

UNIVERSITY OF SOUTHAMPTON

Force Field Feature Extraction for Ear Biometrics

By

David J. Hurley

A thesis submitted for the degree of
Doctor of Philosophy

Department of Electronics and Computer Science
Faculty of Engineering and Applied Science

September 2001

“While the objective in defining feature space is to reduce the dimensionality of the pattern space yet maintaining discriminatory power for classification purposes, successful transformations along these lines are still in their infancy” [1].

Harry C. Andrews, USC, 1972
Mathematical Techniques in Pattern Recognition

University of Southampton

Abstract

FORCE FIELD FEATURE EXTRACTION
FOR EAR BIOMETRICS

by David J. Hurley

Supervisors: Dr. Mark Nixon and Dr. John Carter

Department of Electronics and Computer Science

The overall objective in defining feature space is to reduce the dimensionality of the original pattern space, whilst maintaining discriminatory power for classification. To meet this objective in the context of ear biometrics a novel force field transformation is introduced in which the image is treated as an array of mutually attracting particles that act as the source of a Gaussian force field. In a similar way to Newton's Law of Universal Gravitation pixels are imagined to attract each other according to the product of their intensities and inversely to the square of the distance between them. Underlying the force field there is a scalar potential energy field, which in the case of an ear takes the form of a smooth surface that resembles a small mountain with a number of peaks joined by ridges. The peaks correspond to potential energy wells and to extend the analogy the ridges correspond to potential energy channels. The directional properties of the force field are exploited to automatically locate these wells and channels, which then form the basis of a set of characteristic ear features. The new features are robust especially in the presence of noise, and have the advantage that the ear does not need to be explicitly extracted from its background. The directional properties of the ensuing force field lead to two equivalent extraction techniques; one is algorithmic and based on field lines, while the other is analytical and based on the divergence of force direction. The technique is validated by performing recognition on a database of ears selected from the XM2VTS face database. This confirms not only that ears do indeed appear to have potential as a biometric, but also that the new approach is well suited to their description.

Acknowledgements

I would like to express my gratitude to my supervisor Dr. Mark Nixon for his encouragement, inspiration, and infectious good humour in making this research such an enjoyable journey of discovery. Thanks are also due to my second supervisor Dr. John Carter for his invaluable feedback, especially for his contribution of FFT convolution expertise. I would also like to thank my friends and colleagues for their generous help and advice, especially the enthusiasm and encouragement of Dr. Jeff Wood, Dr. Mark French, and Dr. Yannick Marchand.

Table of Contents

Acknowledgements	i
Table of Contents.....	ii
List of Figures	vi
List of Tables.....	viii
Anatomy of the Human Ear	1
Context and Contributions.....	2
Thesis Hypothesis.....	7
Author's Publications	8
Thesis Outline	9
1 Introduction.....	10
<i>1.1 Famous Impersonations.....</i>	<i>12</i>
1.1.1 The Case of Martin Guerre.....	12
1.1.2 The Tichborne Claimant	13
1.1.3 Grand Duchess Anastasia.....	13
<i>1.2 Personal Identification Evolution</i>	<i>15</i>
1.2.1 Alphonse Bertillon.....	15
1.2.2 Henry Faulds	16
1.2.3 Sir Francis Galton	16
1.2.4 Alfred Iannarelli	17
1.2.5 Fingerprint Evidence Convictions	18
1.2.6 Kunze Ear Decision	19
<i>1.3 Automated Personal Identification.....</i>	<i>20</i>
1.3.1 Biometric Parameters.....	21

1.3.2	Biometric Technologies	24
2	The Force Field Transform.....	27
2.1	<i>Potential Energy Transform Definition.....</i>	28
2.2	<i>Potential Energy Function.....</i>	28
2.3	<i>Potential Energy Surface.....</i>	29
2.4	<i>Potential Wells and Channels.....</i>	29
2.5	<i>Dome Shape Explanation.....</i>	30
2.6	<i>Force as Gradient of Potential</i>	31
2.7	<i>Force Field Transform Definition.....</i>	32
2.8	<i>Force Field Transform Linearity</i>	33
2.9	<i>Potential Transform Linearity</i>	35
2.10	<i>Force Field Computation by Convolution.....</i>	36
2.11	<i>Invertibility.....</i>	38
2.11.1	Numerical Testing for Invertibility.....	38
2.11.2	Condition Number and Invertibility	41
2.11.3	Invertibility Conclusions.....	42
2.12	<i>Invariant Properties</i>	43
2.12.1	Illumination Invariance	43
2.12.2	Scale Invariance.....	44
2.12.3	Noise Tolerance.....	44
3	Force Field Feature Extraction.....	45
3.1	<i>Force Field Magnitude.....</i>	47
3.2	<i>Exploratory Test Pixels.....</i>	49
3.3	<i>Field Lines</i>	50
3.4	<i>Channel and Well Formation.....</i>	51
3.5	<i>Channel and Well Formation for an Ear.....</i>	52

3.6	<i>Invariance Measures</i>	54
3.7	<i>Initialisation Invariance</i>	55
3.8	<i>Scale Invariance</i>	57
3.9	<i>Noise Tolerance</i>	58
3.10	<i>Uniqueness of Channels and Wells</i>	59
3.11	<i>Convergence Feature Extraction</i>	62
3.12	<i>Convergence and Marr-Hildredth</i>	66
3.13	<i>Convergence and Curvature</i>	67
4	Ear Recognition	68
4.1	<i>Ear Database</i>	70
4.2	<i>Template Matching</i>	73
4.3	<i>Classification Results</i>	75
4.4	<i>Statistical Analysis of Results</i>	77
4.5	<i>Edge Detection Template Matching</i>	80
5	Further Work	84
5.1	<i>Autocentricity</i>	84
5.2	<i>Feature Matching and Indexing</i>	84
5.3	<i>Generating an Ear Database</i>	85
5.4	<i>Miscellaneous Further Work</i>	86
5.5	<i>Polar Functions</i>	87
5.6	<i>Synthetic Shape Recognition</i>	90
6	Conclusions	98
6.1	<i>Force Field Transform</i>	98
6.2	<i>Feature Extraction</i>	99
6.3	<i>Ear Recognition</i>	100

Appendices	101
<i>A1 Alternative Proof of Transform Linearity.....</i>	<i>101</i>
<i>A2 Representative Sample of Ear Results.....</i>	<i>103</i>
<i>A3 Binarised Thresholded Convergence Functions</i>	<i>114</i>
<i>A4 XM2VTS Identification Numbers of Ear Subjects.....</i>	<i>118</i>
References	119

List of Figures

Figure 1 Anatomy of the human ear.....	1
Figure 2 Iannarelli ear measurements.....	17
Figure 3 Potential function of an isolated pixel	28
Figure 4 Energy surface for an ear viewed from below the lobe.....	29
Figure 5 Energy surface as a sum of components	30
Figure 6 Force field calculation at the centre of a 3×3 pixel image.....	32
Figure 7 Geometry of a simple 4-pixel image.....	33
Figure 8 Skew symmetry is due to mutual attraction between pixels.....	35
Figure 9 Force field by convolution in Mathcad.....	37
Figure 10 Energy transform matrix inversion for 2×2 image.....	39
Figure 11 Force field magnitude for an ear image	47
Figure 12 Effect of quantised versus continuous direction of test pixel movement	49
Figure 13 Force field due to a small white cross.....	50
Figure 14 Formation of potential channels and wells for a cross shaped image	51
Figure 15 Field line, channel, and well formation for an ear.	52
Figure 16 Inter-well distances as the basis of an invariance measure.....	54
Figure 17 Initialisation invariance for 261×170 pixel ear image.....	55
Figure 18 Scale invariance. Image ratios (a):(b):(c) :: 3.4:2:1	57
Figure 19 Zero mean Gaussian noise tolerance.....	58
Figure 20 Uniqueness of wells and channels as ear features.....	60
Figure 21 Divergence of a vector field.....	63
Figure 22 Convergence field.....	65
Figure 23 Correspondence between channels and convergence.....	65
Figure 24 Principal field lines.....	69
Figure 25 Feature extraction for subject 000	71
Figure 26 Feature extraction on subject 001	72
Figure 27 Binarised thresholded convergence function.....	73
Figure 28 A selection of binarised and thresholded convergence functions	74
Figure 29 Cross-correlation matrix depicted as a grey scale image	75
Figure 30 Nearest neighbour classification matrix.....	76
Figure 31 Normal distribution fitted to frequency distribution of cross-correlations	78
Figure 32 Probability density functions of intra-class and inter-class correlations	78
Figure 33 Receiver operator characteristic	79
Figure 34 Sample of results of thresholded and binarised edge detection	80
Figure 35 Nearest neighbour classification matrix for edge detection.....	81
Figure 36 Normal distribution fitted to frequency distribution for edge detection.....	82

Figure 37 Probability density functions for edge detection	82
Figure 38 Receiver operator characteristic for edge detection.....	83
Figure 39 Mean and variance polar functions in operation.....	89
Figure 40 Twelve basic pentominoe shapes.....	90
Figure 41 Scaling and rotation of the letter 'F'	91
Figure 42 Force field magnitudes corresponding to the letter 'F'	91
Figure 43 Convergence fields corresponding to the letter 'F'	92
Figure 44 Mean distance polar functions corresponding to the letter 'F'	93
Figure 45 Nearest neighbour classification matrix.....	94
Figure 46 Normal distribution fitted to frequency distributions.....	96
Figure 47 Probability density functions for pentominoe recognition	96
Figure 48 Receiver operator characteristic for pentominoe recognition	97

List of Tables

Table 1	Determinants corresponding to the first 32 square images.....	40
Table 2	Determinants corresponding to the first 8 non-square images.....	40
Table 3	Summary of trials resulting in pairs of identical maximum correlation values.....	76
Table 4	Statistics for intra-class and inter-class convergence correlations	77
Table 5	Statistics for intra-class and inter-class edge detection correlations	81
Table 6	Statistics for intra-class and inter-class angle comparisons.....	95
Table 7	XM2VTS identification numbers of ear subjects.....	118

Anatomy of the Human Ear

The ear does not have a completely random structure, it is made up of standard features just like the face. The parts of the ear are less familiar than the eyes, nose, mouth, and other facial features but nevertheless are always present in a normal ear. These include the outer rim (helix), the ridge (antihelix) that runs inside and parallel to the helix, the lobe, and the distinctive u-shaped notch known as the intertragic notch between the ear hole (meatus) and the lobe. Figure 1 shows the locations of the anatomical features in detail.

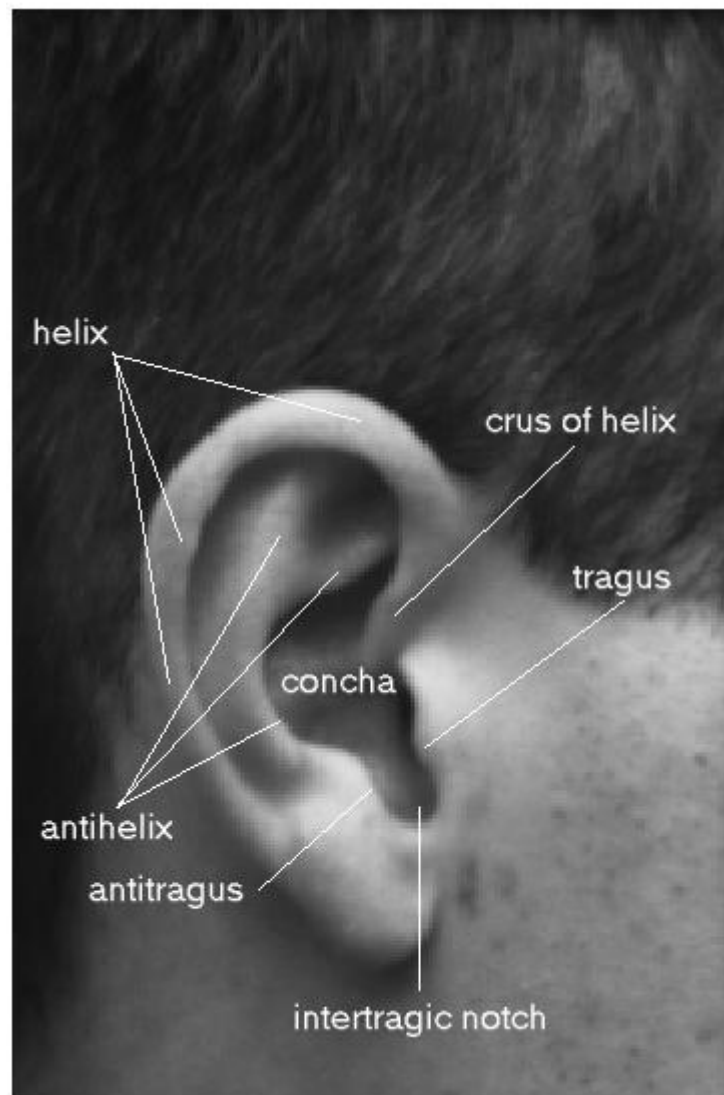


Figure 1 Anatomy of the human ear

Context and Contributions

In machine vision, ear biometrics refers to the automatic measurement of distinctive ear features with a view to identifying or confirming the identity of the owner. It has received scant attention compared with 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 Holland where more than 250 ear print convictions are secured each year [2], and to a lesser extent in the United States, where an ear classification system has been developed by Alfred Iannarelli, and has been in use for more than 40 years [3,4].

More recently in the UK, a burglar was convicted of murder on the basis of an ear print left on a window at the scene of the crime [5,6]. The US Immigration and Naturalisation Service apparently recognises the value of the ear for identification, since it requires that the right ear be clearly visible in identification photographs [7]. The University of Glasgow forensic face reconstruction expert, Prof. Peter Vanezis, who is pioneering ear print evidence in the UK, has attempted to resolve the Grand Duchess Anastasia mystery by comparing the claimant's ear to an old photograph of the Duchess [2,8]. A feasibility study for an automated system for ear identification has been undertaken by Burge and Burger at Johannes Kepler University in Austria [9].

The potential of the ear for recognition was recognised and advocated as long ago as 1890 by the French criminologist Alphonse Bertillon [13]: -

“ The ear, thanks to these multiple small valleys and hills that plough across it, is the most important 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 life in the womb.”

*Alphonse Bertillon
Legal photography, Paris, Gauthier-Villars, 1890*

Ears have certain advantages over the more established biometrics; as Bertillon pointed out, they have a rich and stable structure that is preserved well into old 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 direct contact fingerprint scanning, and is not likely 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, although less so than the face or gait.

An ear recognition system could be used like other biometric systems, say for access control. A database or register would be prepared by processing images of the ears of authorised personnel, to extract a set of characteristic features for each image. Personnel wishing to enter would have their ears scanned at the entrance and the image would be processed and compared for a match against the register. The stored features would have to be sufficiently distinct so as to be able to distinguish one ear from all the others, and sufficiently robust so that the same features would be produced every time the ear is scanned. These are conflicting requirements and present a challenge to the system designer.

There are a number of techniques with potential to find and describe a human ear for computer vision recognition. Clearly, there are application constraints, such as occlusion by hair, but here the concern is with basic technique. For security systems this problem is easily overcome by simply holding the hair to one side during image capture. Ear extraction could use an active contour [14,15] but with initialisation problems that could be relieved by a dual active contour [16], but this still requires establishment of inner and outer contours. Once the ear image has been extracted, we need to describe its shape for recognition using low-level feature extraction where we seek to describe the image with invariance properties. This can be achieved by edge detection [17,18,19], which is invariant for change in the overall illumination level. It can also be achieved by phase congruency [20,21], which can tolerate

local change in image intensity. We could also use more global techniques such as moments [23] or shape from shading [24]. Burge and Burger [9] have combined Canny edge detection with graph matching to describe the ear. A deformable model of the human ear might be used for description, where individual ears would be characterised by the model parameter values that correspond to the deformations needed to make the model fit the given ear shape.

To address these issues, a novel two-stage approach has been developed to provide ear extraction and description concurrently in a reliable and robust manner. Our new approach has new properties and affords a new description with different advantages, to be described later. The two stages in the new approach are: *Image to Force Field Transformation*, and *Potential Well and Channel Extraction*. Firstly, the entire image is transformed into a force field by supposing that each pixel exerts an isotropic force on all the other pixels, which is proportional to the pixel's intensity and inversely proportional to the distance squared to each of the other pixels. Secondly, the directional property of the force field is exploited to automatically locate a small number of potential wells and channels, which correspond to local energy peaks and ridges respectively in the scalar potential energy surface that underlies the vector force field, and it is these wells and channels that form the basis of the new features.

Two equivalent feature extraction techniques have been developed; an algorithmic approach, based on field lines, and an analytical approach, based on the divergence of force direction. In the first approach, notional test pixels spaced evenly around the perimeter of the force field flow towards the centre, following the force direction, tracing out field lines. The field lines flow into a small number of channels that eventually terminate in wells, perhaps joining up with other channels on the way. The channel structure can be thought of as modulating the natural flow of the field lines towards a single well at the centre of the field, which is what would occur if the image consisted of only one constant grey level throughout. This centric property is due to the basic dome shape of the energy field, which will be explained later.

In the second approach, as a result of analysing the mechanism of the field line approach, it was discovered that this approach has an analogous mathematical description, with its own unique advantages. The analysis revealed that the channels and wells correspond to ridges and peaks in a mathematical function, revealing that they are really manifestations of the same phenomenon. This function, which will be called convergence, maps the force field to a scalar field where the output is the additive inverse of the divergence of the direction of the force field. The convergence function provides a richer description than just channels and wells because it includes negative versions of channels and wells, corresponding to positive divergence ridges and peaks respectively.

The high dimensionality of the original image is traded for a much simpler feature description in the form of a set of channels and wells. In the early stages of this research it was hoped that the wells alone would suffice as ear descriptors, which would have led to a particularly attractive solution in the form of a small set of coordinates. However, it became clear that features for some ears only included two wells, which is clearly inadequate for realistic description and classification. Also, the technique when applied to arbitrary shapes of constant grey level sometimes only found one well. For this reason the richer but more complicated description offered by the channels and wells together became the main focus of the research.

The features thus far described are not in a simple form suitable for direct comparison; indeed the precise nature of a channel has not been defined, other than to suggest that it is the path shared by two or more field lines that happen to converge. Nevertheless, channels generally take the form of a small number of curves, perhaps meeting each other at one or more central wells. There is an implicit and reasonable assumption that such a collection of curves would be easier to compare than comparing the original image, perhaps using techniques from differential geometry [22]. There is also an assumption that such curves are sufficiently distinctive and robust to act as good descriptors. There is little point in comparing them if they do not adequately describe the ear shapes upon which they are based.

