The Ear as a Biometric

D. J. Hurley[1] B. Arbab-Zavar[2] and M. S. Nixon[3]

[1] University of Southampton `djh@analyticalengines.co.uk`
[2] University of Southampton `baz05r@ecs.soton.ac.uk`
[3] University of Southampton `msn@ecs.soton.ac.uk`

1 Introduction

The potential of the human ear for personal identification was recognized and advocated as long ago as 1890 by the French criminologist Alphonse Bertillon. In his seminal work on biometrics he writes [7],

> "The ear, thanks to these multiple small valleys and hills which furrow across it, is the most significant factor from the point of view of identification. Immutable in its form since birth, resistant to the influences of environment and education, this organ remains, during the entire life, like the intangible legacy of heredity and of the intra-uterine life".

Ear biometrics has received scant attention compared to the more popular techniques of automatic face, eye, or fingerprint recognition. However, ears have played a significant role in forensic science for many years, especially in the United States, where an ear classification system based on manual measurements was developed by Iannarelli, and has been in use for more than 40 years [25], although the safety of ear-print evidence has recently been challenged [28, 14]. Rutty et al. have considered how Iannarelli's manual techniques might be automated [34] and a European initiative has looked at the value of ear prints in forensics [17].

Ears have certain advantages over the more established biometrics; as Bertillon pointed out, they have a rich and stable structure that changes little with age. The ear does not suffer from changes in facial expression, and is firmly fixed in the middle of the side of the head so that the immediate background is predictable, whereas face recognition usually requires the face to be captured against a controlled background. Collection does not have an associated hygiene issue, as may be the case with contact biometrics, and is unlikely to cause anxiety as may happen with iris and retina measurements. The ear is large compared with the iris, retina, and fingerprint and therefore is more easily captured at a distance.

Burge et al. [5, 6] were amongst the first to describe the ear's potential as a biometric using graph matching techniques on a Voroni diagram of curves extracted from the Canny edge map. Moreno et al. [30] tackled the problem with some success using neural networks and reported a recognition rate of 93% using a two-stage neural network technique. Hurley et al. used force field feature extraction [18, 22, 23] to map the ear to an energy field which highlights "potential wells" and "potential channels" as features. By achieving a recognition rate of 99.2%, [23] this method proved to yield a much better performance than PCA when the images were poorly registered. The approach is also robust to noise; adding 19dB of Gaussian noise actually improved the performance to 99.6% [24]. Abdel-Mottaleb et al. [1] used the force field transform to obtain a smooth surface representation for the ear and then applied different surface curvature extractors to gather the features.

Statistical holistic analysis, especially Principal Components Analysis (PCA), has proved to be one of the most popular approaches to ear recognition. Victor et al. [40] applied PCA to both face and ear recognition and concluded that the face yields a better performance than the ear. However, Chang et al. [8] conducted a similar experiment and reached a different conclusion: no significant difference was observed between face and ear biometrics when using PCA. The image dataset in [40] had less control over earrings, hair, lighting etc. and as suggested by Chang et al., this may account for the discrepancy between the two experiments. Chang et al. also reported a recognition rate of 90.9% using a multimodal approach. Zhang et al. [48] developed a system combining Independent Components Analysis (ICA) with a Radial Basis Function (RBF) network and showed that better performance can be achieved using ICA instead of PCA. However being pure statistical measures, both PCA and ICA offer almost no invariance and therefore require very accurate registration in order to achieve consistently good results.

Yuizono et al. [47] treated the recognition task as an optimisation problem, proposing a system using a specially developed genetic local search targeting the ear images. Given that their work does not include any feature extraction process, it has no invariant properties. Some studies have focused on geometrical approaches [31, 13]; Mu et al. [31] reported an 85% recognition rate using such an approach. Alvarez et al. [3] proposed and intend to implement an ovoid model for segmentation and normalization of the ear.

Yan et al. [45, 43] captured 3D ear images using a range scanner and used Iterative Closest Point (ICP) registration for recognition to achieve a 97.8% recognition rate. Chen et al. proposed a 3D ear detection and recognition system using a model ear for detection, and using ICP and a local surface descriptor for recognition, reporting a recognition rate of 90.4% [9, 12, 10, 11].

A number of multimodal approaches to ear recognition have also been considered [8, 42, 26, 35]. Iwano et al. [26] combined ear images and speech using a composite posterior probability, and showed that the performance improves using ear images in addition to speech in the presence of noise. In this study, PCA was applied to extract the ear features. Chang et al. [8] and

Rahman et al. [35] proposed multimodal biometric systems using PCA on both face and ear. Both studies reported an increase in performance when using multimodal biometrics instead of individual biometrics, achieving multi-modal recognition rates of 90.9% and 94.4% respectively. Yan et al. [42] conducted multi-modal experiments to test the efficacy of various combinations of 2D-PCA, 3D-PCA, and 3D-Edges with the recognition results shown in Table 1. For further details of multi-modal ear and face biometrics see the chapter by Bowyer. An introductory survey of ear biometrics has been provided by Pun

Table 1. Yan et al. Multi-Modal Recognition Results

2d-pca,	3d-pca,	3d-edge,	3d-pca+3d-edge,	2d-pca+3d-edge,	2d-pca+3d-pca,	all 3
71.9%	64.8%	71.9%	80.2%	89.7%	89.1%	90.6%

et al. [33].

In related studies Akkermans et al [2] developed an ear biomeric system based on the acoustic properties of the ear. They measure the acoustic transfer function of the ear by projecting a sound wave at the ear and observing the change in the reflected signal. Scandia Corp. patented a similar technique [37].

We will start this chapter with a review of the anatomy and physiology of the ear and how this is likely to affect its biometric properties. The ear biometrics field is still so small that we will be able to touch on most of the main techniques. In particular, we will describe PCA in some detail as this has proved to be one of the most popular techniques. Despite its intricate mathematical nature, it is quite easy to implement and even easier to use, and should allow the reader to do some simple experiments with ear biometrics in order to confirm their biometric potential. Finally, we will consider the future of ear biometrics and related issues such as 2D and 3D ear databases.

