a reduced HT the method is not constrained by the high computation and memory requirements of the 5D HT for ellipses and given the properties of the HT in handling noise and occlusion it obtains satisfactory results enrolling occluded ears. Our enrolment method can be used to enrol ears from similar head-profile images. This generalization property has been demonstrated through the use of the method for enrolling ears of UND dataset.

In our other work we have introduced a new model-based ear recognition [11] in which the ear is recognised from the parts selected via a model. Combining our enrolment with the model we have built an automatic ear biometric system in which the objective of handling occlusion is carried through in enrolment via the HT and in feature extraction via the use of our new model which is contrary to holistic methods that describe images by their overall appearance.

## References

1. Iannarelli, A.: Ear Identification. Paramount Publishing Company, Freemont, California (1989)
2. Hurley, D.J., Arbab-Zavar, B., Nixon, M.S.: The ear as a biometric. In: Jain, A., Flynn, P., Ross, A., (eds.) Handbook of Biometrics (Forthcoming, 2007)
3. Choras, M.: Image feature extraction methods for ear biometrics—a survey. In: 6th International Conference on Computer Information Systems and Industrial Management Applications (CISIM 2007), Poland, pp. 261–265 (2007)
4. Yan, P., Bowyer, K.W.: Biometric recognition using 3d ear shape. IEEE Transactions on Pattern Analysis and Machine Intelligence 29, 1297–1308 (2007)

5. Chen, H., Bhanu, B.: Human ear recognition in 3d. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 718–737 (2007)
6. Alvarez, L., Gonzalez, E., Mazorra, L.: Fitting ear contour using an ovoid model. In: *Proc. of 39 IEEE International Carnahan Conference on Security Technology*, pp. 145–148 (2005)
7. Messer, K., Matas, J., Kittler, J., Luettin, J., Maitre, G.: Xm2vtsdb: The extended m2vts database. In: *Proc. AVBPA, Washington D.C* (1999)
8. Yan, P., Bowyer, K.W.: Empirical evaluation of advanced ear biometrics. In: *IEEE Computer Society Workshop on Empirical Evaluation Methods in Computer Vision*, San Diego (2005)
9. Aguado, A.S., Montiel, E., Nixon, M.S.: On using directional information for parameter space decomposition in ellipse detection. *Pattern Recognition* 29, 369–381 (1996)
10. Hsu, R., Abdel-Mottaleb, M., Jain, A.: Face detection in color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 696–706 (2002)
11. Arbab-Zavar, B., Nixon, M.S., Hurley, D.J.: On model-based analysis of ear biometrics. In: *IEEE Conference on Biometrics: Theory, Applications and Systems (BTAS 2007)*, Washington DC (to appear, 2007)