To qualify as good ear features the essential character of the features must be preserved when the images are scaled, and they should not suffer

unduly from noise problems, as may happen for example with moments [23]. It would also be an advantage if a certain amount of occlusion by hair could be tolerated. It is confirmed that the new force field features scale correctly when the image is resized, and the technique is also found to be very insensitive to noise, due to the force at any pixel position being a function of all the other pixels, implying large scale filtering. The field line extraction process is very robust since the number and precise location of the initial test pixels does not matter as long as the image is adequately covered; field lines always flow towards wells, usually forming channels on the way. Thus the new extraction technique does not suffer from the usual initialisation problems encountered by other extraction techniques such as active contours [14,15].

Mathematical modeling techniques traditionally used in physics have recently attracted the attention of researchers in computer vision; for example [25] describes the use of Vector Potential to extract corners by treating the Canny edge map of the image as a current density. Also, a recent approach [26] used a Potential Field model in a medial axis transform. Active contours [14,15,16] operate by treating image segmentation as an energy minimisation problem, and Xu [27] extends the active contour model by replacing the external local force with a force field derived from the image edge map.

A selection of four samples taken from each of sixty-three subjects drawn from the XM2VTS face profiles database has been used to test the viability of the technique [28]. A classification rate of better than 91% has been achieved so far, using just simple template matching on the basic channel shapes, demonstrating the merit of the technique at least at this scale. It can reasonably be expected that the use of more sophisticated channel matching techniques would lead to even better results. As such, a new approach has been demonstrated with success in a new application domain.

Thesis Hypothesis

The hypothesis is threefold in that there are three separate contributions; each allocated a separate chapter. In a sentence we aim to show that

1. **Ear Recognition** is viable,
2. using a **New Force Field Feature Extraction** technique,
3. based on a **New Force Field Transform**.

In more detail we aim to show that in the context of ear biometrics, we have developed a new linear transform that transforms an ear image with very powerful smoothing and without loss of information, into a smooth dome shaped surface, whose special shape facilitates a new form of force field feature extraction that extracts the essential ear signature without the need for explicit ear extraction, and in the process we aim to verify the recognition potential of the human ear for biometrics.

One potential difficulty is that we are attempting to prove the new transform and extraction technique at the same time as trying to demonstrate that ears are sufficiently distinctive to act as a biometric source. No large scale study of the feasibility of ear recognition has ever been carried out, and it could even be argued that the new technique is being hampered from showing its full image processing potential by limiting its use to ear recognition.

Author's Publications

Hurley, D. J., Nixon, M. S. and Carter, J. N. (2001), Force Field Energy Functionals for Image Feature Extraction, Invited paper – *Image and Vision Computing*, Special Issue on BMVC 99.

Hurley, D. J., Nixon, M. S. and Carter, J. N. (2000), A New Force Field Transform for Ear and Face Recognition, *Proceedings IEEE International Conference on Image Processing ICIP2000*, pp. 25-28

Hurley, D. J., Nixon, M. S. and Carter, J. N. (2000), Automatic Ear Recognition by Force Field Transformations, *Proc IEE Colloquium: Visual Biometrics* pp. 8/1-8/5.

Hurley, D. J., Nixon, M. S. and Carter, J. N. (1999), Force Field Energy Functionals for Image Feature Extraction. *Proc. 10th British Machine Vision Conference BMVC99* pp. 604-613.

Thesis Outline

Chapter 1 Introduction

The first chapter reviews the history of ear identification in the wider context of personal identification, from historical cases of impersonation, then on to forensic science, through to modern automated personal identification in the form of biometrics.

Chapter 2 Force Field Transform

This chapter defines the force field and potential energy transforms and establishes some useful properties. They are shown to be linear and invertible under certain circumstances. Feature extraction without the requirement for explicit ear extraction is shown to be due to the basic dome shape of the energy field. Scale and illumination invariance is analysed, followed by a noise sensitivity analysis.

Chapter 3 Force Field Feature Extraction

Two equivalent forms of feature extraction are presented; one algorithmic and based on field line direction; the other is a convergence function based on force direction divergence. The function is shown to be similar in form to the Marr-Hildreth operator, but with important differences.

Chapter 4 Ear Recognition

Rudimentary ear recognition is carried out using template matching of potential channels in the form of binarised and thresholded convergence functions. The results are analysed and compared with the results for simple edge detection.

Chapter 5 Further Work

A number of outstanding issues not covered in the main body are briefly considered. A technique based on the statistical moments of the convergence functions, and which returns sets of closed contours that are readily adapted to be scale and rotation invariant, is discussed.

1 Introduction

This section sets the stage by reviewing the fascinating historical development of personal identification from renowned historical cases of impersonation, to the development of anthropometrics and forensic science, through to present state-of-the art biometric systems and techniques.

Some famous impersonation cases are reviewed starting with the 16th century French case of Martin Guerre who was hanged for impersonating a missing husband. In England there is the Victorian case of the *Tichborne Claimant*, Arthur Orton who was sentenced to 14 years for impersonating the missing heir to a large Hampshire fortune. Lastly we will take a look at the case of Anna Anderson-Manahan, who claimed to be the Grand Duchess Anastasia of Russia, the sole survivor of the Romanov family, who was believed to have escaped the cruel fate of her parents and sisters. An attempt was even made to resolve the issue by comparing photographs of their ears [8]. It should come as no surprise that all these cases involved considerable inheritances, which emphasises the need for reliable personal identification to establish rightful ownership or authority.

The Tichborne claimant case was very widely reported and followed in Victorian times and may have driven early attempts to systematically identify people. Indeed Faulds, who was the first to suggest fingerprint identification, actually speculates in his letter [29] published in *Nature* in 1880, whether in fact fingerprints might be used to resolve the Tichborne case. During the second half of the 19th century all the pioneers in personal identification were laying its foundations; Adolphe Quetelet the Belgian statistician [30], the Scottish missionary, Dr. Henry Faulds in Japan [29], Sir William Herschel in India, grandson of the famous astronomer of the same name [31], Alphonse Bertillon in France [13,33,34], Sir Francis Galton in England [35,36,37], and also Sir Edward Henry in India [38].

The system of criminal personal identification and classification known as Bertillonage after its founder, eventually gave way to fingerprinting in the early part of the last century. Fingerprint evidence has since then taken its place in the public perception as indisputable scientific evidence. However,

we shall review some recent court cases that have thrown serious doubts on the universal assumption that fingerprint evidence is infallible. The question as to whether ears can be used to uniquely identify an individual naturally arises and this is also considered in the light of a recent US appeals court ruling on the admissibility of earprint evidence.

We round off this chapter by considering the question of automated personal identification in the form of biometrics and the evolving technologies of electronic signatures. Some underlying biometrics issues such as biometric uniqueness are considered by viewing biometrics in a classification context. Biometric parameters are reviewed, followed by a review of the present state of the art in biometric technologies, including some systems that are presently deployed and some that are under development.

1.1 Famous Impersonations

One of the earliest and most notorious cases of personal identification is recounted in the biblical tale of two women in Jerusalem who were brought before King Solomon to resolve the disputed ownership of a baby, whereupon Solomon, resorting to his renowned wisdom, ordered that the baby be halved, trusting that the true mother would give up her baby rather than accept such a drastic solution [39]. Nowadays, with the advent of DNA testing [40], any such dispute would easily be resolved without resorting to such drastic internecine measures. Nevertheless, as some of the famous identity disputes below will show, up until recently proving one's identity has not always been as straight forward as most of us would take for granted.

1.1.1 The Case of Martin Guerre

The first case concerns the impersonation of 16th century French farmer Martin Guerre by a soldier, Arnaud du Tilh, who closely resembled him, and allegedly left him for dead on the battlefield after obtaining his life story. The impostor successfully masqueraded as Guerre, fooling his father and even his wife. But when he tried to claim Martin's inheritance, the father denounced him and the case resulted in two trials. Arnaud gave a remarkable account of himself in the trials with many witnesses testifying on his behalf and it seemed that he was about to win his case when the real Martin suddenly returned and denounced him, resulting in Arnaud's hanging.

Since this remarkable event was first reported in 1561 by the trial judge Jean de Coras [41], it has found a place in French popular culture and has been retold in various books and films. Also in France there is the case of a woman who claimed to be the deceased marquise de Douhault, who became known as *la femme sans nom*, the woman with no name, after the courts not only dismissed her claim, but also declared that she was not Anne Buiret, as claimed by the relatives of the marquise. Given this background, it is perhaps not surprising that France gave birth to the father of anthropometric personal identification, Alphonse Bertillon.

1.1.2 The Tichborne Claimant

One of the most extraordinary cases in British legal history concerns the case of Arthur Orton, who claimed to be the heir to the Tichborne family fortune after the real heir had been missing and presumed drowned for more than 11 years. Roger Tichborne, born in Hampshire in 1829, was the heir to the considerable Tichborne family fortune, which would have made him a very wealthy man. At age 23, having spent most of his youth in France, he embarked on a world tour. His ship left Rio de Janeiro but never made land and was presumed lost at sea in 1854. All on board were eventually declared dead but his mother had never given up hope of finding her son alive and placed advertisements in the Australian newspapers in 1865, some three years after his father had died. In response to the advertisement a man came forward claiming to be her son. The mother and some of her family believed his story, but most of the family believed him an imposter and so he took the trustees to court to claim his inheritance. After a civil trial lasting more than three months, the jury found for the defendants and he was then charged with perjury. After a criminal trial lasting more than 6 months he was found guilty and sentenced to 14 years in prison.

1.1.3 Grand Duchess Anastasia

Even more recently there is the case of Anna Anderson-Manahan who was believed to be the youngest daughter of Tsar Nikolai Romanov II of Russia. A few years after the execution of the Tsar and his family, Anna Anderson was rescued from a suicide attempt in a German canal. She was taken to a Berlin mental hospital where in 1921 she declared that she was the Grand Duchess Anastasia, claiming to have been rescued by one of the guards and taken to Romania. She made her way to Germany to seek out her relatives but became depressed when she could not find them and tried to commit suicide. Only a few of the Romanov relatives believed her story and it was not until 1967 that the courts, in response to the relative's request to settle the estate, declared that she was not Anastasia. One interesting piece of supportive evidence in Anna's favour came in the form of gait identification. Gled Botkin, who was a

childhood playmate of Anastasia's, and the son of the family doctor, was when he met her ten years later, immediately convinced by the way she walked that she was indeed Anastasia.

In 1993 forensic scientist Peter Vanezis, based at the University of Glasgow, performed an identification experiment based on matching of ears from photographs of Anna Anderson and Anastasia [8]. The test confirmed that the ears in both cases were similar, but a positive ear identification is less meaningful than a negative one; it would have been far more conclusive if the test showed striking differences.

1.2 *Personal Identification Evolution*

Despite the evidence to the contrary afforded by the extraordinary cases just mentioned, it appears that the most reliable form of personal identification historically consisted of confirmation by a close friend or relative that you are, or are not, who you claim to be. This form of identification includes not only physiological details mainly in the form of appearance but also a whole host of behavioural details including mannerisms, speech peculiarities, and gait. And then there is shared knowledge such as family history or common experiences shared by friends. However, this very reliable form of identification has not always been possible or convenient, especially when there is an inheritance at stake. Sometimes it is advantageous not to be identified as one's former self. Such was commonly the case for French habitual criminals, because evidence of previous offenses would automatically guarantee a longer sentence. It was to prevent such misrepresentation that Alphonse Bertillon devised his famous anthropometric system of cataloguing prisoners.

1.2.1 Alphonse Bertillon

One of the earliest systems of measurement and recording of personal traits for identification was devised by Alphonse Bertillon who worked in the French penal system [13,32,34]. In an early version of the system described by Sir Francis Galton [36] four measurements were taken from each inmate, the width and length of the head, and the lengths of the foot and middle finger. Each of the measurements was classified as large, medium, or small, leading to a total of 81 classes. Other distinguishing features were noted on a card, which was then pigeonholed, thus creating an anthropometrically indexed filing system. Later he extended the system to include eleven measurements, one of which was the length of the ear. Galton correctly pointed out that Bertillon's system suffered from correlation of measurements in an individual; a long foot is likely to be accompanied by a long finger [36]. However, Bertillon's system was generally successful and was widely used throughout the world until it was replaced by fingerprinting in the early part of the 20th century.

1.2.2 Henry Faulds

Henry Faulds was a Scottish doctor who spent some years working in Japan as a missionary. In 1880 he published a letter in Nature [29] describing how his discovery of finger-marks, made on a small piece of ancient pottery while the clay was still soft, had driven him to investigate the skin-furrows in human fingers generally. He described in detail the loops, whorls, and minutiae that are used today in fingerprint identification. He went on to describe how to collect prints from a subject, by pressing the finger on a sheet of tin spread thinly with printer's ink and then pressing the tin sheet on to a sheet of slightly damp paper. He made one remark that starkly contradicts contemporary claims for the uniqueness of fingerprints: -

“The dominancy of heredity through these infinite varieties is sometimes very striking. I have found unique patterns in a parent repeated with marvelous accuracy in his child”

Two other remarkable events were also reported. He used fingerprints to help identify a servant who had stolen some rectified spirit, and he also used prints to prove the innocence of a suspect in another case. So, not only was Faulds the first person to suggest forensic use of fingerprints, but he was also the first person to prove guilt and to prove innocence, using fingerprinting.

1.2.3 Sir Francis Galton

Francis Galton gave a lecture at the Royal Institution on May 25, 1888 during which he described a machine he had developed called a *mechanical selector*, for automatically measuring the profile of the human face [36]. During the second part of his lecture he went on to consider other personal characteristics that might be used to identify a person. He made an interesting remark that the makers of glass eyes had reported thousands of varieties in the markings of the iris of the eye, a somewhat prophetic remark since the iris is now one of the most promising biometrics [48]. He went on to briefly mention the identification potential of ears: -

“I shall not dwell now upon these, nor on such peculiarities as those of hand-writing, nor on the bifurcations and interlacements of the superficial veins, nor on the shape and convolutions of the ear”

In the same lecture he went on to describe his interest in fingerprints and also made reference to Sir William Herschel’s work in India over the previous 28 years. Galton published a book in 1892 simply titled *fingerprints* [35].

1.2.4 Alfred Iannarelli

Alfred Iannarelli is an American consultant in forensic investigation and criminal investigation who developed a system of ear classification used by American law enforcement agencies. In late 1949, after spending two years with a military security and law enforcement unit and while attending a police academy in Oakland, California 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, which he published in book form [4] in 1964 and updated in 1988.

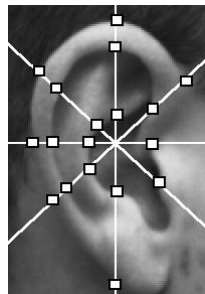


Figure 2 Iannarelli ear measurements

Essentially his system consists of taking 12 measurements around the 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 intersection between the antihelix and 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 [3].

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”

1.2.5 Fingerprint Evidence Convictions

It is well over a hundred years since the first conviction was secured on the basis of fingerprint evidence [29] and as a result of countless real and fictional police dramas most people take the infallibility of fingerprint evidence for granted. If your fingerprints match those found at the scene of the crime, then you must have put them there. Yet in recent years serious doubts have begun to arise as to the infallibility of fingerprint evidence [10,42]. Certainly, when the patterns of all ten fingers are considered as an ensemble, it seems most reasonable to suppose that they are unique. In reality however, convictions are very often secured with just a single fingerprint, and even then the latent fingerprint is often only a partial print and of poor quality.

Such was the remarkable fate of Scottish police officer, Shirley McKie, who was wrongly convicted of perjury on the strength of fingerprint evidence that placed her at the house of Marion Ross, 55, who was murdered at home in Kilmarnock in 1997. McKie, who denied ever being at the scene, had her conviction overturned [11] after an American fingerprint expert cast doubt on the fingerprinting technique used by the Scottish Criminal Records Office. The SCRO officers who gave evidence were subsequently suspended, pending a review of the fingerprint matching process.

Science historian Simon Cole points out that no serious scientific study has ever been carried out to confirm the uniqueness of fingerprints, and that no

error rates are available for fingerprint identification. He further suggests that fingerprint examiners may match prints to the wrong person as much as 20 per cent of the time [10,42].

1.2.6 Kunze Ear Decision

In the 1993 Daubert case, the US Supreme Court ruled that scientific evidence has to be testable, subject to scrutiny by other scientists, must be accompanied by quantitative estimates of uncertainty, and must be generally accepted by the scientific community [43]. This ruling led to pre-trial Daubert hearings, where a judge tests the admissibility of scientific evidence. The admissibility of ear print evidence was recently put to such a test.

In Washington State in 1997 David Wayne Kunze was convicted of murder and was sentenced to life imprisonment on the basis of two expert witnesses testifying that a latent ear print found on a bedroom door could only have been made by Kunze. The murder conviction was subsequently appealed and the appeal court ruled [3] that the trial court erred by allowing the expert witnesses to testify that Kunze was the likely or probable maker of the latent print, on the basis that latent ear print identification is not generally accepted in the forensic science community. A point of interest is that one of the two expert witnesses was the veteran Dutch ear print police officer who has pioneered ear print evidence in Holland where more than 250 ear print convictions are secured annually [2]. In response to the US appeals court ruling, a large-scale study involving 10,000 subjects has been proposed by Prof. Andre Moenssens to determine the variability of the ear across the population [12].

1.3 Automated Personal Identification

We can look forward to one day enjoying a cashless, keyless, passless society thanks to biometrics. As our society becomes increasingly electronically networked so will be the need for ever-more reliable and secure biometric personal identification. Many people now routinely purchase products and services over the Internet, and banks and building societies are lining up to offer electronic banking services. Yet the lack of a legally binding electronic equivalent of the traditional signature has stood in the way of the full-scale take up of electronic commerce [45]. This need is being addressed by the development of electronic signature protocols, which have recently been given legal recognition by legislation in the USA [44] and Europe [45,46].

In the same way as the behavioural link of the traditional signature to its author inspires confidence in its authenticity, so will the ultimate goal of including biometric authentication instill public confidence in the electronic signature. Just as an individual signature can be very simple or elaborate, so individuals could adjust their own authentication policies, requiring as much or as little biometric content in the authentication process, as they see fit.

Biometrics in the present context refers to the automatic identification or confirmation of the identity of a person based on their physiological and behavioural characteristics. It can be considered as a classification problem in the overall context of pattern recognition, where each individual represents one class. A successful biometric must therefore have sufficient complexity to achieve adequate class separation of a very large number of classes whilst at the same time it must be sufficiently stable across samples of the same class to ensure low misclassification rates.

Most biometric systems operate at a single resolution; a face-based system does not work at the level of detail of an iris or retina based system, even though these are included in the face. It is clear from the outset therefore that such systems are unlikely to assign a single class to each individual on the planet. It is well known that famous people have look-alikes, so it is reasonable to assume that this equally applies to each of us, although we are unlikely to ever meet them. However, the look-alikes are not likely to have

the same freckle pattern or the same iris pattern, so by including more and more information either by multi-resolution or multi-modal biometrics the class size can be expected to be reduced to a size of just one if required, but at a higher cost.