2 Evidence and Support for Ears as a Biometric

The structure of the ear is not quite so random as Bertillon seems to suggest; it has a definite structure just like the face. Most people when asked could easily draw the outline of the ear but only the experienced artist would be able to reproduce from memory its detailed intricate structure. As shown in Figure 1, the shape of the ear tends to be dominated by the outer rim or helix, and also by the shape of the lobe. There is also an inner helix or antihelix which runs roughly parallel to the outer helix but forks into two branches at the upper extremity. The inner helix and the lower of these two branches forms the top and left side of the concha, named for its shell-like appearance. The bottom of the concha merges into the very distinctive intertragic notch, which

due to its very sharp bend at the bottom can form a useful reference point for biometrics purposes. Note also the crus of helix where the helix intersects with the lower branch of the antihelix. This is one of the points used by Iannarelli as a reference point for his measurement system, the other point being the antitragus or the little bump on the left of the intertragic notch [25]. The front of the concha opens into the external ear canal or acoustic or auditory meatus, more commonly referred to as the ear hole, although this is usually somewhat concealed by the flesh around and above the tragus. It is interesting to note [32] that the embryonic ear has a small number of about 6 individual growth nodules which eventually develop along with the foetus to become the fully formed auricle in the newborn infant, striking a note with Bertillon's earlier observation.
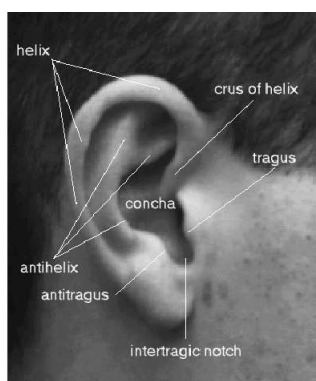

Fig. 1. Anatomy of the ear. In addition to the familiar rim or helix and ear lobe, the ear also has other prominent features such as the anti-helix which runs parallel to the helix, and a distinctive hairpin-bend shape just above the lobe called the intertragic notch. The central area or concha is named for its shell-like appearance.

Figure 2 shows a small sample of human ears indicating the rich variety of different shapes. Notice that some ears have well formed lobes, whereas others have almost none. These latter are called "attached lobes" and make measurement of the length of the ear difficult.

Because of the tendency of the inner and outer helices to run parallel, there is quite a degree of correlation between them which detracts somewhat from the biometric value of the ear; indeed it could also be argued that the concha is simply the space that remains when the other parts have been accounted for, so that it is also highly correlated to its neighbouring parts and therefore contributes less independent information than might appear to be the case at first.

The outer ear called the auricula or pinna forms only part of the total ear organ which has evolved to locate, collect, and process sound waves. Many other mammals like horses, dogs, and cats can articulate their ears to better

Fig. 2. Examples of the human ear shape. Notice that helices, concha, intertragic notch, etc. are present in all the examples, but that some ears have so called attached lobes, where the lobes are poorly formed or are almost non-existent.

locate particular sound sources. Fortunately for the purpose of biometrics we humans can hardly articulate our ears; our ears are held rigidly in position by cartilaginous tissue which is firmly attached the bone at the side of the head. The ear owes its semi-rigid shape due this stiff tissue which underlies its soft flesh.

The face has roughly the same visual complexity as the ear. Quite simple changes in the parameters which define the size and shape of the eyes, nose, mouth, and cheek-bones can lead to a wide range of facial appearances. In this we regard perfect symmetry as a mark of beauty, but we should note that the ear lacks all symmetry. It is also worth noting that since the face is symmetrical about its centre-line, therefore its structure really only represents half-a-face from a biometrics perspective because the information on the left side reflects that on the right. The ear has no symmetry and therefore does not suffer from this drawback giving it an advantage over the face, and of course the face is contorted during speech and when expressing emotions, and its appearance is often altered by make-up, spectacles, and beards and moustaches, whereas the ear does not move and only has to support earrings, spectacle frames, and sometimes hearing aids, although of course it is often occluded by hair. As such, the ear is much less susceptible to covariate interference than many other biometrics, with particular invariance to age.

3 Approaches to Ear Biometrics

3.1 The early work of Iannarelli and Forensic Ears

Alfred Iannarelli developed a system of ear classification used by American law enforcement agencies. In late 1949 he became interested in the ear as a means of personal identification in the context of forensic science. He subsequently developed the Iannarelli System of Ear Identification [25]. As shown in Figure 3 his system essentially consists of taking a number measurements around the
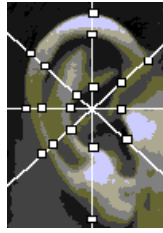
Fig. 3. Iannarelli's manual ear measurement system.

ear by placing a transparent compass with 8 spokes at equal 45 intervals over an enlarged photograph of the ear. The first part of registration is achieved by ensuring that a reference line touches the crus of helix at the top and touches the innermost point on the tragus at the bottom. Normalisation and the second step of registration are accomplished by adjusting the enlargement mechanism until a second reference line exactly spans the concha from top to bottom. Iannarelli has appeared personally as an expert witness in many court cases involving ear evidence, or is often cited as an ear identification expert by other expert witnesses [28]. In the preface to his book Iannarelli states,

> "Through 38 years of research and application in earology, the author has found that in literally thousands of ears that were examined by visual means, photographs, ear prints, and latent ear print impressions, no two ears were found to be identical - not even the ears of any one individual. This uniqueness held true in cases of identical and fraternal twins, triplets, and quadruplets"

When Iannarelli suggests that "not even the ears of any one individual are unique" he has unwittingly touched on the nub of the biometrics problem. It is not an advantage, as he seems to suggest, that the ear samples from the same individual are not unique. On the contrary the less these samples are unique, then the less are we entitled to claim that an individual's biometric is unique. If we think of individuals' samples as forming points in a feature space, then these points will form clusters for each individual. It is the extent to which these different clusters are separated from one and other and the extent to which the individual clusters are closely grouped around their own averages, that determines how good a particular biometric system performs. In recent times attempts have been made to automate Iannarelli's system [34].