To complete this introduction to automated personal identification, some practical biometrics issues will be reviewed in terms of biometrics parameters and in terms of typical systems already deployed such as face, eye, hand, and fingerprint; and those that are still at the laboratory stage, such as gait, ear, and DNA.

1.3.1 Biometric Parameters

Here we discuss some of the most relevant parameters likely to be encountered in literature relating to biometric technology.

FAR or the false accept rate is the fraction of impostors that are passed in error as genuine by a biometric system.

FRR or the false reject rate is the fraction of genuine applicants that are rejected in error by a biometric system.

EER or the equal error rate is the rate of either FAR or FRR when both rates are equal.

ROC or receiver operator characteristic is a parametric plot of FAR against FRR with a view to adjusting the system threshold to meet a given security policy.

Attributes are the physiological and behavioural characteristics that can be used as biometrics. Physiological examples include weight, height, body shape, face shape, hand shape, skin colour and texture, odour, hair colour and style, limb proportions, retina, iris, DNA, fingerprint, and of course ear shape. Behavioural examples include gait, body posture, gestures (including hand,

arm, and shoulder gestures), speech, and handwriting, keyboard typing pattern, heartbeat, respiration pattern, and eye blinking pattern.

Universality is the extent to which the biometric is readily collectible across the population as a whole. Ears are sometimes covered by hair especially with women. Or they might be partly obscured by jewelry, but we could equally say that eyes might be obscured by spectacles or sunglasses.

Permanence refers to the extent to which the attribute does not change over time. The fingerprint and the ear are very persistent, hardly changing at all throughout adulthood. The face and the iris are much less persistent, one changing with expression and the other responding to changes in ambient lighting. Other less persistent biometrics include speech, hairstyle and colour, odour, weight, and profile.

Acceptability refers to the degree of intrusion or inconvenience caused to the user. Many users will object on hygiene grounds to contact based biometrics such as fingerprint and hand. Scanning the iris and retina may cause anxiety. Users will become impatient if the process takes more than a few seconds.

Collectability refers to ease of acquisition. The retina requires considerable cooperation because the subject is usually required to look through an eyepiece at a target. On the other hand gait, face, and ear measurements can be collected more easily from a greater distance. Some biometrics such as keyboard typing patterns can be collected covertly.

Uniqueness is how unlikely a description is to occur more than once in the population. The iris is believed to be one of the most unique biometrics. As mentioned with fingerprints, the degree of uniqueness will also depend on the number of fingers being considered.

Circumventability is the ease with which the system can be fooled for example by holding a picture of someone else's face or ear in front of your

own. Counter-measures include vitality testing, for example it is known that the pupil diameter undergoes small "hippus" oscillations [47] once or twice per second, even under uniform lighting. It is unlikely that ears could be tested for liveness, unlike the iris or gait. This might restrict use in high-security applications.

1.3.2 Biometric Technologies

Here we review a selection of current biometric technologies together with a small selection of those that are still at the research stage.

Face

One of the least obtrusive and most collectible biometrics is the human face; we are used to having our faces compared with our photographs in passports and identity cards [49]. However, the face changes considerably from birth to old age, and even ephemerally due to make-up and expression changes. Even though strictly not part of the face, alterations in hairstyle can radically change the appearance of the face.

Fingerprint

Because fingerprints have been used extensively in forensic science for a long time, it is not surprising that this is also one of the most mature biometrics in use today [50]. Automatic fingerprint identification systems are available both for forensic science and biometrics applications. Fingerprints have very high persistence but have an associated criminal stigma and can be perceived to be unhygienic, as contact with a sensor is usually required. Systems based on both the global appearance and minutiae of the friction ridges have been developed and laser-scanning systems are also in use.

Voice

Speech recognition and speech synthesis are two enduring areas of research, and this has naturally resulted in speech-based biometrics. There is a choice here whether the aim is to recognise characteristics of an individual voice or whether to base it on chosen pre-arranged words [51]. Speech would clearly be one of the most natural and unobtrusive biometrics but is easily defeated by mimicking and using voice recordings.

Hand

The geometry of the hand can be used as a biometric by capturing its shape in the form of a silhouette [52]. The user is required to place the hand in a receptacle that has pegs to separate the fingers. Also in use are systems that produce a thermogram of the back of the hand using the vein pattern as the biometric [53].

Retina

The vein pattern in the retina at the back of the eye can be used for identification but requires considerable cooperation on the part of the subject who has to look through an eyepiece and furthermore is required to look at a particular location in the visual field [54]. However, to offset this the signature is believed to be highly unique.

Iris

The iris, which gives the eye its colour, would at first seem to be an unlikely choice as a biometric. Closer inspection reveals considerable texture detail in the form of modulation of a basic radial pattern of tiny muscles that cause the pupil to adjust its aperture in response to ambient light variations. Indeed, Daugman and Downing use demodulation to extract an iris signature, which is believed to be one of the most unique biometric signatures available [55,56].

Gait

Gait is still in the research stages but is very promising especially as it holds out the prospect of being one of the most remotely collectible signatures. This would clearly be very advantageous in surveillance situations where known criminals could be automatically detected even if they attempt to disguise their appearance [57]

Ear

There are no known working ear based biometric systems in use at this time, however preliminary research has been carried out by Burge and Burger on its

viability as a biometric [9]. There are rumours that the US armed services use ears for recognition, but this has not been confirmed.

Odour

Biometrics based body odour is still very much at the research stage [58]. Research suggests that body odour can indeed be used for identification [59], but problems can be expected due to use of perfumes and deodorants and changes in diet and health.

Hand Writing

The written signature is an obvious candidate as a behavioral biometric since it is readily collected and the user will need very little practice [60]. Also, in addition to the geometry of the signature, the sound patterns and velocity and acceleration of the stroke can be incorporated in the feature set, thereby making the technique more immune to forgery.

DNA

DNA is believed to be one of the most reliable and positive forms of personal identification, although identical twins share the same sequence [40]. At the moment its use for identification is confined to forensic science where it is believed that it might eventually replace fingerprinting [10]. At this time the procedure is too slow to be of use for biometrics and is very intrusive.

2 The Force Field Transform

This chapter provides the mathematical foundation for the new transforms and establishes some useful properties. The transforms are defined and are shown to be linear transformations and to be invertible under certain circumstances. The basic concepts underpinning the transforms and the mathematics used to describe them can be found in various introductory works on physics [61,62] and electromagnetics [63,64].

The energy field equations are first introduced, and then it is shown how a potential energy surface is composed of a summation of elementary potential functions corresponding to isolated pixels. The energy surface is then used to define potential wells and channels. An explanation is offered as to why the energy surface has an underlying dome shape, which in turn gives the force field an advantageous centric property. It is shown that force can also be viewed as the gradient of energy, thus allowing the force field to be calculated by differentiating the energy field, and also allows some properties established about one to be generalised to the other. Although the fields can be derived by direct application of the defining equations, it is shown that treating the process as a convolution, and using the Convolution Theorem to perform the calculation in the frequency domain, can gain a considerable speed advantage. The question of transform invertibility is considered before rounding off the section by establishing some invariant properties and analysing sensitivity to noise.

2.1 Potential Energy Transform Definition

The image is transformed by treating the pixels as an array of N particles that act as the source of a Gaussian potential energy field. It is assumed that there is a spherically symmetrical potential energy field surrounding each pixel, where $E_i(\mathbf{r}_j)$ is the potential energy imparted to a pixel of unit intensity at the pixel location with position vector \mathbf{r}_j by the energy field of a remote pixel with position vector \mathbf{r}_i and pixel intensity $P(\mathbf{r}_i)$, and is given by

$$E_i(\mathbf{r}_j) = \frac{P(\mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (1)$$

where the units of pixel intensity, energy, and distance are arbitrary, as are the co-ordinates of the origin of the field. It should be noted that the energy field is notional and is not intended to model the propagation of light or anything else. If an exploratory unit intensity test pixel is moved around in the energy field generated by a given pixel, energy will be exchanged if the net effect is to change the distance of the test pixel from the given pixel. Thus the field consists of concentric rings of equal potential energy known as equipotentials. If the test pixel moves to a different location on the same equipotential ring, no energy is exchanged. If it moves to a different equipotential, an amount of energy will be exchanged equal to the difference in energy between the two rings.

2.2 Potential Energy Function

The potential energy function of a single isolated pixel looks like the shape shown in Figure 3.

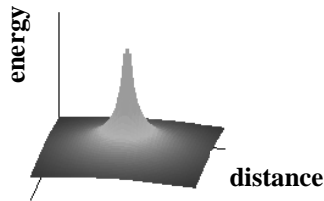


Figure 3 Potential function of an isolated pixel

Horizontal cross sections of the potential function correspond to equipotential rings and vertical cross sections correspond to the 2-dimensional double-sided $1/r$ inverse function shape.

2.3 Potential Energy Surface

Now to find the total potential energy at a particular pixel location in the image, the scalar sum is taken of the values of the overlapping potential energy functions of all the image pixels at that precise location and is given by

$$E(\mathbf{r}_j) = \sum_{i=0}^{N-1} E_i(\mathbf{r}_j) = \sum_{i=0}^{N-1} \frac{P(\mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (2)$$

This summation is then carried out at each image pixel location to generate the complete transformation, which is a smoothly varying surface due to the fact that the underlying potential functions have smooth surfaces. The result of this process for the energy transform for an ear image is shown in Figure 4, where the same surface has been depicted from a variety of different perspectives below the lobe.

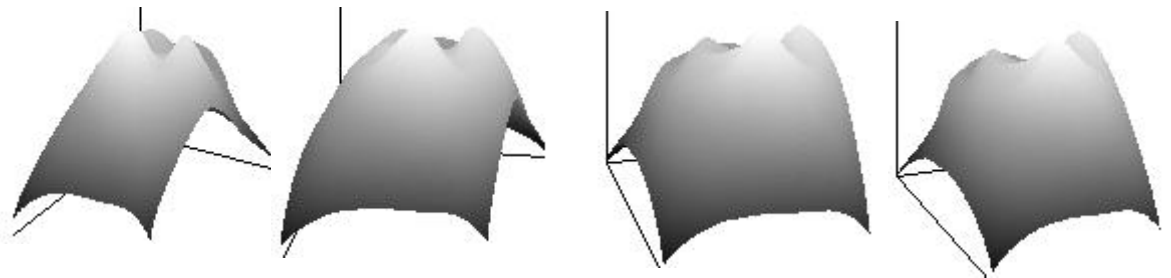


Figure 4 Energy surface for an ear viewed from below the lobe

2.4 Potential Wells and Channels

The potential surface undulates, forming local peaks or maxima, with ridges leading into them. These peaks we will call potential energy wells since, by way of analogy, if the surface were to be inverted and water poured over it, the

peaks would correspond to small pockets where water would collect. Notice that the highest of the three obvious peaks in Figure 4 has a ridge that slopes gently towards it from the smaller peak to its left. This corresponds to a potential energy channel, because to extend the analogy, water that happened to find its way into its inverted form would gradually flow along the channel towards the peak.

2.5 Dome Shape Explanation

It will be noticed that the energy surface has an elongated dome shape, which is modulated by the channels and wells; it looks like a small mountain with a few peaks and ridges. The reason for the dome shape can be easily understood by considering the case where the image has just one grey level throughout. In this situation, the energy field at the centre would have the greatest energy share because test pixels at that position would have the shortest average distance between themselves and all the other pixels, whereas test pixels at the edges would have the greatest average distance to all the other pixels, and therefore the least total energy imparted to them. Because the transform is linear the energy surface can be considered to be a sum of two components, corresponding to the mean value and variation of the image. This is illustrated in Figure 5.

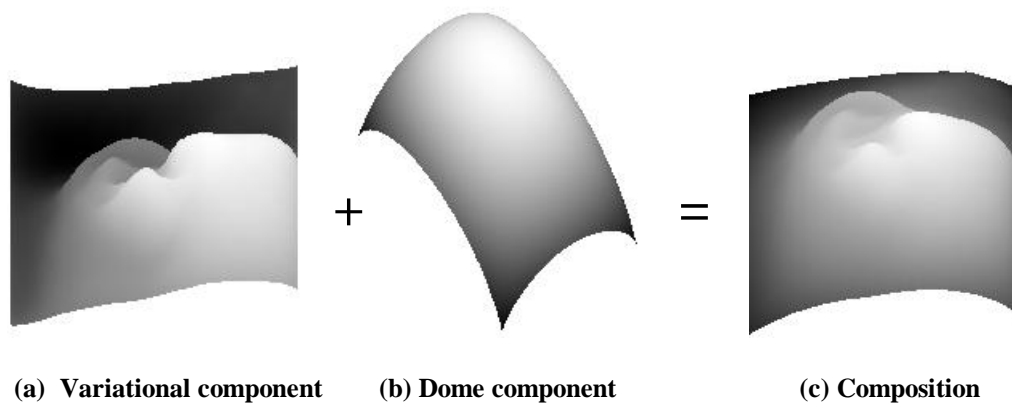


Figure 5 Energy surface as a sum of components

It is the mean value component that gives the energy surface its dome shape and the variational component that modulates this basic shape causing peaks

and ridges. The dome shape leads to automatic feature extraction, since by way of the water analogy once again, water when introduced at the edge of the dome shape would always flow towards the centre, finding its way into the channels on the way, and eventually ending up in one or other of the wells.

2.6 Force as Gradient of Potential

Associated with the scalar energy field there is a vector force field and the fields are related by the fact that the force at a given point is equal to the additive inverse of the gradient of the potential energy surface at that point. This relationship, shown in Equation 3, allows the force field to be easily calculated by differentiating the energy field, and allows some conclusions drawn about one field to be extended to the other.

$$\mathbf{F}(\mathbf{r}) = -grad(E(\mathbf{r})) = -\nabla E(\mathbf{r}) \quad (3)$$

The derivation of this equation can be found in [63] and [64].

2.7 Force Field Transform Definition

The force field can also be defined directly with its own set of equations. The defining equations are more complicated than those of the energy field but the concept is more intuitive. The image is transformed by treating the pixels as an array of N mutually attracting particles that act as the source of a Gaussian force field. In a similar way to Newton's Law of Universal Gravitation, the pixels are considered to attract each other according to the product of their intensities and inversely to the square of the distances between them. Each pixel is assumed to generate a spherically symmetrical force field so that the force $\mathbf{F}_i(\mathbf{r}_j)$ exerted on a pixel of unit intensity at the pixel location with position vector \mathbf{r}_j by a remote pixel with position vector \mathbf{r}_i and pixel intensity $P(\mathbf{r}_i)$ is given by

$$\mathbf{F}_i(\mathbf{r}_j) = P(\mathbf{r}_i) \frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3} \quad (4)$$

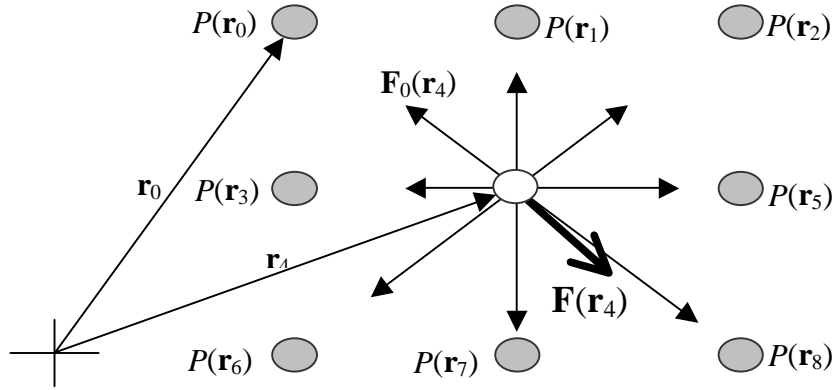


Figure 6 Force field calculation at the centre of a 3x3 pixel image

The total force $\mathbf{F}(\mathbf{r}_j)$ exerted on a pixel of unit intensity at the pixel location with position vector \mathbf{r}_j is the vector sum of all the forces due to the other pixels in the image and is given by,

$$\mathbf{F}(\mathbf{r}_j) = \sum_{i=0}^{N-1} \mathbf{F}_i(\mathbf{r}_j) = \sum_{i=0}^{N-1} \left(P(\mathbf{r}_i) \frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3} \right) \quad (5)$$

This calculation is illustrated graphically in Figure 6 for the total force acting on a unit value test pixel at the centre of a simple 3×3 pixel image. The actual pixel $P(\mathbf{r}_4)$ is replaced with a unit value test pixel, which is depicted in white. The test pixel at the centre of the group is attracted by eight surrounding pixels whose forces are depicted by light arrows. Pixel $P(\mathbf{r}_8)$ exerts the greatest force as indicated by the longest arrow, whereas pixel $P(\mathbf{r}_3)$ exerts the smallest force as indicated by the shortest arrow. The total force acting on the test pixel is the vector sum of the forces exerted by the eight surrounding pixels and is depicted in the diagram by a heavy arrow. In order to calculate the force field for the entire image, Equation 5 should be applied at every pixel position in the image.

2.8 Force Field Transform Linearity

It will now be shown that the force field transform is a linear transformation. One way of doing this is to show that the transformation has an equivalent matrix equation since linear transformations between finite-dimensional vector spaces are precisely those transformations that have a matrix representation [65,66,67]. It could also be argued that the transformation is linear since the force acting on a unit text pixel is defined as a vector summation, which is a linear operation. However, the matrix representation provides further insight and may be used to investigate whether the transformation is invertible or not.

The matrix representation of the transform will be found by considering the case of a simple 4-pixel image as depicted in Figure 7.

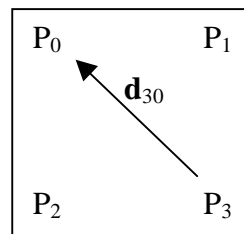


Figure 7 Geometry of a simple 4-pixel image

Equation 5 is applied at each pixel location to determine the net force acting at each of those positions. Here, the force acting at each location is the sum of the three forces exerted by the other three pixels. For example the net force $\mathbf{F}(\mathbf{r}_0)$ acting at the location of pixel P_0 is given by,

$$\mathbf{F}(\mathbf{r}_0) = \mathbf{d}_{01}P(\mathbf{r}_1) + \mathbf{d}_{02}P(\mathbf{r}_2) + \mathbf{d}_{03}P(\mathbf{r}_3) \quad (6)$$

where the coefficients \mathbf{d}_{ji} correspond to the inverse square displacement vectors given by,

$$\mathbf{d}_{ji} = \frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3} \quad (7)$$

Now this sum can be expressed as the inner product of a row vector of inverse square displacement vectors, times a column vector of pixel intensities given by,

$$(0 \quad \mathbf{d}_{01} \quad \mathbf{d}_{02} \quad \mathbf{d}_{03}) \begin{pmatrix} P(\mathbf{r}_0) \\ P(\mathbf{r}_1) \\ P(\mathbf{r}_2) \\ P(\mathbf{r}_3) \end{pmatrix} = \mathbf{d}_{01}P(\mathbf{r}_1) + \mathbf{d}_{02}P(\mathbf{r}_2) + \mathbf{d}_{03}P(\mathbf{r}_3) = \mathbf{F}(\mathbf{r}_0) \quad (8)$$

where the first element of the row vector is zero, to ensure that $P(\mathbf{r}_0)$ makes no contribution.

Now the net force acting at each of the other three pixel positions can also be expressed as inner products and if their row vectors are stacked to form a matrix, the overall force field calculation can be expressed as a matrix equation as follows,

$$\begin{pmatrix} 0 & \mathbf{d}_{01} & \mathbf{d}_{02} & \mathbf{d}_{03} \\ \mathbf{d}_{10} & 0 & \mathbf{d}_{12} & \mathbf{d}_{13} \\ \mathbf{d}_{20} & \mathbf{d}_{21} & 0 & \mathbf{d}_{23} \\ \mathbf{d}_{30} & \mathbf{d}_{31} & \mathbf{d}_{32} & 0 \end{pmatrix} \begin{pmatrix} P(\mathbf{r}_0) \\ P(\mathbf{r}_1) \\ P(\mathbf{r}_2) \\ P(\mathbf{r}_3) \end{pmatrix} = \begin{pmatrix} \mathbf{F}(\mathbf{r}_0) \\ \mathbf{F}(\mathbf{r}_1) \\ \mathbf{F}(\mathbf{r}_2) \\ \mathbf{F}(\mathbf{r}_3) \end{pmatrix} \quad (9)$$

The leading diagonal of zeros reflects the fact that no pixel attracts itself. This is a skew-symmetric matrix, of the form

$$A = -A^T \quad (10)$$

where the skew symmetry is accounted for by the fact that we are dealing with a fully connected network but with a pair of equally weighted directed edges connecting every pair of nodes.

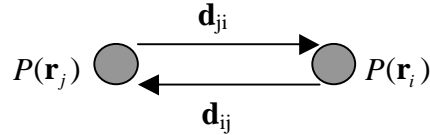


Figure 8 Skew symmetry is due to mutual attraction between pixels

It is easy to see that this reasoning can be applied to an image of any size and therefore it follows that the transformation is linear since it has a matrix representation.

2.9 Potential Transform Linearity

Because of the simple relationship of Equation 3, or by similar reasoning to that given in Section 2.8, the scalar potential energy transform is also a linear transformation. The matrix representation of the transform corresponding to the case of the simple 4-pixel image is given by,

$$\begin{pmatrix} 0 & d_{01} & d_{02} & d_{03} \\ d_{10} & 0 & d_{12} & d_{13} \\ d_{20} & d_{21} & 0 & d_{23} \\ d_{30} & d_{31} & d_{32} & 0 \end{pmatrix} \begin{pmatrix} P(\mathbf{r}_0) \\ P(\mathbf{r}_1) \\ P(\mathbf{r}_2) \\ P(\mathbf{r}_3) \end{pmatrix} = \begin{pmatrix} E(\mathbf{r}_0) \\ E(\mathbf{r}_1) \\ E(\mathbf{r}_2) \\ E(\mathbf{r}_3) \end{pmatrix} \quad (11)$$

where the coefficients d_{ji} correspond to the inverse distance scalars given by,

$$d_{ji} = \frac{1}{|\mathbf{r}_j - \mathbf{r}_i|} \quad (12)$$

This is a symmetric matrix with a leading diagonal consisting entirely of zeros but the matrix elements are now scalars instead of vectors.

2.10 Force Field Computation by Convolution

The matrix representation provides a useful insight into the nature of the transformation but would not be used in practice to calculate the force, as it would be very inefficient to say nothing of the storage requirements. In practice the defining equations would be applied directly, or for even greater efficiency, the process can be treated as a convolution of the image with the force field corresponding to a unit value test pixel and then invoking the Convolution Theorem to perform the calculation as a multiplication in the frequency domain, the result of which is then transformed back into the space domain. The force field equation for an $M \times N$ pixel image becomes

$$forcefield = \sqrt{M \times N} \mathfrak{I}^{-1}[\mathfrak{I}(unit\ forcefield) \times \mathfrak{I}(image)] \quad (13)$$

where \mathfrak{I} stands for the Fourier Transform and \mathfrak{I}^{-1} for its inverse. The usual care must be taken to ensure that dimensions of the unit sample force field are twice those of the image dimensions and that sufficient zero padding is used to avoid aliasing effects.

The *Mathcad* code fragment in Figure 9 shows the convolution method for computing the force field associated with an $M \times N$ pixel image. The overall strategy is to directly compute only the force field associated with a unit value test pixel and then to use this to calculate the field associated with each image pixel. Simply scaling it by the pixel value and centring it on the pixel position achieves this. The overall force field is then simply the sum of all the overlapping force fields.

This calculation amounts to the convolution of a $2N-1 \times 2M-1$ unit value force field top-left aligned in a $3N-2 \times 3M-2$ field of zeros, with the image also top-left aligned in such a field. The Convolution Theorem is employed so that rather than actually performing convolution in the space domain, the calculation is handled as a multiplication in the frequency domain, and the answer transformed back into the space domain.

Lines 1 to 7 compute **upf**, the force field associated with a unit value test pixel placed in the centre of a $2N-1 \times 2M-1$ field of zeros.

Line 8 pads **upf** with zeros so that it becomes located at the top left corner of a $3N-2 \times 3M-2$ field of zeroes.

Lines 9 and 10 place the image **pic** at the top left corner of a $3N-2 \times 3M-2$ field of zeroes, **inp**.

Line 11 performs the transform, multiply, and invert cycle in accordance with $\mathbf{oup} = \sqrt{MN} \mathfrak{F}^{-1}[\mathfrak{F}(\mathbf{upf}) \times \mathfrak{F}(\mathbf{inp})]$ where \mathfrak{F} stands for Fast Fourier Transform.

Line 12 finally extracts the $M \times N$ required force field from the centre of the $3N-2 \times 3M-2$ result.

```
ff(pic) := | sr ← 2·(rows(pic) - 1)
            | sc ← 2·(cols(pic) - 1)
            | r ← rows(pic) - 1
            | c ← cols(pic) - 1
            | for rr ∈ 0..sr
            |   for cc ∈ 0..sc
            |     upfrr,cc ←  $\frac{(r + c \cdot 1j) - (rr + cc \cdot 1j) + 0j}{(|r + c \cdot 1j - (rr + cc \cdot 1j)|)^3}$ 
            | upf3·rows(pic)-3, 3·cols(pic)-3 ← 0
            | inp ← pic
            | inp3·rows(pic)-3, 3·cols(pic)-3 ← 0
            | oup ←  $\sqrt{\text{rows(inp)} \cdot \text{cols(inp)}} \cdot \text{icfft}(\overrightarrow{(\text{cfft(upf)} \cdot \text{cfft(inp))})})$ 
            | ff ← submatrix(oup, r, 2·r, c, 2·c)
```

Figure 9 Force field by convolution in Mathcad

2.11 Invertibility

The application of Equation 2 at each of the N pixel positions leads to a system of N equations in N unknowns. Now if the N equations are linearly independent, then it follows that the system of equations can be solved for the pixel values, given the energy values. In other words, the transform would be invertible. It may not be so obvious why there should be a need to invert the transform, but invertibility is an important property in its own right, and worth investigating. One useful advantage that accrues if the transform proves to be invertible is that the original image can be completely recovered from the energy surface, thus establishing that the transform is information preserving.

2.11.1 Numerical Testing for Invertibility

There are a number of different but equivalent tests that can be applied to determine whether or not the transform is invertible, or equivalently, if the corresponding representation matrix is invertible. For a square matrix A the following statements are equivalent [67]: -

1. The matrix A is invertible.
2. The rows and columns of A are linearly independent.
3. The determinant of A is non-zero.
4. All the eigenvalues of A are non-zero.
5. $A^T A$ is positive definite.

We will only consider the question of the invertibility of the energy transforms and not of the force field transforms because the representation matrices for the force field transforms are skew symmetric, and all odd skew symmetric matrices are singular. To see this for an $N \times N$ skew symmetric matrix, notice that

$$\det(A) = \det(-A^T) = (-1)^N \det(A) \quad (14)$$

Now when N is odd, $\det(A) = -\det(A)$ so that $\det(A) = 0$, making A singular.

Statement 2 says that the transform will be invertible if the system of N equations in N unknowns is linearly independent, which is what we have

already considered. Now while it might not be a proof, it can at least be said that the rows of the matrix were derived independently in the sense that none was made up from a combination of the others; each of the equations was derived by the application of Equation 2 at its own unique pixel position.

Statements 3 allows us to test numerically for the invertibility of the transforms corresponding to particular image sizes. As an example, Figure 10 demonstrates transform invertibility by showing that the energy transform representation matrix of a simple 2×2 pixel image can be inverted. This is done by showing that the product of the matrix with its inverse yields the corresponding identity matrix.

$$\left[\begin{array}{c} \text{Inverse of Matrix} \\ \frac{-1}{4} \cdot \begin{bmatrix} 2\sqrt{2} & -2 & -2 & 0 \\ -2 & 2\sqrt{2} & 0 & -2 \\ -2 & 0 & 2\sqrt{2} & -2 \\ 0 & -2 & -2 & 2\sqrt{2} \end{bmatrix} \end{array} \right] \cdot \left[\begin{array}{c} \text{Matrix} \\ \begin{bmatrix} 0 & 1 & 1 & \sqrt{2} \\ 1 & 0 & \sqrt{2} & 1 \\ 1 & \sqrt{2} & 0 & 1 \\ \sqrt{2} & 1 & 1 & 0 \end{bmatrix} \end{array} \right] = \left[\begin{array}{c} \text{Identity Matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \right] \quad (15)$$

Figure 10 Energy transform matrix inversion for 2×2 image

Exploiting statement 3, a numerical test for invertibility can be carried out by constructing the appropriate matrix and computing the determinant. If the determinant is non-zero the matrix is non-singular, and therefore invertible. All the determinants of matrices corresponding to the sequence of square images ranging from 2×2 pixels to 33×33 pixels have been computed and have been found to be non-zero. It has also been verified that all non-square image formats up to 7×8 pixels have associated non-singular matrices.

Table 1 shows the determinants of the energy transform representation matrices corresponding to a sequence of square images ranging from 2×2 pixels to 33×33 pixels illustrating that since these values are non-zero the associated matrices and hence the corresponding transforms are invertible. Table 2 shows the result for the matrices corresponding to the first 8 non-square images, indicating that these also are invertible. Some of the values are

rather small, but well above the resolution of the arithmetic system used in their calculation.

SIZE	DETERMINANT		SIZE	DETERMINANT
2x2	-1.75		18x18	-0.00136
3x3	0.7951		19x19	-0.0008128
4x4	0.3864		20x20	0.09267
5x5	-1.102		21x21	-0.00569
6x6	0.4426		22x22	0.000561
7x7	0.1995		23x23	7.198e-006
8x8	0.0002739		24x24	-0.03551
9x9	1.628		25x25	0.05521
10x10	0.4425		26x26	7.446
11x11	-0.4079		27x27	-4.993e-006
12x12	-0.00225		28x28	-0.00442
13x13	0.8954		29x29	0.1265
14x14	-0.4363		30x30	-1.364
15x15	-1.465		31x31	-0.4055
16x16	0.1187		32x32	-1.104e-008
17x17	-0.00876		33x33	-0.00749

Table 1 Determinants corresponding to the first 32 square images

		ROWS						
		1	2	3	4	5	6	7
C O L U M N S	1	1	-1	1	-0.15972	-0.61227	0.64451	-0.06688
	2	-1	-1.75	0.1673	1.54061	0.28055	-1.35533	-0.75388
	3	1	0.1673	0.79512	-1.21593	-0.29127	-0.59947	-0.49042
	4	-0.15972	1.54061	-1.21593	0.38635	-1.30503	-0.29998	-0.61057
	5	-0.61227	0.28055	-0.29127	-1.30503	-1.10237	-0.03096	1.32986
	6	0.64451	-1.35533	-0.59947	-0.29998	-0.03096	0.44255	-0.03878
	7	-0.06688	-0.75388	-0.49042	-0.61057	1.32986	-0.03878	0.19951
	8	-0.46868	0.9216	0.80471	-0.43568	0.25514	-0.0832	-0.08339

Table 2 Determinants corresponding to the first 8 non-square images

2.11.2 Condition Number and Invertibility

In numerical analysis the *condition number* is a measure of the stability of a computational problem. For example the *relative condition number* for the problem of evaluating the real function $f(x)$ when a perturbation $x + \delta x$ is introduced is given by the ratio of the relative changes in the function and the dependent variable as given by

$$\text{Relative Condition Number} = \left| \frac{f(x + \delta x) - f(x)}{f(x)} \right| \bigg/ \left| \frac{\delta x}{x} \right| \quad (16)$$

A large condition number indicates that the problem is not well behaved when it is said to be *ill-conditioned*. More generally the condition number of a linear system $Ax = b$ is given in terms of the norm of the matrix A and its inverse, as given by

$$\text{Condition Number} = \|A\| \times \|A^{-1}\| \quad (17)$$

A matrix is ill-conditioned if the condition number is too large, and singular if it is infinite. If the condition number is one, the matrix is said to be perfectly conditioned. In practical terms if one row of a matrix is almost but not quite linearly dependent on another, then the matrix will be ill-conditioned. For example using the function $\text{Cond}(\phi)$ based on the L1 norm we have

$$\begin{aligned} \text{Cond} \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} &= 20 \\ \text{Cond} \begin{bmatrix} 1 & 2 \\ 1 & 2.01 \end{bmatrix} &= 1207.01 \\ \text{Cond} \begin{bmatrix} 1 & 2 \\ 1 & 2.001 \end{bmatrix} &= 12007.001 \end{aligned} \quad (18)$$

We see that the condition number grows larger as the second row becomes more linearly dependent on the first row. In the limit the condition number tends to infinity, indicating a singular matrix. Even though the condition number is defined in terms of the matrix inverse, it can be estimated without calculating the inverse and due to the inherent instability in calculating the determinant, should prove to be a better test for invertibility.

2.11.3 Invertibility Conclusions

Notwithstanding questions of machine accuracy, these results suggest that the energy transform is indeed invertible for most image sizes and aspect ratios. Since the transform is invertible it may be concluded that the original image is in principle completely recoverable from the potential energy surface. This is an important result from an information theoretic point of view, because it implies that the transformed image conserves all the information contained in the original image. Even if there are some particular combinations of aspect ratio and size that yield singular matrices, this should not detract from the overall conclusion that the transformation preserves information.

It would be nice to provide a general proof of invertibility based on the equivalent statements of invertibility, however such a proof has thus far remained elusive, and the search must be postponed to further work. We must also remember that it simply may not be true that these matrices are invertible in general, and hence a proof will be impossible. Of course, a proof to this effect would also be nice.

2.12 Invariant Properties

For the approach to be of practical use, we require to show that the resulting force and energy fields are both illumination invariant and scale invariant. By analysing noise sensitivity, we show that a major advantage of the technique is its considerable noise tolerance. By the nature of the underlying equations, we can analyse the properties for the energy field and the results generalise naturally to force.

2.12.1 Illumination Invariance

Should the image intensity be scaled by a factor k then we have that

$$\begin{aligned} E(\mathbf{r}_j) &= \sum_{i \in 0, N-1 | i \neq j} E_i(\mathbf{r}_j) = \sum_{i \in 0, N-1 | i \neq j} \frac{kP(\mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{r}_j|} \\ &= k \sum_{i \in 0, N-1 | i \neq j} \frac{P(\mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{r}_j|} \end{aligned} \quad (19)$$

We see that scaling the image intensity does not alter the resulting force field distribution but, as expected, only scales the energy intensity by the same factor. As such, the resulting force and potential fields are invariant under multiplicative changes in illumination.

Similarly, change in overall brightness by addition gives

$$\begin{aligned} E(\mathbf{r}_j) &= \sum_{i \in 0, N-1 | i \neq j} E_i(\mathbf{r}_j) = \sum_{i \in 0, N-1 | i \neq j} \frac{P(\mathbf{r}_i) + k}{|\mathbf{r}_i - \mathbf{r}_j|} \\ &= \sum_{i \in 0, N-1 | i \neq j} \frac{P(\mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i \in 0, N-1 | i \neq j} \frac{k}{|\mathbf{r}_i - \mathbf{r}_j|} \end{aligned} \quad (20)$$

This suggests that addition of k brightness levels to each pixel results in an increment to the steepness of the dome shaped surface component shown in Figure 5(b). This will have the effect of a slight non-linear scale stretching of the feature distribution, but otherwise will preserve its form.

2.12.2 Scale Invariance

Should the image size be scaled by a factor s then we have that

$$\begin{aligned} E(s\mathbf{r}_j) &= \sum_{i \in 0, N-1 | i \neq j} E_i(s\mathbf{r}_j) = \sum_{i \in 0, N-1 | i \neq j} \frac{P(s\mathbf{r}_i)}{|s\mathbf{r}_i - s\mathbf{r}_j|} \\ &= \frac{1}{s} \sum_{i \in 0, N-1 | i \neq j} \frac{P(s\mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{r}_j|} \end{aligned} \quad (21)$$

We see that scaling the image distribution not only results in a scaled force field distribution but, as expected, larger distances imply weaker forces, as the denominator is proportional to scale. As such, the resulting force and potential fields are scale invariant.

2.12.3 Noise Tolerance

Next we seek to demonstrate tolerance to noise. Given that the image is corrupted by additive zero-mean Gaussian noise v , for field point \mathbf{r}_j we have that

$$\begin{aligned} E(\mathbf{r}_j) &= \sum_{i \in 0, N-1 | i \neq j} E_i(\mathbf{r}_j) = \sum_{i \in 0, N-1 | i \neq j} \frac{P(\mathbf{r}_i) + v_i}{|\mathbf{r}_i - \mathbf{r}_j|} \\ &= \sum_{i \in 0, N-1 | i \neq j} \frac{P(\mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i \in 0, N-1 | i \neq j} \frac{v_i}{|\mathbf{r}_i - \mathbf{r}_j|} \end{aligned} \quad (22)$$

We see that the resulting energy at field point \mathbf{r}_j has two terms, one is the uncorrupted energy and the other is the weighted summation of all the noise in the image. However, the magnitude of the noise term with respect to the energy term is diminished due to the cancellation effect of zero-mean noise summation. Thus the technique does enjoy very good noise tolerance, due to the powerful averaging inherent in the transformation. Note that non-zero mean noise can be considered as zero-mean noise with a deterministic offset, without loss of generality.

3 Force Field Feature Extraction

The force field and energy field transforms have been defined and some of their properties analysed. They have been shown to be linear transformations and the energy transform has been shown to be invertible for a variety of image sizes and shapes, which leads us to the important conclusion that the transforms are information preserving. The energy transform of an ear image takes the form of a smooth surface consisting of just a few wells with channels converging upon them. Because the transform is information preserving, we know that this smooth surface contains all the information of the original image. It is therefore argued that these channels and wells, since the surface is otherwise smooth, provide much of the descriptive information about the surface, and as such make good ear features. Therefore in this chapter we will show how the directional properties of the force field can be exploited to automatically locate these channels and wells, which will then form the basis of the new ear features.