3.2 Burge and Burger Proof of Concept

Burge and Burger [5, 6] were the first to investigate the human ear as a biometric in the context of machine vision. Inspired by the earlier work of Iannarelli [25], they conducted a proof of concept study where the viability of the ear as a biometric was shown both theoretically in terms of the uniqueness

and measurability over time, and in practice through the implementation of a computer vision based system. Each subject's ear was modeled as an adjacency graph built from the Voronoi diagram of its Canny extracted curve segments. They devised a novel graph matching algorithm for authentication which takes into account the erroneous curve segments which can occur in the ear image due to changes such as lighting, shadowing, and occlusion. They found that the features are robust and could be reliably extracted from a distance. Figure 4 shows the extracted curves, Voronoi diagram, and neighbour graph for a typical ear. They identified the problem of occlusion by hair as
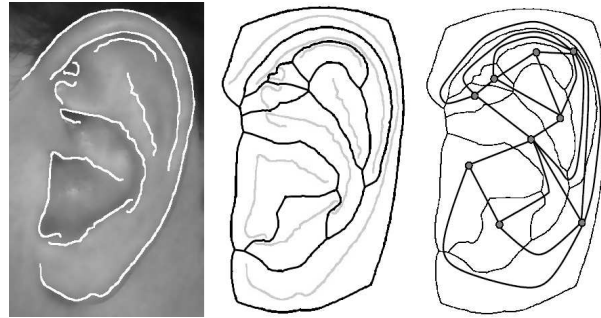

Fig. 4. Graph model: Stages in building the ear biometric graph model. A generalized Voronoi diagram (centre) of the Canny extracted edge curves (left) is built and a neighborhood graph (right) is extracted.

a major obstacle and proposed the use of thermal imagery to overcome this obstacle.

3.3 Principal Components Analysis

Principal Components Analysis, closely related to Singular Value Decomposition, has been one of the most popular approaches to ear recognition [40, 8, 23, 26, 41, 35]. It is an elegant, easy to implement and easy to use technique, so we will attempt to describe it in sufficient detail for the reader to be able to understand and implement it readily with a view to being able to set up a simple ear recognition experiment to confirm the basic biometric potential of the ear. The underlying mathematics can be found in [39, 27].

We will first show how images can be looked upon as vectors, and how any picture can be constructed as a summation of elementary picture-vectors. We will then show how PCA can process these vectors to achieve image compression, and how this in turn can be used for biometrics.

We are familiar with the real coordinate space R^3 where any point can be represented as a linear combination of 3 unit value basis vectors mutually at right angles to each other. For example, the point (3,4,5) can be expressed as,

$$3(1,0,0) + 4(0,1,0) + 5(0,0,1) = (3,0,0) + (0,4,0) + (0,0,5) = (3,4,5)$$

We could also express any point as the sum of non-standard basis vectors, providing that none of the chosen basis vectors is a linear combination of the other two. For example, we can also write,

$$(3,4,5) = 1.333(1,2,3) + 0.333(2,3,1) + 0.333(3,1,2)$$

Now if we admit the possibility of negative value pixels, then pictures can also be treated as vectors so that any picture can be expressed as a linear combination of unit value basis picture-vectors. For example, a trivial four element picture can be expressed as,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = 1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + 3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + 4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

In the example which follows taken from [23] we will be dealing with 111x73 pixel images. This would require 111x73 = 8103 sparse elementary picture-vectors, each with only one pixel set to 1 and the remaining pixels set to 0, and a set of 8103 weights to specify a particular picture, obviously not resulting in any compression advantage.

In this real example we use a subset of the XM2VTS face profiles database [29], consisting of 4 ear images for each of 63 subjects giving us a total of 252 images . Now here is how the "magic" of PCA works. By taking one of the four samples from each of the 63 subjects we produce a special *projection matrix* **P** which enables us to compute a set of 63 weights for each of the 252 images which when used to scale a set of 63 special picture-vectors already encoded in **P** produces a reasonable facsimile of the original image. Instead of requiring 8103 weights we can make do with only 63 which is a very high degree of compression of well over 100:1, albeit lossy compression. These weights form convenient 63 element feature vectors representing each picture and are perfect for biometric comparison as they allow us to calculate the Euclidian distance between pictures by doing a simple vector subtraction.

We will now give the details of the calculations involved. In order to carry out matrix multiplication of the 111x73 picture-vectors we first have to encode them as 8103x1 column vectors by stacking the 73 columns on top of each other. Any results can be recoded as rectangular matrices for display purposes.

The projection matrix is calculated as follows

Let **p** be any of the 63 first of four picture samples
Let **m** be the average over the 63 pictures i.e.$(\sum \mathbf{p})/63$
Let $\mathbf{d} = \mathbf{p} - \mathbf{m}$ be the difference between each picture and the average
Let **D** be the array formed by the 63 columns of difference pictures **d**
Then the projection matrix is given by,

$$\mathbf{P} = \mathbf{D}S(\mathbf{D^T D}) \tag{1}$$

where $S(\mathbf{M})$ is a function that returns a matrix whose columns are the normalised eigenvectors of matrix $\mathbf{M}$

The basis-pictures or *eigenvectors* are simply the columns of $\mathbf{P}$
The weights for picture $\mathbf{p}$ are given by

$$\mathbf{w} = \mathbf{d^T P} \tag{2}$$

The compressed image for a given picture $\mathbf{p}$ is given by

$$\mathbf{c} = \mathbf{P w^T} + \mathbf{m} \tag{3}$$

Figure 5 shows the first 36/63 eigenvctors, whereas Figure 6 shows the projections and eigenvector spectra for 3 subjects. Notice the that the leftmost projections are the best facsimiles because they been used in forming the projection matrix. Notice also that the eigenvector spectra, consisting of the 63 weights, do not rapidly diminish to zero, in fact all of these 63 weights are used for comparison. Each set of 63 weights is treated as a vector and the Euclidian distances between these vectors are used as a suitable metric,

$$distance = \|\mathbf{w}_i - \mathbf{w}_j\| \tag{4}$$

The means and standard deviations of the inter-class and intra-class distributions can then be calculated to gauge the efficacy of the technique. The spreads or standard deviations of the two distributions should be small compared to the separation of their means for a good biometric. It is customary to consider the 63 samples used in forming $\mathbf{P}$ as having been "sacrificed" and not to include them in the biometric comparison so that only $252 - 63 = 189$ ears would be used. In this experiment a recognition rate of 186/189 or 98.4% was achieved [23].