Two distinct methods of force field feature extraction are presented, each with its own advantages and peculiarities. The more intuitive of the two is an algorithmic approach based on field lines, where notional test pixels trace out field lines, which flow across the force field, discovering its structure in the form of channels and wells. In this approach wells and channels appear to be distinct entities, and indeed in the preliminary stages of this research, it was hoped that wells alone would suffice as compact ear features. However, this proved not to be the case and the reasons will be discussed.

The second method is an analytical approach, which resulted from an analysis of the field line approach. The analysis revealed that the channels and wells correspond to ridges and peaks respectively in a mathematical function, showing that they are really separate manifestations of the same phenomenon. This function, which we call convergence, maps the force field to a scalar field where the output is the additive inverse of the divergence of the direction of the force field. The convergence function provides a more comprehensive description than just channels and wells because it includes negative ridges and peaks, which we will refer to as antichannels and antiwells respectively.

Even though the force field magnitude is not used in locating the new features, the magnitude has interesting properties itself. The force magnitude leads to a form of edge detection with powerful intrinsic smoothing, which we call homogeneous cancellation, and this will be described firstly before going on to describe feature extraction.

The concept of a unit value test pixel, which hitherto has only been used as a stationary field gauge of local energy and force, is now extended by allowing it to follow the direction of the force field, tracing out field lines, eventually forming channels and wells. A simple force field due to a small white cross is demonstrated and the field due to a larger version of the same cross is used to illustrate field line formation of channels and wells.

Field line feature extraction is demonstrated on ears and its advantages are considered. The form and location of the field line features is shown to be independent of the choice of test pixel starting positions, and the channels are also shown to scale in proportion to the image size. The features are shown to enjoy good noise tolerance due the high degree of intrinsic smoothing inherent in the transformation. The uniqueness of the features for a variety of different ears is considered, where the difficulty of using just wells as a description will be illustrated. How to actually compare channels and wells will be addressed in the next chapter on ear recognition.

3.1 Force Field Magnitude

The new features rely only on the directional properties of the force field; they do not take account of the force magnitude. This is because the exploratory test pixels, to be described in Section 3.2, always move in the direction of the local force, without regard to its magnitude. However, the force magnitude clearly also contains information about the image, so it is appropriate to consider this before going further.

Figure 11(b) shows the magnitude of a force field that has been generated from an ear image Figure 11(a). The result is still clearly recognisable as an ear, albeit an unusual one. The transformation appears to provide a remarkable degree of intrinsic smoothing and there also appears to be something akin to edge detection. The latter is attributable to a process that we call homogeneous cancellation where local forces are highly symmetrical in areas of constant pixel intensity and so tend to cancel. The inverse square nature of the field reduces the effect of forces away from the locality. An imbalance of symmetry in areas of rapid intensity change results in net forces that cause peaks in the magnitude response.

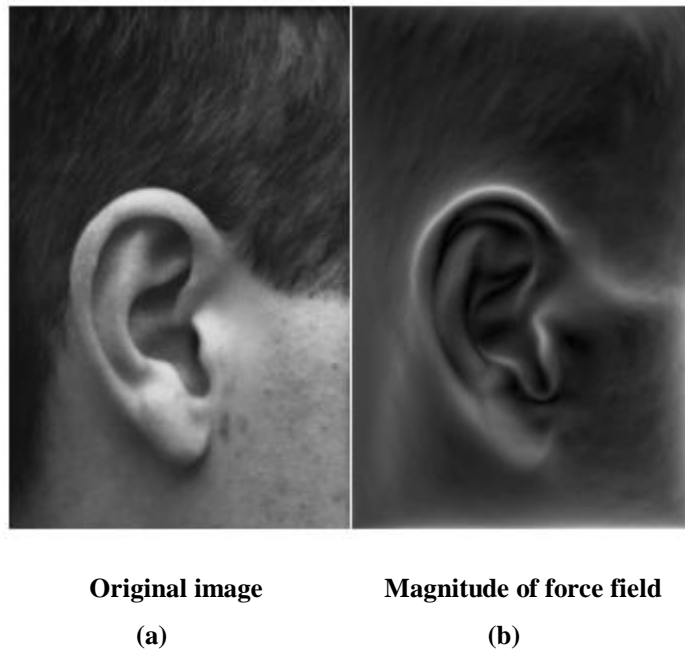


Figure 11 Force field magnitude for an ear image

Notice how this effect appears to result in a false groove, appearing as a dark band, along the length of the uppermost helix especially along the right hand side. Whereas the original image shows solid flesh, the transformed image exhibits a form of edge detection so that the outer and inner edges sandwich a dark band between them. Notice also that the two brightest spots in the original image of the ear to the left and right of the intertragic notch, just above the lobe, have become de-emphasised in the transformed image. The outline of this notch is also clearly depicted as a result of the transformation.

3.2 Exploratory Test Pixels

The concept of a unit value exploratory test pixel is introduced to assist with field line description. This idea is borrowed from Physics, where it is common to refer to unit value test particles when describing the Gaussian force fields associated with gravitational masses and electrostatic charges. When such notional test pixels are placed in a force field and allowed to follow the field direction their trajectories are said to form field lines. When this process is carried out with many different starting points a set of field lines will be generated that capture the general flow of the force field.

The test pixel positions are advanced in arbitrary increments of one unit of distance corresponding to a single pixel width, and test pixel locations are maintained as real numbers, producing a smoother trajectory than if they were constrained to occupy exact pixel grid locations. This effect of quantised versus continuous test pixel direction of movement is illustrated in Figure 12(a), which shows the effect of moving in only one of 8 possible directions, and in Figure 12(b), which shows the improvement in trajectory smoothness obtained if this restriction is removed.

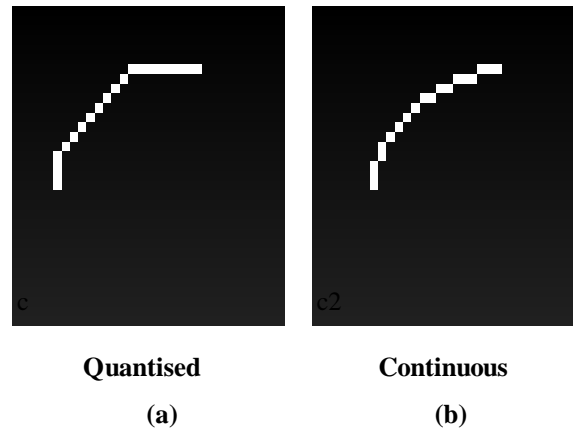


Figure 12 Effect of quantised versus continuous direction of test pixel movement

3.3 Field Lines

Figure 13(a) shows the force field corresponding to a small white cross-shaped pattern of pixels on a black background, shown in Figure 13(b). This is an 11×11 pixel image in which arrows depict the magnitude and direction of the forces at each of the 121 pixel positions in the image. Careful examination reveals that there is no force associated with the centre pixel. This is because the symmetry of the shape causes all forces at the centre of the image to be cancelled. The length of each arrow is proportional to the magnitude of the force, although the magnitude of the force is not actually used for our purpose, whereas the arrows point in the force direction. Notice that the forces align themselves along the field lines that in turn bend towards the arms of the cross where the pixels are located.

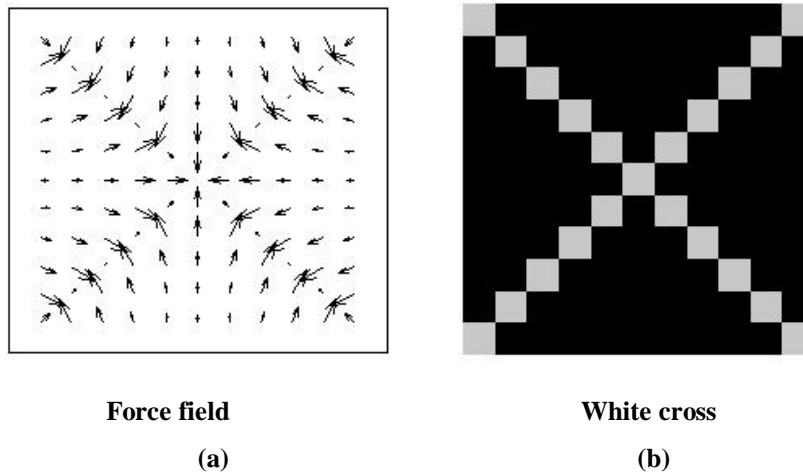


Figure 13 Force field due to a small white cross

3.4 Channel and Well Formation

An important property of field lines is that they never cross, for the simple reason that the force vector at any point in the field is unique. If two or more trajectories should happen to arrive at the same pixel location they will follow the same trajectory from that point onwards. As other trajectories meet this path then they will also follow it thus forming potential energy channels.

Now the force acting at any point in a force field is proportional to the gradient of the energy surface at that point; hence there is no force acting at the bottom of a potential well, since the gradient at that point is flat. Now the surface surrounding a well slopes towards the bottom, so that all the field lines surrounding a well slopes towards the bottom, so that all the field lines surrounding a well will all lead to the bottom. Thus a test pixel placed in a potential well will always work its way towards the bottom of the well, where it will be trapped and cannot escape. For this reason all the field lines will eventually terminate in one or other of the wells on an energy surface.

Channel and well formation is illustrated using a larger scale version of the image of a white cross on a black background; Figure 14(a) shows the image of the cross. Four channels can be seen forming along the arms of the cross, where Figure 14(b) shows the process at an early stage of formation and Figure 14(c) and (d) show an intermediate stage and the completed process. In this example, we see that every field line terminates in a single well at the centre of the force field. All but four of the field lines arrive at the well via one or other of the four channels. The field lines that arrive directly are those that start from the midpoints of the sides of the square and whose lateral forces cancel due to symmetry.

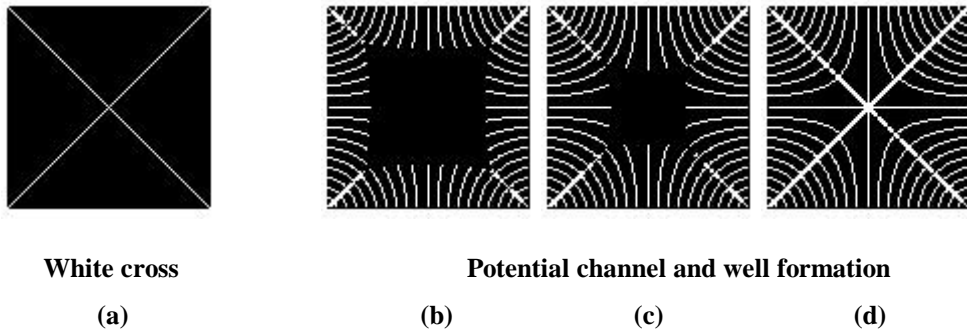


Figure 14 Formation of potential channels and wells for a cross shaped image

3.5 Channel and Well Formation for an Ear

Figure 15 demonstrates the field line approach to feature extraction for an ear image. A set of 50 test pixels has been arranged in an ellipse shaped array around the ear, and then iterated to generate a set of field lines. Even though the associated energy surface, shown in Figure 4, indicates the presence of three wells and one channel, it can be seen that field lines are more sensitive and have extracted four wells whose positions are indicated in Figure 15(c). The well locations have been extracted by observing clustering of test pixel co-ordinates so that Figure 15(c) is simply obtained by plotting the terminal positions of all the co-ordinates. Notice how field lines flow into channels and continue onwards until they terminate in wells. See how fourteen field lines cross the upper helix, and how each line joins a common channel that follows the curvature of the helix rightwards, finally terminating in a well.

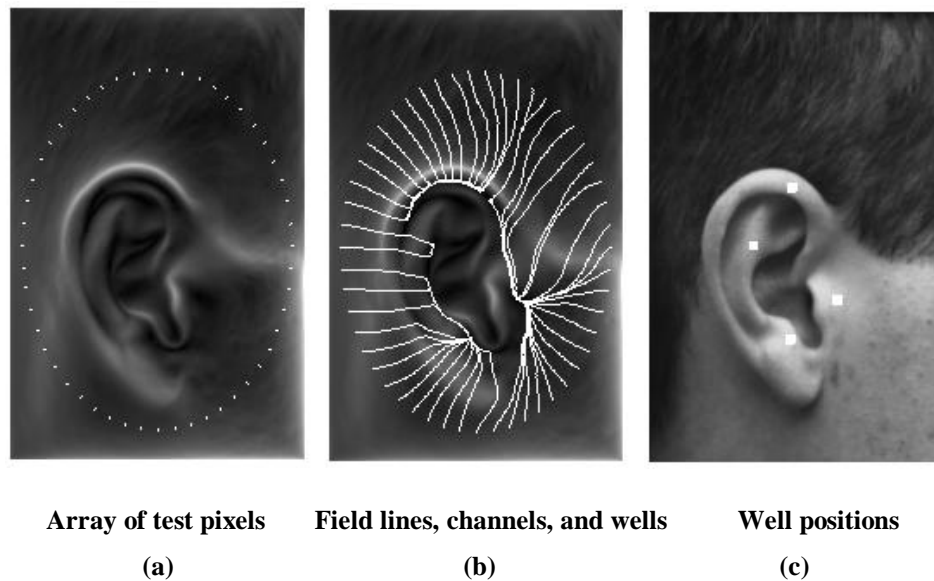


Figure 15 Field line, channel, and well formation for an ear.

Two isolated field lines flow in parallel from the left edge of the image to form a short channel terminating in a well at the left-centre of the ear. Thirteen field lines flow from the bottom left of the image to form a well at the bottom of the ear. The uppermost three of these field lines form a channel leading to the well, whereas the other ten flow directly into the well. This is in contrast

with the well at the top of the image, which has a clearly defined channel that follows the curvature of the upper rim of the ear. Twenty-two field lines flow from the right-hand side of the image to form the one remaining well at the right-centre of the ear.

3.6 Invariance Measures

To assess the results of tests for initialisation invariance and scale invariance, we shall use a measure of the normalised standard deviation of the distance between well positions, $Nsdist$. The distance between every pair of wells is calculated from the well vector coordinates, and then the standard deviation of the distances is divided by the mean distance, thus ensuring a scale invariant and rotation invariant measure. Figure 16 shows how the distances between all wells have been plotted and also shows the extracted and ordered distances as a bar graph in the top left corner.

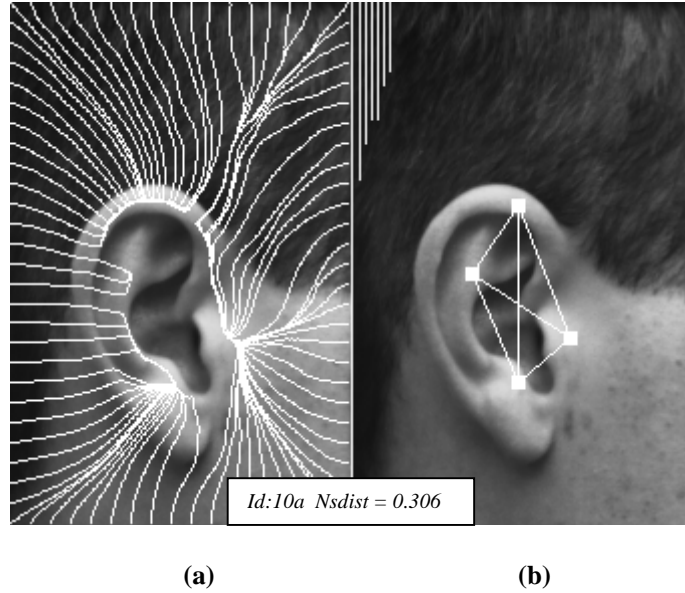


Figure 16 Inter-well distances as the basis of an invariance measure

For W wells with D inter-well distances $dist_i$, we have: -

$$Nsdist = \frac{\text{stdev}_{i \in D}(dist_i)}{\text{mean}_{i \in D}(dist_i)} \quad (23)$$

where

$$D = \frac{W^2 - W}{2} \quad (24)$$

3.7 Initialisation Invariance

Figure 17 demonstrates initialisation invariance. As in Figure 15 a set of 50 test pixels has been arranged in an ellipse shaped array and iterated to generate a set of field lines. To demonstrate initialisation invariance the centre of the ellipse has been displaced upwards in Figure 17(a), and downwards in Figure 17(b), whereas Figure 17(c) shows an initialisation along the edges of the image at intervals of 10 pixels. It can be seen that the force field structure is essentially preserved across the three images and that the location of the potential wells is the same in each case. This result is hardly surprising since the force field is not altered merely because we choose to enter it at different locations.