3.4 Force Field Transform

Hurley et al. [18, 20, 22, 23] have developed an invertible linear transform which transforms an ear image into a force field by pretending that pixels have a mutual attraction proportional to their intensities and inversely to the square of the distance between them rather like Newton's Universal Law of Gravitation. Underlying this force field there is an associated energy field which in the case of an ear takes the form of a smooth surface with a number of peaks joined by ridges as shown in Figure 8. The peaks correspond to potential energy wells and to extend the analogy the ridges correspond to potential energy channels. Since the transform also turns out to be invertible, all of the original information is preserved and since the otherwise smooth surface is modulated by these peaks and ridges, it is argued that much of the information is transferred to these features and that therefore they should make good features.
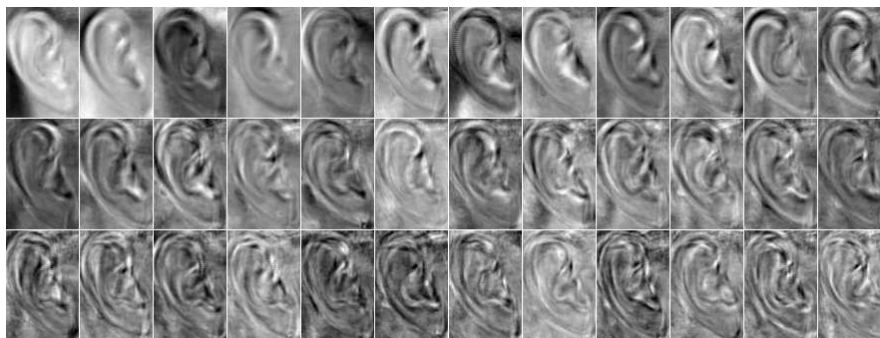
Fig. 5. The first 36 of the set of 63 eigenvectors for the subset of 63 ear images selected from the 252 image database. The first of the four samples from each of the 63 subjects was used in forming the projection matrix. These are the basis picture-vectors which will be scaled by the computed weights to produce the compressed or projected images.
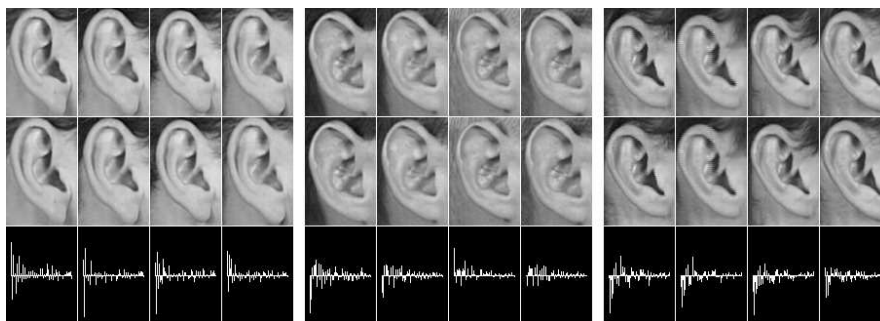

Fig. 6. PCA projections and eigenvector spectra for 3 subjects. The top rows show the original images whilst the middle rows are their corresponding projections into the eigenvector subspace. The bottom row depicts the eigenvector spectrum for each image consisting of the 63 weights used to render its projection.
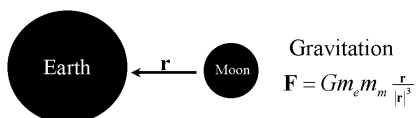

Fig. 7. Newton's Universal Law of Gravitation. The earth and moon are mutually attracted according to the product of their masses m_e and m_m respectively, and inversely proportional to the square of the distance between them. G is the gravitational constant of proportionality.

$$\mathbf{F}(\mathbf{r}_j) = \sum_i \left\{ P(\mathbf{r}_i) \frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3} \right\} \forall i \neq j, 0 \forall i = j \tag{5}$$

$$E(\mathbf{r}_j) = \sum_i \frac{P(\mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{r}_j|} \forall i \neq j, 0 \forall i = j \tag{6}$$

Two distinct methods of extracting these features are offered. The first

Fig. 8. Generating an ear energy surface by convolution. The energy field for an ear (right) is obtained by locating a unit value potential function (left) at each pixel location and scaling it by the value of the pixel and then finding the sum of all the resulting functions. For efficiency this is actually calculated in the frequency domain.

method depicted in Figure 9a is algorithmic, where test pixels seeded around the perimeter of the force field are allowed to follow the force direction joining together here and there to form channels which terminate in potential wells. The second method depicted in Figure 9b is analytical, and results from an analysis of the mechanism of the first method leading to a scalar function based on the divergence of the force direction. The second method was used to obtain a recognition rate of over 99% on a database of 252 ear images consisting of 4 time lapsed samples from each of 63 subjects, extracted from the XM2VTS face profiles database [29].

Equations 5 and 6 show how the force and energy fields are calculated at any point $\mathbf{r}_j$. These equations must be applied at every pixel position to generate the complete fields. In practice this computation would be done in the frequency domain using Equation 7 where $\Im$ stands for FFT.

$$Energy = \sqrt{MN} \left\{ \Im^{-1} \left[\Im\left(potential\right) \times \Im\left(image\right) \right] \right\} \tag{7}$$

Convergence provides a more general description of channels and wells in the form of a mathematical function in which wells and channels are revealed to be peaks and ridges respectively in the function value. This function maps the force field $\mathbf{F}(\mathbf{r})$ to a scalar field $C(\mathbf{r})$, taking the force as input, and returning the additive inverse of the divergence of the force direction, and is defined by,

$$C(\mathbf{r}) = -\mathrm{div}\mathbf{f}(\mathbf{r}) = -\lim_{\Delta A \to 0} \frac{\oint \mathbf{f}(\mathbf{r}) \cdot d\mathbf{l}}{\Delta A} = -\nabla \cdot \mathbf{f}(\mathbf{r}) = -\left(\frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} \right) \tag{8}$$

where $\mathbf{f}(\mathbf{r}) = \frac{\mathbf{F}(\mathbf{r})}{|\mathbf{F}(\mathbf{r})|}$ is the force direction, ΔA is incremental area, and $d\mathbf{l}$ is its boundary outward normal. This function is real valued and takes negative values as well as positive ones where negative values correspond to force direction divergence. Note that the function is non-linear because it is based on force direction and therefore must be calculated in the given order.
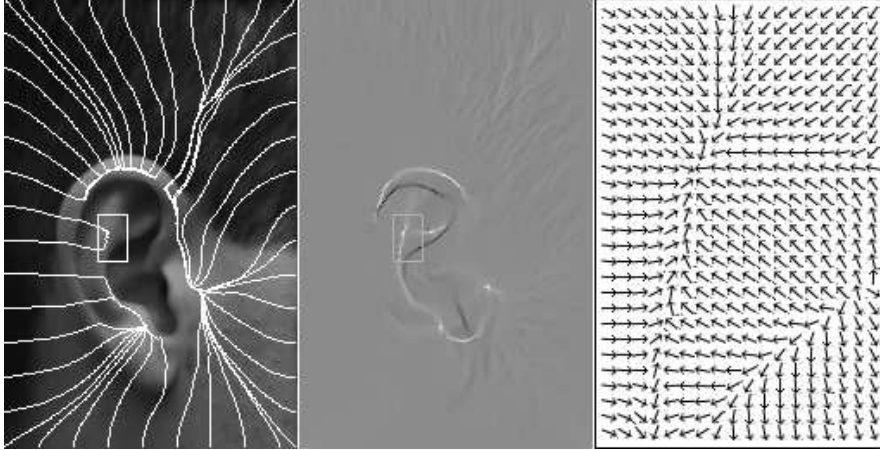

Fig. 9. Force and convergence fields for an ear. The force field for an ear (left) and its corresponding convergence field (centre). The force direction field (right) corresponds to the small rectangular inserts surrounding a potential well on the inner helix.

3.5 Three Dimensional Ear Biometrics

The auricle has a rich and deep three dimensional structure, so it is not surprising that a number of research groups have focused their attention in this direction.

Yan and Bowyer ICP Approach

Yan et al. [46, 42, 44, 45, 43] use a Minolta VIVID 910 range scanner to capture both depth and colour information. The device uses a laser to scan the ear, and depth is automatically calculated using triangulation. They have developed a fully automatic ear biometric system using ICP based 3D shape matching for recognition, and using both 2D appearance and 3D depth data for automatic ear extraction which not only extracts the ear image but also separates it from hair and earrings. They achieve almost 98% recognition on a time-lapse database of 1,386 images over 415 subjects, with an equal error rate of 1.2%. The 2D and 3D image datasets used in this work are available

to other research groups. For further details see the chapter by Flynn in the appendix.

Ear extraction 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.

Ear pit detection includes: (i) geometric prepossessing to locate the nose tip to act as the hub of a sector which includes the ear with a high degree of confidence; (ii) skin detection to isolate the face and ear region from the hair and clothes; (iii) surface curvature estimation to detect the pit regions depicted in black in the image; (iv) surface segmentation and classification, and curvature information to select amongst possible multiple pit regions using a voting scheme to select the most likely candidate. The detected ear pit is then used to initialize an active contour algorithm to find the ear outlines. Both 2D colour and 3D depth are used to drive the contour, as using either alone is inadequate since there are cases in which there is no clear colour or depth change around the ear contour.
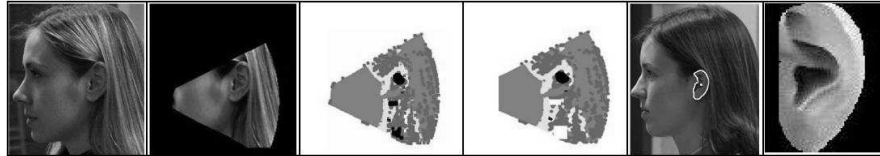

Fig. 10. 3D ear extraction. From left to right, skin detection and most likely sector generation, pit detection and selection, ear outline location, 3D ear extraction
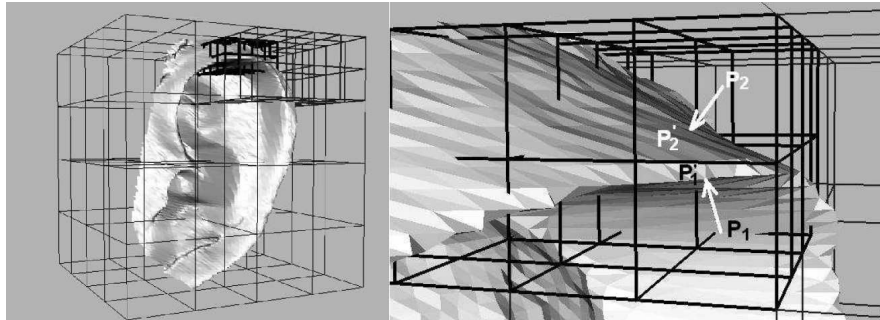

Fig. 11. Voxelization: Left: 3D Image space is partitioned into voxels. Right: Two voxel centres P_1 and P_2 and their closest points on the gallery surface P_1' and P_2'.