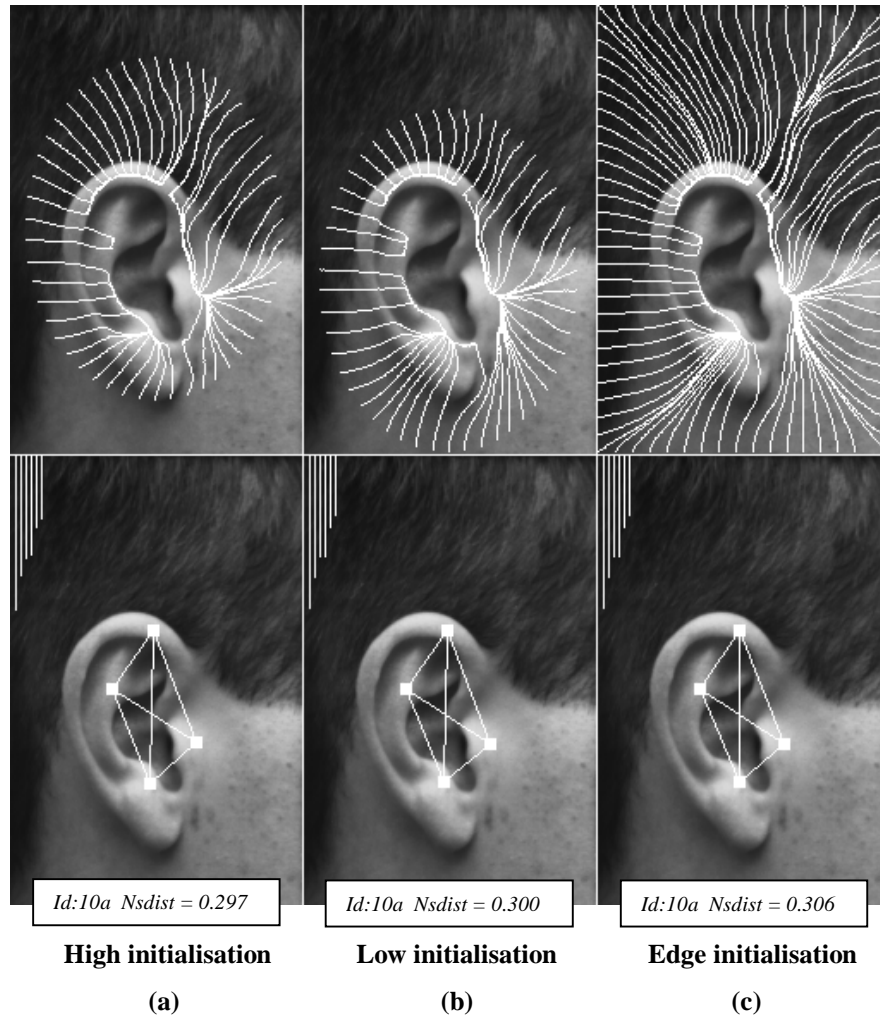


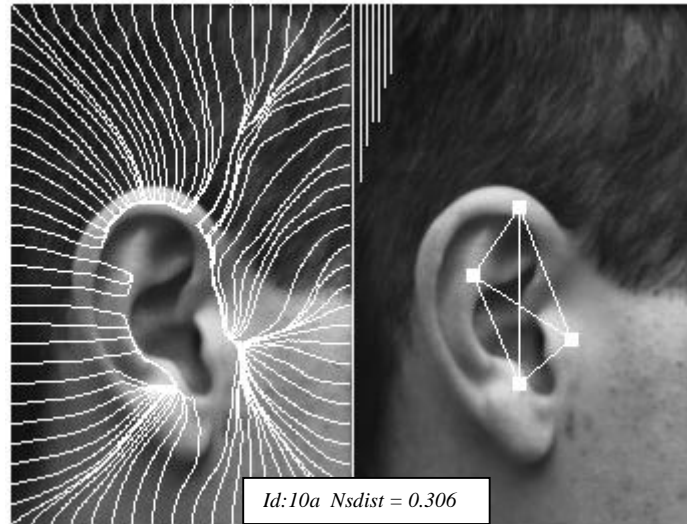
Figure 17 Initialisation invariance for 261x170 pixel ear image

It can be seen that it is a matter of chance whether the ellipse starting points happen to coincide with one of the edge initialised field lines. This can be seen in Figure 17(b), where the left hand well is formed from three field lines as opposed to two for the other initialisations. Clearly the result could vary with the number of ellipse points, especially if the number were chosen to be very small, where wells naturally would be omitted as no field lines would find them. This suggests simply that the number of initialisation points should be sufficiently large and sufficiently widely distributed so as to ensure that all wells will be extracted. The algorithm could perhaps be refined to use a regular grid of closely spaced initialisation points, perhaps eliminating one of each pair of test pixels when they come within a predefined region of convergence.

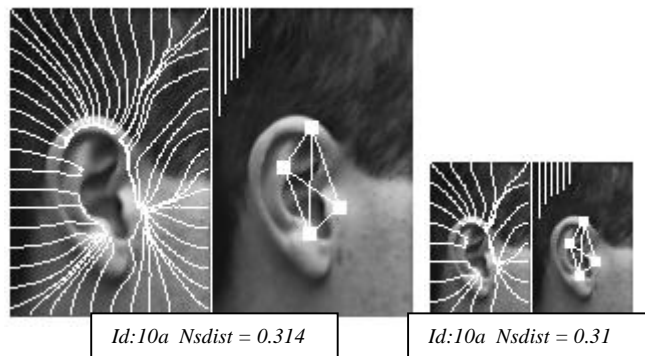
This demonstration also illustrates that the ear features can be extracted without the need for explicit ear extraction. Provided that the ear is roughly in the middle of the image, we see that the field lines will always seek out the same set of channels and wells. This can be attributed to the centric property of the force field, which derives from the dome shape of the energy field. The centric property comes at a cost, as it tends to de-emphasise details away from the centre towards the edge of the image. The reason for this is that the slope of the dome is steepest near the edge of the image and therefore the force is stronger there, and tends to pull test pixels across regions, which might otherwise form channels if the force were not so strong there. This aspect will be discussed in greater detail in the section on further work.

3.8 Scale Invariance

Figure 18 demonstrates scale invariance. The image in Figure 18(a) is at a resolution of 261×170 pixels and this is down scaled in the ratio of 3.4:2:1 so that Figure 18(b) has a resolution of 154×100 pixels and Figure 18(c) has a resolution of 77×50 pixels. It can be seen that the structure of the force field is essentially preserved when an image is captured and transformed at lower resolution. This is an important result because it confirms that scale space techniques can be employed so that a low-resolution image could be used to locate a target's position and a higher resolution version could then be used to refine feature information.



(a) 261×170



(b) 154×100

(c) 77×50

Figure 18 Scale invariance. Image ratios (a):(b):(c) :: 3.4:2:1

3.9 Noise Tolerance

Noise tolerance is demonstrated in Figure 19 where noise has been modeled as additive with a zero mean Gaussian distribution. It can be seen that the force field structure is very tolerant in the presence of severe noise.

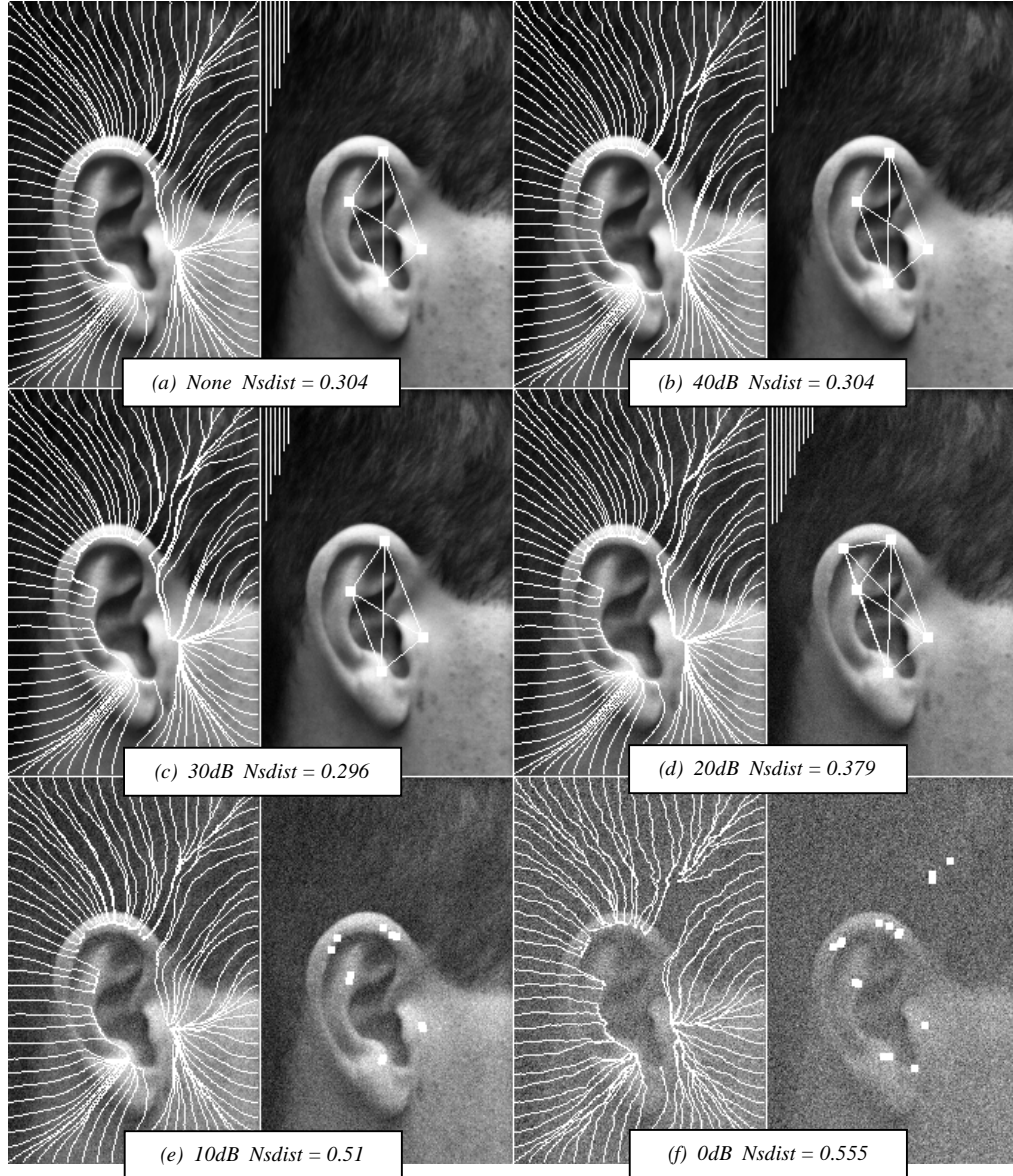


Figure 19 Zero mean Gaussian noise tolerance

Figure 19(a) has no added noise but is included for comparison. Figure 19(b) has a signal to noise (S/N) ratio (calculated as $20\log_{10}(S/N)$) of 40dB and appears to be unaffected, but careful examination shows that the well on the

left centre of the image now has four field lines converging upon it instead of three lines in the uncontaminated image. Figure 19(c) with 30dB S/N appears to be exactly the same as Figure 19(b), but interestingly the *Nsdist* measure has altered slightly from 0.304 to 0.296, indicating some slight shifting of well positions. In Figure 19(d), which has 20dB S/N, we now see that the channel running along the upper helix has disintegrated resulting in two wells instead of one. Figure 19(e) has 10dB S/N where we see that the wells have now become more chaotic but the channels are still quite recognisable. Figure 19(f) has 0dB S/N and we see that the wells have become really degenerate and also that the channels are beginning to look rather ragged, but are still surprisingly recognisable.

This confirms that the new technique has excellent noise tolerance, in accordance with the earlier observation that the transformation has intrinsic smoothing properties, and it is easy to see that even with very severe noise, a cleanup algorithm could be used to extract the channel information. This result can be viewed as quite significant because noise is a major problem in image processing especially when differential techniques are employed.

3.10 Uniqueness of Channels and Wells

Figure 20 demonstrates force field feature extraction for a set of eight different ears (i.e. eight more subjects), indicating that the channel descriptions are rich and varied, but we also see here the reason why wells alone will not suffice as compact ear descriptors. These ears all differ from the ear of the previous subject since its value of *Nsdist* was within 0.296-0.314 for scaling, initialisation, and S/N greater than 30 dB. As such, that ear would not be confused with these eight different subjects. However, ears of subjects id:14 and id:15 both have the same value for *Nsdist* and as such would be confused. This is the difficulty associated with discriminating with a scalar. But this scalar has sufficed for our purpose here.

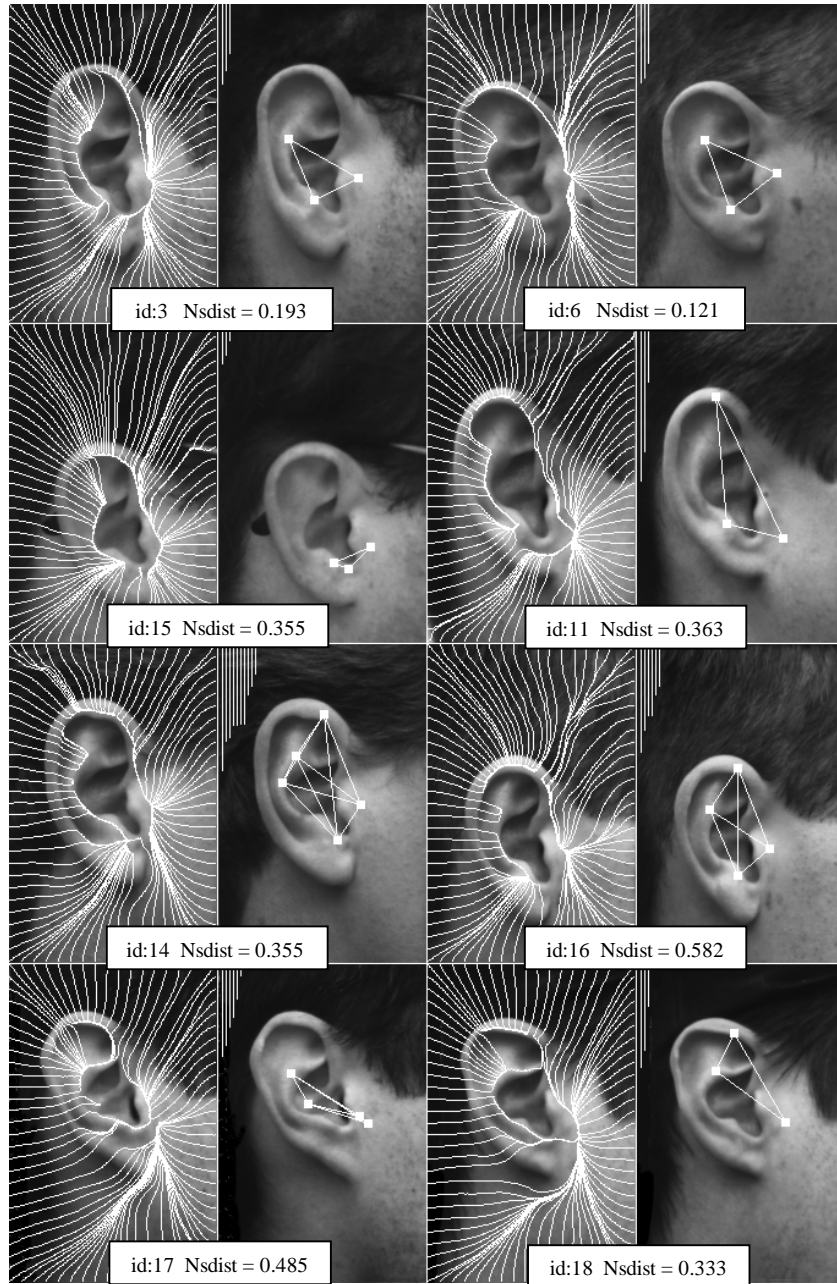


Figure 20 Uniqueness of wells and channels as ear features

The reason that wells are degenerate is twofold; their number is quite often insufficient for classification, and some are so close together that they could easily be absorbed as one. As shown in Figure 20(id:3), (id:6), (id:11), (id:15), and (id:18) most of the images have only produced three wells, which is barely enough for classification. In Figure 20(id:15) the three wells are very

close together, and the one in the middle has only two field lines leading to it. This well could easily merge with the right hand well, which would result in only two wells, clearly inadequate for classification. We also note that the triangle formed by these three wells, although very different in size, is quite similar to the triangle shown in Figure 20(id:11), and this is confirmed by the close agreement of their *Nsdist* measures of 0.355 and 0.363 respectively. The triangle formed by the three wells in Figure 20(id:18) with *Nsdist* = 0.333 is to a lesser extent similar to both of them.

On the other hand the channels provide a much more varied and rich description that looks far more promising as a biometric. Notice that in all cases that channels form along the upper helix and also along the antihelix, although the line of the antihelix is not rigidly followed, especially towards the bottom, where the antihelix joins the intertragic notch. In this small region, channel formation appears to behave with medial axis properties; its course being influenced both by the corner of the notch and the helix to the left. This property would go some way towards answering a potential criticism that the field lines seem to ignore the helix at the lower left of the ear. Indeed this medial axis property tends to dominate in regions where there are no strong edges or sometimes between edges. This is quite noticeable in Figure 20(id:11) and (id:17) where channel formation can be observed just under and to the left of the intertragic notch, where both of these ears have very weak lobes.

Whilst we have rejected the wells as a direct form of a compact ear biometric, we must not dismiss them completely, as they are so central to the channel formation process. It may also be the case that if force field feature extraction is applied to general object recognition, that wells may play a more stable and important role.

3.11 Convergence Feature Extraction

Here we introduce the analytical method of feature extraction as opposed to the field line method. This method came about as a result of analysing in detail the mechanism of field line extraction. As shown in Figure 22(c), when the arrows usually used to depict a force field are replaced with unit magnitude arrows, thus modeling the directional behaviour of exploratory test pixels, it becomes apparent that channels and wells arise as a result of patterns of arrows converging towards each other, at the interfaces between regions of almost uniform force direction. As the divergence operator of vector calculus measures precisely the opposite of this effect, it was natural to investigate the nature of any relationship that might exist between channels and wells and this operator. This resulted in the discovery of not only a close correspondence between the two, but also showed that divergence provided extra information corresponding to the interfaces between diverging arrows.

The concept of the divergence of a vector field will first be explained, and then used to define the new function. The function's properties are then analysed in some detail, and the close correspondence between the two forms of feature extraction is illustrated by superimposing one on the other.

The divergence of a vector field is a differential operator that produces a scalar field representing the net outward flux density at each point in the field. For the vector force field $\mathbf{F}(\mathbf{r})$ it is defined as,

$$\text{div } \mathbf{F}(\mathbf{r}) = \lim_{\Delta V \rightarrow 0} \frac{\oint \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S}}{\Delta V} \quad (25)$$

where $d\mathbf{S}$ is the outward normal to a closed surface S enclosing an incremental volume ΔV . In two-dimensional Cartesian coordinates it may be expressed as follows [63,64],

$$\text{div } \mathbf{F}(\mathbf{r}) = \nabla \cdot \mathbf{F}(\mathbf{r}) = \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} \right) \quad (26)$$

Figure 21 illustrates the concept of divergence graphically. In Figure 21(a) we see an example of positive divergence where the arrows flow outwards from the centre, and in Figure 21(b) we see negative divergence, where the arrows flow inwards, whereas in Figure 21(c) there is no divergence because all the arrows are parallel.