3D shape matching: ICP [4] has been widely used for 3D shape matching due to its simplicity and accuracy, however it is computationally expensive.

Given a source point set P and a model point set M, ICP iteratively calculates the rigid transform T that best aligns P and M. At the ith iteration, the transform T_i is the transform that minimizes the mean square differences between the corresponding points of P_i and M. The corresponding points are the closest points between the two point-sets. P_i is then updated using T_i.

Yan et al. [46] have developed an efficient ICP registration method called "Pre-computed Voxel Closest Neighbours" which exploits the fact that subjects have to be enrolled beforehand for biometrics. Since the most time consuming part of the ICP algorithm is finding the closest points between the probe and the gallery (of order $Np * logNm$) the main idea of this method is to approximate each point of the probe with a nearby point whose nearest point in the gallery point set is pre-computed. They proposed a quantised 3D volume using voxels, as shown in Figure 11. Placing the 3D probe image into this volume, each point of the probe falls into a voxel. Each probe point is then approximated by the voxel centre wherein it is placed. For each voxel the closest point in 3D space on the gallery surface is computed ahead of time. Figure 11 shows the closest points to the two voxel centres P_1 and P_2.

Chen and Bhanu Local Surface Patch Approach

Chen et al.[9, 12, 10, 11] have also tackled 3D ear biometrics using a Minolta range scanner as the basis of a complete 3D recognition system on a database of 52 subjects consisting of two images per subject. The ears are detected using template matching of edge clusters against an ear model based on the helix and antihelix, and then a number of feature points are extracted based on local surface shape. A signature called a "Local Surface Patch" based on local curvature is computed for each feature point and is used in combination with ICP to achieve a recognition rate of 90.4%

Feature points extraction Shape index S_i is a quantitative measure of surface shape [16] based on principal curvatures which classifies surface shape as one of 9 basic types represented by values in the interval [0,1].

$$S_i\left(p\right) = \frac{1}{2} - \frac{1}{\pi}\tan^{-1}\frac{k_1\left(p\right) + k_2\left(p\right)}{k_1\left(p\right) - k_2\left(p\right)} \tag{9}$$

where k_1 and k_2 are the maximum and minimum principal curvatures respectively. Chen et al. then choose as feature points those where the index is locally maximum or minimum.

Local Surface Patch A local surface patch (LSP) [9] comprises the neighbourhood of points N around a feature point P which are close enough to the feature point in Euclidean distance and surface normal.

$$N = \{N_i : N_i \ pixel, \|N_i - P\| \leq \varepsilon_1, a\cos(n_p \bullet n_{n_i}) < A\} \tag{10}$$

For each feature point, shape index values of its LSP points and the dot product of surface normal vectors of the feature point and its LSP points are

computed, and accumulated in a 2D histogram. The 2D histogram accumulates this information in bins along two axes. These two axes are the shape index with range [0,1] and the dot product of surface normal vectors which is in the range [-1,1]. A surface type of "concave", "convex", or "saddle" is also allocated to each LSP. Taken together the 2D histogram, the surface type and the centroid of the local surface patch make up a distinctive signature for each patch.
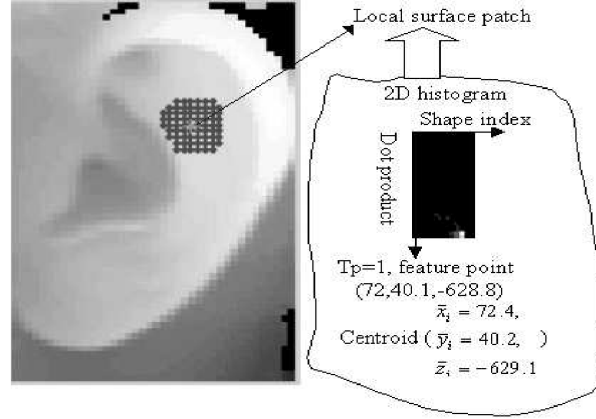

Fig. 12. Local Surface Patch. The LSP constitutes a characteristic signature consisting of a 2D histogram, a surface type, and a centroid.

Recognition This is a two stage process based on LSP for coarse alignment and ICP for fine alignment of probe and gallery images. Probe images are compared against all images in the gallery; each comparison is started by identifying the best match for each probe LSP in the gallery. Assuming that the true set of matches which pairs the patches that depict similar features in both probe and gallery is a subset of the total matches, a geometric constraint is applied to divide the matches into groups where each pair of matches in a group must satisfy the following condition,

$$d_{C_1,C_2} = |d_{P_1,P_2} - d_{G_1,G_2}| < \varepsilon_2 \qquad (11)$$

where $C_1 = \{P_1, G_1\}$ and $C_2 = \{P_2, G_2\}$ are the matches for probe and gallery patches P and G respectively, and d_{P_1,P_2} and d_{G_1,G_2} are the Euclidean distances between patch centroids. The above constraint guarantees that a group of matches preserves the mutual position of the patches. In other words d_{P_1,P_2} should be consistent with d_{G_1,G_2}. Note that with this definition a match can be placed in more than one group. The biggest group is then declared as the true match subset.

Starting with an initial rigid transform based on the true match subset, ICP is applied to find the refined alignment between the probe and the gallery image. Having compared all the gallery images to the probe, the gallery image with least root mean square (RMS) error is classified as the correct match.

3.6 Acoustic Ear Recognition

Akkermans et al. [2] have exploited the acoustic properties of the ear for recognition. It turns out that the ear by virtue of its special shape behaves like a filter so that a sound signal played into the ear is returned in a modified form. This acoustic transfer function forms the basis of the acoustic ear signature. An obvious commercial use is that a small microphone might be incorporated into the earpiece of a mobile phone to receive the reflected sound signal and the existing loudspeaker could be used to generate the test signal.
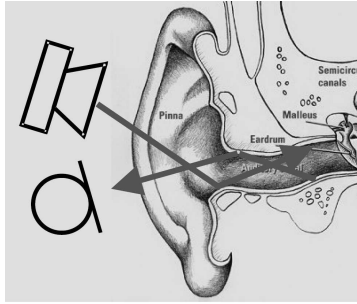

Fig. 13. An ear signature is generated by probing the ear with a sound signal which is reflected and picked up by a small microphone. The shape of the pinna and the ear canal determine the acoustic transfer function which forms the basis of the signature.

Akkermans et al. measure the impulse response of the ear by sending a noise signal $n(t)$ with a spectrum $N(\omega)$ into the pinna and ear canal and measuring the response $r(t)$. Next, the response is transformed into the frequency domain by using an FFT to calculate the output frequency spectrum $R(\omega)$. Finally, an estimate is obtained of the transfer function $H(\omega) = R(\omega)/N(\omega)$ where $H(\omega)$ is the cascade of the transfer functions of the loudspeaker, pinna and ear canal, and microphone as shown in Figure 14.

The test database consists of 8 ear signatures collected from each of 31 subjects using headphones and a separate set of 8 signatures from 17 subjects using a modified mobile phone with a small microphone incorporated into the earpiece. The correlation metric,

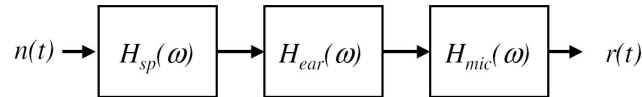$$C = \frac{\mathbf{x}.\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|} \tag{12}$$

$$n(t) \rightarrow \boxed{H_{sp}(\omega)} \rightarrow \boxed{H_{ear}(\omega)} \rightarrow \boxed{H_{mic}(\omega)} \rightarrow r(t)$$

Fig. 14. Calculating the impulse response of the ear

was used for comparison where **x** and **y** are the feature vectors taken relative to the mean of the population. Using Fisher LDA analysis equal error rates of 1.5% - 7% were obtained depending on whether headphones were used or mobile phones.

4 Conclusions and Outlook

The ear as a biometric is no longer in its infancy and it has shown encouraging progress so far - which is improving, especially with the interest created by the recent research into its 3D potential. It enjoys forensics support, it's structure appears individual, and it appears to have less variance with age than other biometrics.

It is also most unusual, even unique, in that it supports not only visual recognition but also acoustic recognition at the same time. This, together with its deep 3-dimensional structure will make it very difficult to fake thus ensuring that the ear will occupy a special place in situations requiring a high degree of protection against impersonation.

The all important question of "just how good is the ear as a biometric" has only begun to be answered. The initial test results, even with quite small datasets, were disappointing, but now we have regular reports of recognition rates in the high 90's on more sizeable datasets. But there is clearly a need for much better intra-class testing, both in terms of the number of samples per subject and of variability over time. However we will not dwell on this topic as it is treated in depth in the chapter in the the appendix on databases by Flynn.

Most of the recent work has focused on the overall appearance or on the shape of the ear, whether it be PCA, force field, or ICP, but it may prove profitable to further investigate if different and particular parts of the ear are more important than others from a recognition perspective. There is also a need to develop techniques with better invariance perhaps more model based, and to seek out high speed recognition techniques to cope with the very large datasets that are likely to be encountered in practice.

We must not forget that the inherent disadvantage of the occlusion of the ear by hair will always be a problem, but even this might be ameliorated by the development of thermal imaging schemes. But one thing is for certain, and

that is that there are many questions to be answered, so we can look forward to many interesting papers addressing these issues.

References

1. M. Abdel-Mottaleb, J. Zhou, Human Ear Recognition from Face Profile Images, *ICB 2006,* pp. 786 - 792.
2. A. H. M. Akkermans, T. A. M. Kevenaar, D. W. E. Schobben, Acoustic Ear Recognition for Person Identification, *Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05)* pp. 219-223
3. L. Alvarez, E. Gonzalez, L. Mazorra, Fitting ear contour using an ovoid model, *Proc. of 39 IEEE International Carnahan Conference on Security Technology,* 2005, pp. 145- 148.
4. Paul J. Besl, Neil D. McKay, A method for registration of 3-D shapes, *IEEE Trans. Pattern Anal. Machine Intell.,* pp. 239-256, 1992.
5. M. Burge, W. Burger, *Ear biometrics in: Jain, Bolle and Pankanti (Eds.), Biometrics: Personal Identification in Networked Society,* Kluwer Academic, Dordrecht, 1998, pp. 273-286.
6. Burge, M., and Burger, W., Ear biometrics in computer vision, *Proc. ICPR 2000,* pp. 822-826, 2002
7. A. Bertillon, *La photographie judiciaire, avec un appendice sur la classification et l'identification anthropometriques,* Gauthier-Villars, Paris, 1890.
8. K. Chang, K.W. Bowyer, S. Sarkar, 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. H. Chen, B. Bhanu, R. Wang, Performance evaluation and prediction for 3-D ear recognition, *Proc. International Conference on Audio and Video based Biometric Person Authentication,* NY, 2005.
10. H. Chen, B. Bhanu, Contour matching for 3-D ear recognition, *Proc. IEEE Workshop on Applications of Computer Vision,* Colorado, 2005.
11. B. Bhanu, H. Chen, Human ear recognition in 3-D, *Proc. Workshop on Multimodal User Authentication,* Santa Barbara, CA, 2003, pp. 91-98.
12. H. Chen and B. Bhanu, Shape Model-based ear detection from side face range images, *Proc. of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops,* 2005, vol. 3, p. 122.
13. M. Choras, Ear Biometrics Based on Geometrical Feature Extraction, *Electronic Letters on Computer Vision and Image Analysis (Journal ELCVIA),* 2005, vol. 5, no. 3, pp. 84-95.
14. www.timesonline.co.uk/article/0,1-973291,00.html *Man convicted of murder by earprint is freed,* January 22, 2004
15. Daubert v. Merrell Dow Pharmaceuticals (92-102), 509 U.S. 579 (1993).
16. C. Dorai and A. Jain, COSMOS-A representation scheme for free-form surfaces, *Proc. IEEE Conf. Computer Vision,* 1995, pp. 1024-1029.
17. L. Meijermana, S. Shollb, F. De Contic, M. Giaconc,C. van der Lugtd, A. Drusinic, P. Vanezis, G. Maata, Exploratory study on classification and individualisation of earprints, *Forensic Science International* 140 (2004) 91-99
18. Hurley, D. J., Nixon, M. S. and Carter, J. N. Force Field Energy Functionals for Image Feature Extraction. *Proc. 10th British Machine Vision Conference,* 1999, pp. 604-613