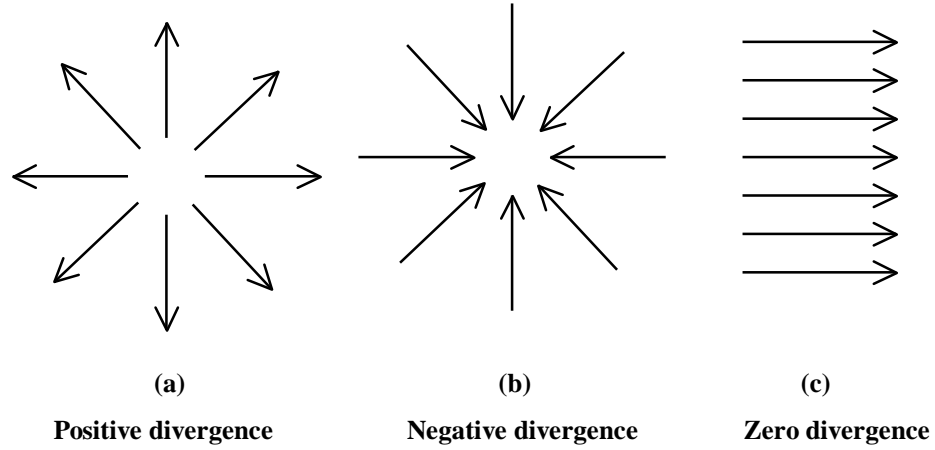


Figure 21 Divergence of a vector field

Having defined divergence we may now use it to define convergence feature extraction. 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. The new function maps the force field to a scalar field, taking the force as input and returning the additive inverse of the divergence of the force direction. The function will be referred to as the force direction convergence field $C(\mathbf{r})$ or just convergence for brevity. A more formal definition is given by

$$C(\mathbf{r}) = -\text{div } \mathbf{f}(\mathbf{r}) = -\lim_{\Delta A \rightarrow 0} \frac{\oint \mathbf{f}(\mathbf{r}) \cdot d\mathbf{l}}{\Delta A} = -\nabla \cdot \mathbf{f}(\mathbf{r}) = -\left(\frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} \right) \quad (27)$$

$$\text{where } \mathbf{f}(\mathbf{r}) = \frac{\mathbf{F}(\mathbf{r})}{|\mathbf{F}(\mathbf{r})|} \quad (28)$$

ΔA is incremental area

$d\mathbf{l}$ is the incremental boundary of ΔA

This function is real valued and takes negative values as well as positive ones

where negative values correspond to force direction divergence. Figure 22(b) shows an example of the convergence field for an ear, with the corresponding field line description shown in Figure 22(a), and a magnified portion of the force direction field shown in Figure 22(c). Small rectangular inserts drawn on the field line and convergence images indicate the portion that has been magnified. In Figure 22(b) the convergence values have been adjusted to fall within the range 0 to 255, so that negative convergence values corresponding to antichannels appear as dark bands, and positive values corresponding to channels appear as white bands. Notice that the antichannels are dominated by the channels, and that the antichannels tend to lie within the confines of the channels. Also notice how wells appear as bright white spots.

The detailed relationship between the convergence function and the channels and wells can be clearly seen by observing the patterns in the force direction field shown in Figure 22(c). Notice how a white ridge in the insert convergence field running from the top centre to bottom left of the insert corresponds to a line of converging arrows in the force direction field. Closer scrutiny of the two figures reveals that there are in fact three channels leading to a single well. The arrows from the top converge downwards and inwards forming the upper vertical channel. Arrows from the right converge leftwards and inwards to form a horizontal channel while the arrows from the bottom converge upwards and inwards forming the bottom vertical channel. These three channels then meet forming a well with the arrows pointing everywhere inwards. Also, clearly visible in the bottom right quadrant of the insert is a ridge of negative convergence that shows up as a black line. We can see that the corresponding arrows in Figure 22(c) clearly diverging.

It is evident that the convergence map provides more information than the field line implementation, in the form of negative versions of wells and channels or antiwells and antichannels, although it may be possible to modify the field line technique to extract this extra information. Also, notice that the two channels leading upwards and rightwards from the well in the insert are not extracted by the field lines because of the shielding effect of the strong channel running along the top of the ear. This illustrates another advantage of convergence over field lines.

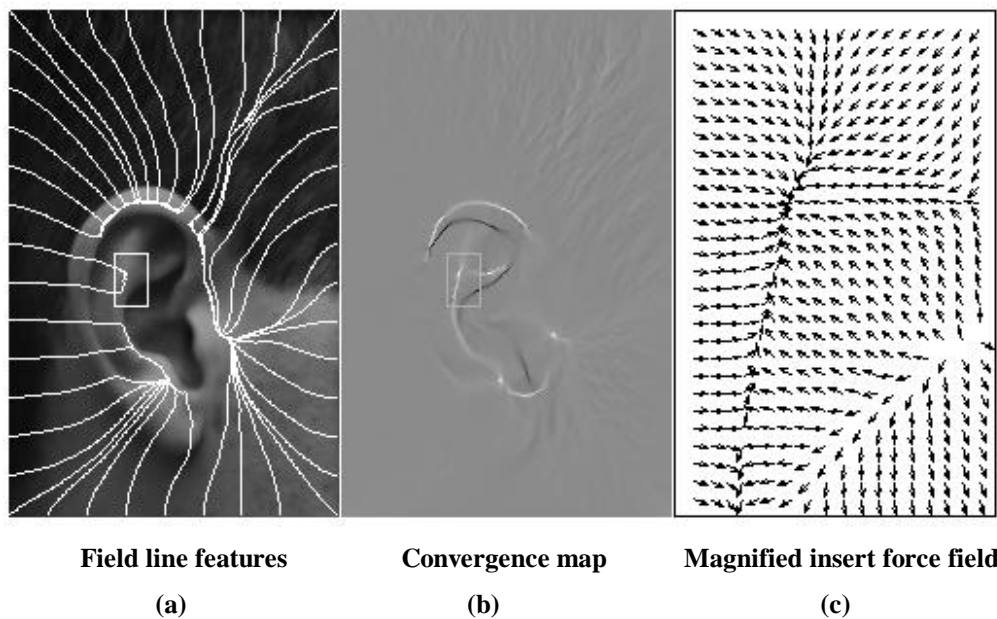


Figure 22 Convergence field

Figure 23 shows the convergence field of an ear with the corresponding field lines superimposed. Figure 23(a) is the field line map, and Figure 23(b) is the convergence map, while Figure 23(c) is the superposition of one on the other. We can see clearly how channels coincide with the ridges in the convergence map and that potential wells coincide with the convergence peaks. Notice the extra information in the centre of the convergence map that is not in the field line map, again illustrating the advantage of convergence over field lines.

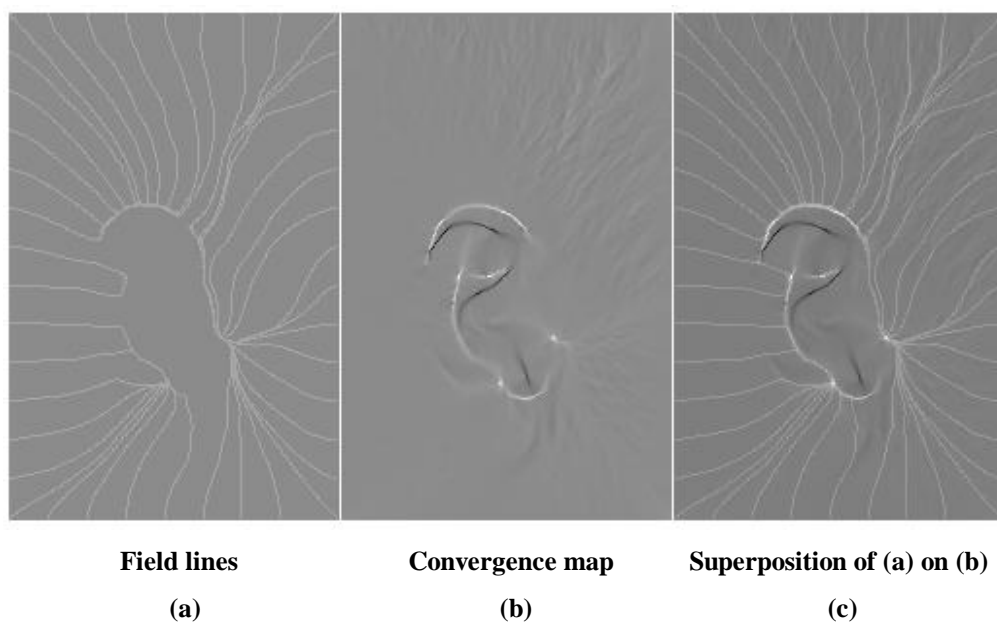


Figure 23 Correspondence between channels and convergence

3.12 Convergence and Marr-Hildredth

There is a striking resemblance between the form of the convergence function and the Marr-Hildredth edge detection operator [17]. The Marr-Hildredth operator takes the divergence of the gradient of a Gaussian smoothed image whereas the convergence operator takes additive inverse of the divergence of the direction of the gradient of the energy transform of the image. The energy transformation can be viewed as smoothing the image with an inverse function $1/r$ shaped kernel instead of a bell shaped Gaussian kernel. The difference is therefore in the form of smoothing used but more importantly using gradient direction rather than gradient for the convergence function, thus making it a non-linear operator.

Analytically the Marr-Hildredth operator is given by

$$MH(\mathbf{r}) = \text{div } \mathbf{g}(\mathbf{r}) = \nabla \cdot \mathbf{g}(\mathbf{r}) = \left(\frac{\partial g_x}{\partial x} + \frac{\partial g_y}{\partial y} \right) \quad (29)$$

Compare this with the convergence operator given by

$$C(\mathbf{r}) = -\text{div } \mathbf{f}(\mathbf{r}) = -\nabla \cdot \mathbf{f}(\mathbf{r}) = -\left(\frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} \right) \quad (30)$$

$$\text{where } \mathbf{g}(\mathbf{r}) = \text{grad}(G(\mathbf{r}) * I(\mathbf{r})) \quad (31)$$

$I(\mathbf{r})$ is the original image.

$$\text{and } G(\mathbf{r}) = e^{-\frac{|\mathbf{r}|^2}{2\sigma^2}} \text{ is a Gaussian kernel.}$$

We are not suggesting that convergence competes with the Marr-Hildredth operator, or even that it accrues any advantage from its similarity of form. We merely observe that we notice the similarity and find it interesting. The kernel used by Marr-Hildredth is optimal in Gaussian noise whereas the inverse $1/r$ kernel is an artifact of its force field nature, and the intrinsic smoothing it affords is merely a fortunate consequence of its formulation.

3.13 Convergence and Curvature

The **shape operator** S of a surface M in \mathfrak{R}^3 [22] measures the rate of change of a unit normal U in any direction on M and thus describes the way the shape of M is changing in that direction. If \mathbf{p} is a point on M , then for each tangent vector \mathbf{v} to M at \mathbf{p} , let

$$S_p(\mathbf{v}) = -\nabla_{\mathbf{v}} U \quad (32)$$

where U is a normal vector field on a neighbourhood of \mathbf{p} in M , then S_p is called the shape operator of M at \mathbf{p} derived from U . The shape operator is a **linear operator** on the tangent plane of M at \mathbf{p} .

The number $k(\mathbf{u}) = S(\mathbf{u}) \cdot \mathbf{u}$ is called the **normal curvature** of M in the \mathbf{u} direction, where \mathbf{u} is a unit vector tangent to M at point \mathbf{p} . The maximum and minimum values of the normal curvature $k(\mathbf{u})$ of M at \mathbf{p} are called the **principal curvatures** of M at \mathbf{p} , and are denoted by k_1 and k_2 respectively. The directions in which these extreme values occur are called **principal directions** and unit vectors in these directions are called **principal vectors**. Provided normal curvature is not constant then the two principal directions are orthogonal, and the principal curvatures are the eigenvalues of S and the principal vectors are the eigenvectors of S .

The **Gaussian curvature** K is the determinant of the shape operator S and **mean curvature** H is half the trace of S . Gaussian and mean curvatures can be expressed in terms of principal curvature by

$$K = k_1 k_2 \quad \text{and} \quad H = (k_1 + k_2)/2 \quad (33)$$

The mean curvature H can be identified with the Laplacian $I_{xx} + I_{yy}$ of an image function $I(x, y)$, and therefore confirms that the Marr-Hildreth operator, otherwise known as the LoG operator (Laplacian of Gaussian) is essentially based on image curvature. However, it must be emphasised that this does not equally apply to the convergence operator because this is a non-linear operator due to the fact that it is based on the direction of the force rather than the force itself.

4 Ear Recognition

It is useful to review intermediate results and conclusions at this stage before proceeding with the task of ear recognition. We have developed dual force field and potential energy transforms, which enable us to transform an image, with powerful intrinsic smoothing but without loss of information, into a smooth dome shaped energy surface. In a similar manner to the popular gradient descent algorithm, we can initialise a distribution of exploratory test pixels and allow them to traverse the surface gradient until they eventually become trapped in a small number of energy wells, forming energy channels along the way. Since these channels and wells dominate the character of the otherwise smooth energy surface, we argue that they should make good ear features. Scale invariance, initialisation invariance and noise tolerance have been confirmed. Provided the ear is roughly centred in the image, explicit ear extraction is not required, due to the powerful autocentring property attributable to the steep sides of the dome shaped surface.

An analytic dual of the force field approach has been developed where we see that wells and channels are really different manifestations of the same phenomenon; they are just peaks and ridges in the value of the convergence function. This approach gives a more comprehensive description in the form of antichannels and antiwells, and also extracts information that may have been shielded by powerful channels in the field line case. We will also see in this section that with the analytic version we can calculate a centroid whose position is fairly stable relative to the features. The stability of the centroid may also be attributed to the dome shape, which tends to exclude outlying data.

The early hope of using just wells as a compact description did not live up to its expectation and the richer but more complicated description provided by the channels will be used instead. The problem of comparing ears has been reduced to one of comparing channels. This should be easier than comparing ears directly but it is still not a trivial task. Certainly the dimensionality of the channel description is considerably less than that of the original image. Figure 24 shows a synthetic image whose principal field lines have been identified,

where these could be defined as the field lines that make the most contribution to a given channel. Assuming that a way can be found of easily extracting such lines they will still need to be compared. Perhaps techniques derived from differential geometry, which deals with the geometry of curves and surfaces, could be exploited to describe the form of the channels [22]. Perhaps contour description techniques could be used to describe and compare the channels, although it might not be as easy as describing and comparing just simple closed contours [68,69]. Principal field lines would appear to offer a ready made list of points and directions to facilitate contour descriptions, while the convergence approach would require thresholding, thinning, and differentiation to extract the same information.

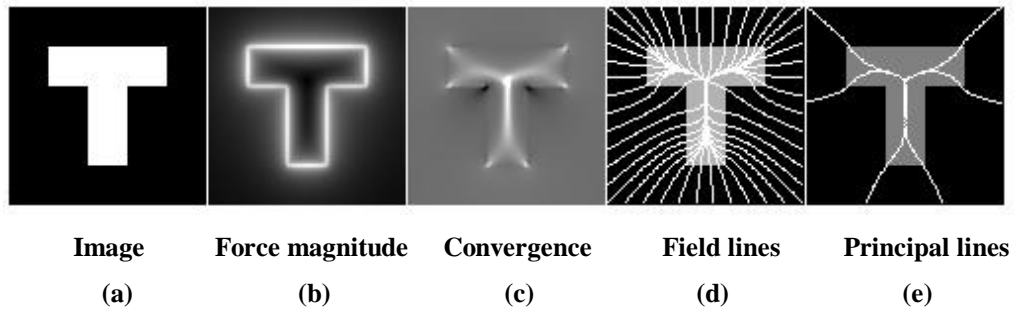


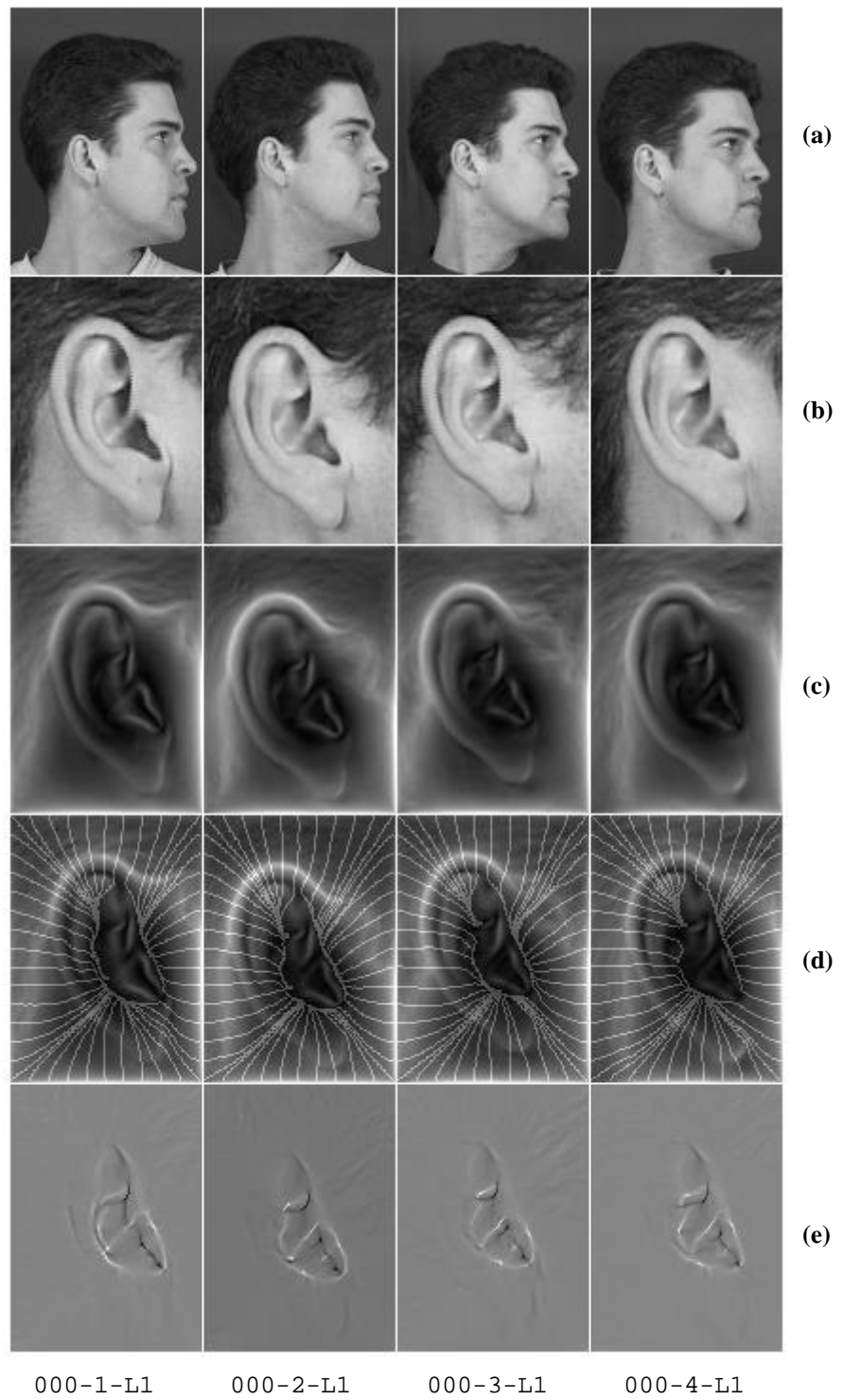
Figure 24 Principal field lines

Considerations along these lines make it clear that the evaluation of the ability of the technique for recognition is very much at the mercy of the method chosen to compare channels. There is no guarantee that the method chosen will be the optimal one. A given technique can at best only establish a lower bound on the recognition capability. For this reason a lower bound on the validity of the technique is established by using just basic template matching of binarised and thresholded convergence functions, and implemented using Fourier cross-correlation techniques. This approach is not invariant to scale and rotation but should nevertheless establish a useful lower bound. More sophisticated techniques will be left for discussion in the chapter on further work. Also, to establish a reference baseline, the same template matching technique is applied to the database using binarised and thresholded edge detection.