19. D. J. Hurley, M. S. Nixon, J. N. Carter, A New Force Field Transform for Ear and Face Recognition. *In Proceedings of the IEEE International Conference on Image Processing ICIP2000,*, 2000, pp. 25-28.
20. D. J. Hurley, M. S. Nixon, J. N. Carter, Force Field Energy Functionals for Image Feature Extraction, *Image and Vision Computing,* Special Issue on BMVC 99, 2002, vol. 20, No.5-6, pp. 311-317
21. D. J. Hurley, M. S. Nixon, J. N. Carter, Automatic Ear Recognition by Force Field Transformations. *In Proceedings of IEE Colloquium: Visual Biometrics* (00/018), 8/1-8/5.
22. D. J. Hurley, Force Field Feature Extraction for Ear Biometrics. *PhD Thesis 2001,* Electronics and Computer Science, University of Southampton.
23. D. J. Hurley, M. S. Nixon, J. N. Carter, Force field feature extraction for ear biometrics, *Computer Vision and Image Understanding,* 2005, vol. 98, pp. 491-512.
24. D. J. Hurley, M. S. Nixon, J. N. Carter, Ear Biometrics by Force Field Convergence, *Proc. AVBPA 2005,* pp. 386-394
25. A. Iannarelli, Ear Identification, *Paramount Publishing Company,* Freemont, California, 1989
26. K. Iwano, T. Hirose, E. Kamibayashi, S. Furui, Audio-Visual Person Authentication Using Speech and Ear Images, *Proc. of Workshop on Multimodal User Authentication, 2003,* pp.85-90.
27. I. T. Jolliffe, Principal Component Analysis *(New York: Springer),* 1986
28. STATE v. David Wayne KUNZE, *Court of Appeals of Washington, Division 2.* 97 Wash. App. 832, 988 P.2d 977, 1999
29. K. Messer, J. Matas, J. Kittler, J. Luettin, G. Maitre, XM2VTSDB: The Extended M2VTS Database, *Proc. AVBPA'99* ,Washington D.C., 1999
30. B. Moreno, A. Sanchez, On the Use of Outer Ear Images for Personal Identification in Security Applications, *IEEE 33rd Annual Intl. Conf. on Security Technology, 1999,* pp. 469-476.
31. Z. Mu, L. Yuan, Z. Xu, D. Xi, S. Qi, Shape and Structural Feature Based Ear Recognition1, *Sinobiometrics 2004,* LNCS 3338, 2004, pp. 663-670.
32. J. L. Northern, M. P. Downs, Hearing in Children, *Lippincott Williams & Wilkins,* Fifth Edition, 2002
33. K. Pun, Y. Moon, Recent advances in ear biometrics, *Proc. of the Sixth International Conference on Automatic Face and Gesture Recognition, 2004,* pp. 164-169.
34. G.N. Rutty, A. Abbas, D. Crossling, Could earprint identification be computerised? An illustrated proof of concept paper, *International Journal of Legal Medicine, 2005,* no.6, pp. 335-343.
35. M. M. Rahman, S. Ishikawa, Proposing a Passive Biometric System for Robotic Vision, *Proc. of the Tenth International Symposium on Artificial Life and Robotics, 2005,* Oita, Japan.
36. R. L. Goode, Auditory Physiology of the external ear, *Physiology of the ear,* San Diego, Calif. : Singular, 2001. pp. 147-159.
37. US Patent 5,787,187. Systems and methods for biometric identification using the acoustic properties of the ear canal. *Scandia.* 1998
38. R. Teranishi E. Shaw, External-Ear Acoustic Models with Simple Geometry, *The Journal of the Acoustical Society of America,* 1968, vol 44, pp. 257-263.
39. M. Turk and A. Pentland, Eigenfaces for recognition, *Journal of Cognitive Neuroscience,* Vol. 3, No. 1, pp. 71-86, Winter 1991.

40. B. Victor, K.W. Bowyer, S. Sarkar, An evaluation of face and ear biometrics, *Proc. ICPR 2002,* pp. 429-432.
41. Y. Wang, H. Turusawa, K. Sato and S. Nakayama, Study on Human Recognition with Ear *Image, Information Processing Society of Japan (IPSJ)* Kyushu Chapter Symposium, 2003.
42. P. Yan, K.W. Bowyer, 2D and 3D ear recognition, *Biometric Consortium Conference,* 2004.
43. P. Yan and K. W. Bowyer, Biometric recognition using three-dimensional ear shape, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* to appear.
44. P. Yan, K. W. Bowyer, ICP-based approaches for 3D ear recognition, *Biometric Technology for Human Identification II, Proceedings of SPIE,* 2005, vol. 5779, pp. 282-291.
45. P. Yan, K. W. Bowyer, Empirical evaluation of advanced ear biometrics, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops,* 2005, page 41.
46. P. Yan, K. W. Bowyer, A Fast Algorithm for ICP-Based 3D Shape Biometrics, *Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID),* October 2005, New York, pp. 213-218.
47. T. Yuizono, Y. Wang, K. Satoh, S. Nakayama, Study on Individual Recognition for Ear Images by Using Genetic Local search, *Proc. of the 2002 Congress on Evolutionary Computation, 2002,* pp. 237-242.
48. H. Zhang, Z. Mu, W. Qu, L. Liu, C. Zhang, A novel approach for ear recognition based on ICA and RBF network, *Proc. of the Fourth International Conference on Machine Learning and Cybernetics,* 2005, pp. 4511-4515.