4.1 Ear Database

In order to create a suitable database of experimental ear images the XM2VTS face profiles database [28] was chosen, from which 63 subjects were selected as being suitable candidates for ear recognition by eliminating those whose ears were covered by hair. The ear database thus consists of 4 samples of each of 63 subjects that were captured during 4 sessions over a 5-month period thus ensuring excellent natural intra-class variation. Diffuse lighting was used during image capture. The ear images extracted are at a resolution of 141×101 pixels, which is a suitable resolution to evaluate the new technique. Table 7, Appendix A4 lists the identification numbers of the subjects selected from the XM2VTS database.

Figure 25 and Figure 26 show how the first two subjects, with the XM2VTS identification numbers 000 and 001, have been processed to extract the convergence maps. The original colour profiles were converted to grey scale (a) and a 141×101 pixel frame was manually adjusted to surround the ear portion and crop the ear images (b). The force field transforms of the ear images were taken, and force field magnitudes calculated (c). Test pixels were arranged around the edge of the force field, at intervals of 10 pixels, and iterated to extract field line features (d). Finally the force fields were converted to convergence fields (e). Note that the second subject shown in Figure 26 actually changes hair colour, not just shade, during database acquisition. This does not affect the structure of the force field or its convergence, as expected.



(a) Profile (b) Ear (c) Force magnitude (d) Field lines (e) Convergence field
Figure 25 Feature extraction for subject 000

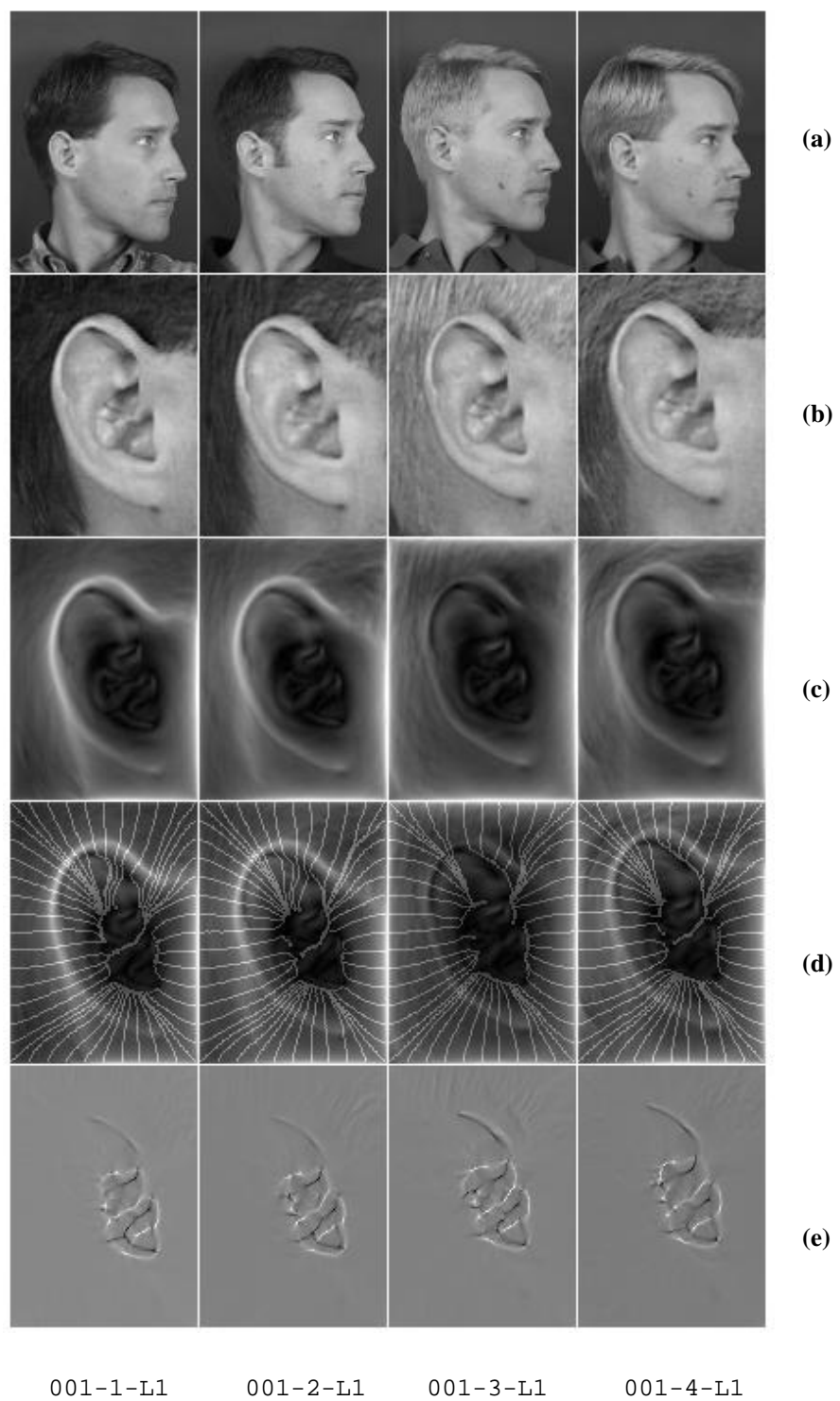


Figure 26 Feature extraction on subject 001

4.2 Template Matching

This section covers the results of ear recognition experiments on the subset of the XM2VTS database. The aim is to validate the technique by using simple template matching to establish a lower bound on the recognition capability. Template matching is implemented using Fourier based cross-correlation on thresholded and binarised convergence functions. Thresholding is not only in terms of the convergence level but also in terms of a radial exclusion zone centred on the convergence centroid shown in Figure 27(b). The centroid of the convergence tends to be stable with respect to the ear features, and this property was exploited to offset the lack of rotation invariance in template matching by adding the ability to rotate the templates about their centroids during matching. The radial exclusion zone as well as removing unwanted outliers such as bright spots caused by spectacles, also results in circular templates, which are more suitable for rotation matching.

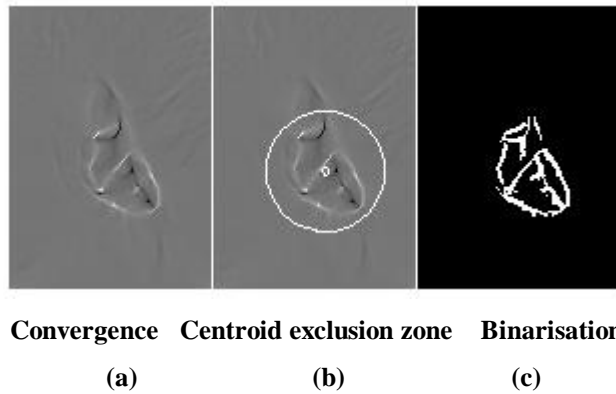


Figure 27 Binarised thresholded convergence function

Figure 27 shows how the convergence functions are thresholded and binarised before Fourier cross-correlation is performed. First the convergence map is thresholded by setting to zero all values with magnitude less than two standard deviations, and then binarising the magnitude of the result. This threshold value was chosen experimentally to obtain the template channel thickness shown in Figure 27(c), which was felt to be a good compromise for template matching. Figure 27(a) shows the convergence map before further processing has been applied, and Figure 27(b) shows where the centroid has been

calculated and used to establish an exclusion zone of 30 pixels. Figure 27(c) shows where the result has been binarised resulting in a circular template of radius 30 pixels.

In order to further facilitate the correlation process the thresholded and binarised convergence functions are uniformly aligned in 60-pixel \times 60-pixel frames. This is achieved by translating each one so that its centroid now lies at the centre of the new frame. The results for the first 8 subjects are shown in Figure 28 and the results for the remainder of the subject population are shown in Appendix A3. These results also confirm the stability of the centroids with respect to the features, since visual inspection of the results confirm that all the templates are correctly aligned in the middle of their frames. If a particular centroid were unstable, then the corresponding template would appear offset.

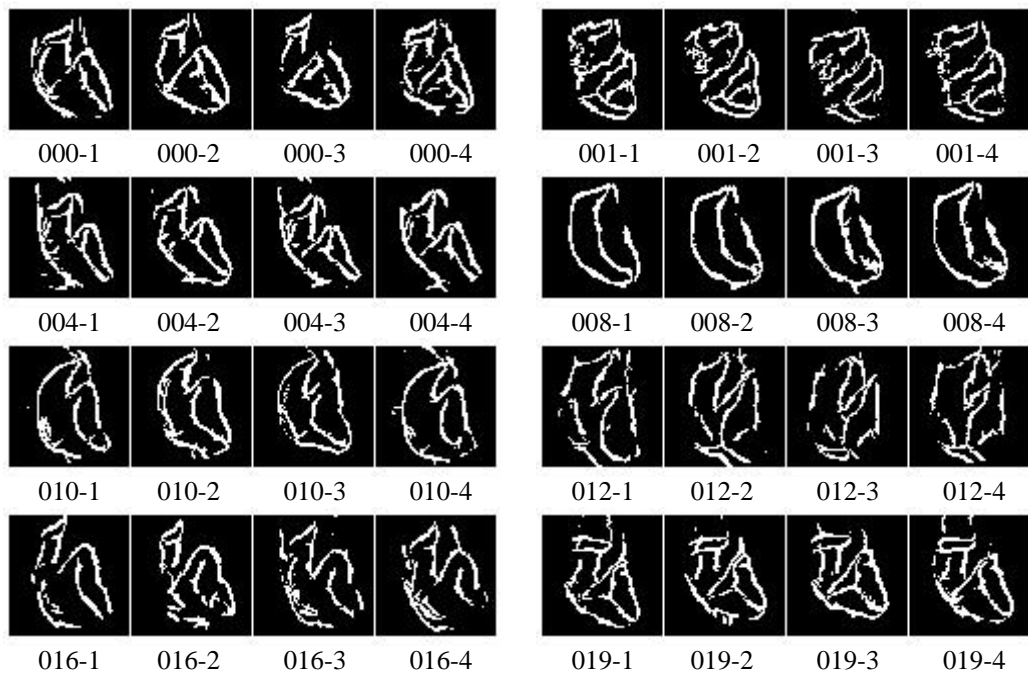


Figure 28 A selection of binarised and thresholded convergence functions

4.3 Classification Results

The binarised convergence maps were exhaustively cross-correlated to yield a 252×252 element array of cross-correlations. The matrix of correlations is too large to be displayed directly, but can be illustrated graphically by displaying it as a grey scale image. The template rotation referred to in Section 4.2 was not used in practice, owing to time constraints.

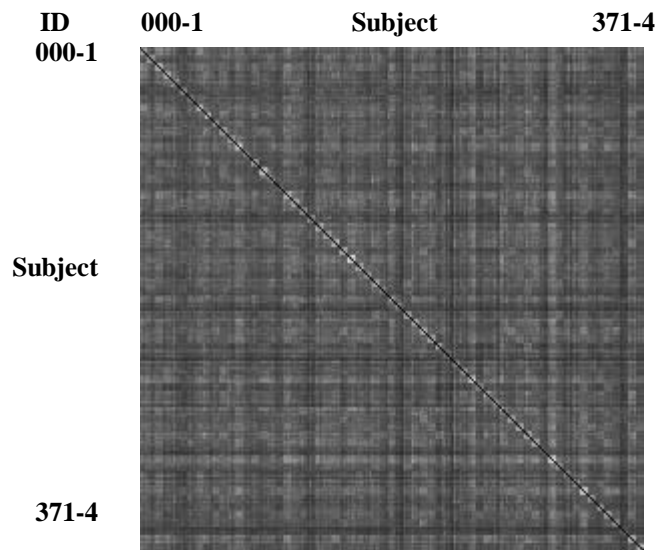


Figure 29 Cross-correlation matrix depicted as a grey scale image

This is shown in Figure 29, where each row represents the cross-correlation of one subject sample against all the other others. The bright squares along the main diagonal correspond to intra-class correlation values and all pixels outside these 4×4 pixel squares correspond to inter-class correlation values. The pixels corresponding to self-correlations, occurring precisely along the main diagonal, have been set to zero, as they supply no useful information. The fact that some rows are either darker or brighter than others is thought to reflect the fact that the corresponding ears have either far from or close to average ear shape where the average is based on the XM2VTS subset.

This method can be taken further, to illustrate nearest neighbour classification, as shown in Figure 30. The position corresponding to the highest value in each row is assigned a white pixel while all the others are set

to black. The self-correlations along the main diagonal are excluded, as they must obviously yield the highest value for each row. As expected most of the white pixels lie within a 4×4 pixel square centred on this main diagonal indicating correct nearest neighbour classification. White pixels lying outside these squares indicate incorrect classifications.

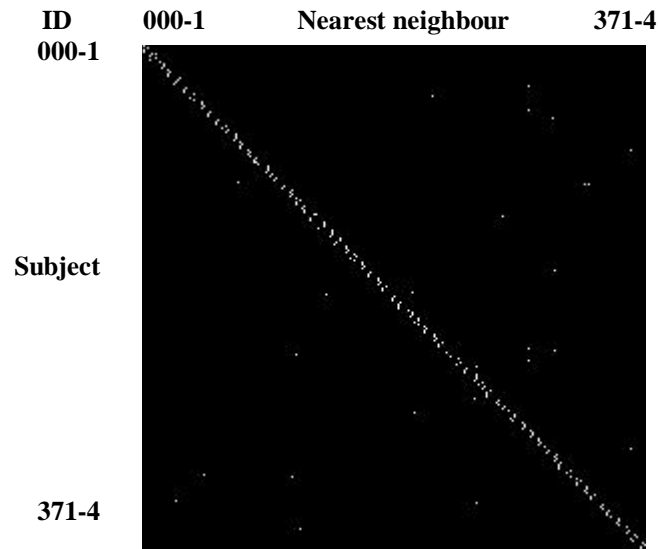


Figure 30 Nearest neighbour classification matrix

Out of the 252 trials 230 have been correctly classified corresponding to a rate of 91.3%. Very careful scrutiny of Figure 30 reveals that there are actually 258 white pixels instead of just 252. This occurs because six of the trials return pairs of identical maximum values, which occurs as a consequence of binarisation of convergence values. This information has been summarised in Table 3, where we see that rows 4, 58, and 84 have pairs of identical values in their correct squares, whereas rows 52 and 215 have one of a pair of identical values in the correct square, and row 69 has a pair both outside the correct square.

Row	4	52	58	69	84	215
Columns	5	53	56	221	86	75
	7	244	57	223	87	214

Table 3 Summary of trials resulting in pairs of identical maximum correlation values

4.4 Statistical Analysis of Results

The mean and standard deviation of the intra-class and inter-class correlations give a good indication of the nature of the results. Table 4 shows these statistics. Note that self-correlations that occur along the main diagonal of the cross-correlation matrix have been excluded from the calculations.

Intra-class		Inter-class	
Mean	Std. deviation	Mean	Std. deviation
221	51	138	25

Table 4 Statistics for intra-class and inter-class convergence correlations

Using these results we may compute the decidability index after Daugman [70]. The decidability index d' measures how well separated the distributions are, since recognition errors are caused by their overlap. The measure aims to give the highest scores to distributions with the widest separation between means, and smallest standard deviations. If the two means are μ_1 and μ_2 and the two standard deviations are σ_1 and σ_2 then d' is defined as

$$d' = \frac{|\mu_1 - \mu_2|}{\sqrt{(\sigma_1^2 + \sigma_2^2)/2}} \quad (34)$$

Substituting the mean and standard deviation from Table 4, we have

$$d' = \frac{|221 - 138|}{\sqrt{(51^2 + 25^2)/2}} = 2.1 \quad (35)$$

Figure 31 shows the frequency distributions for the intra-class and inter-class correlations. Normal distributions corresponding to their respective means and standard deviations have been fitted, suggesting that they might be normally distributed. The cross-correlation axes have been divided into 40 intervals, between minimum and maximum values, resulting in 40 bins or columns.

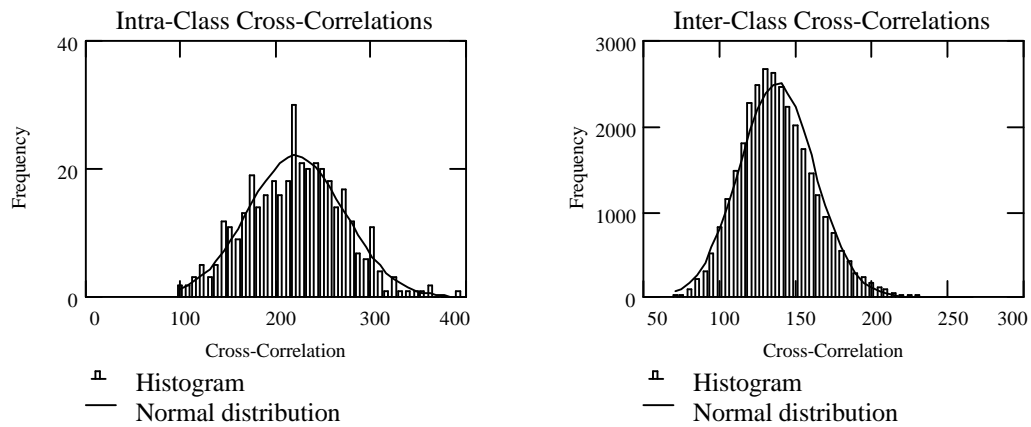


Figure 31 Normal distribution fitted to frequency distribution of cross-correlations

Since the frequency distributions are shown to be normally distributed, we can normalise them and form probability density functions as shown in Figure 32.

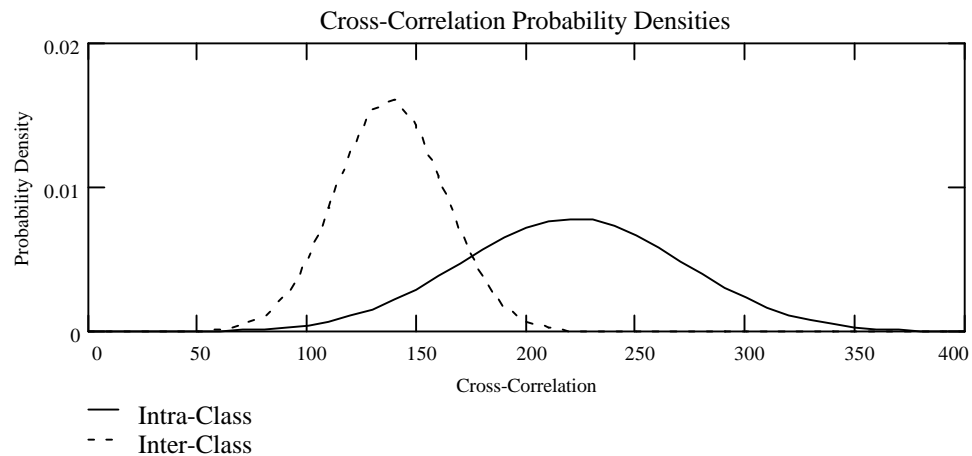


Figure 32 Probability density functions of intra-class and inter-class correlations

The area of intersection formed between the two distributions, compared with the areas of the distributions themselves, gives an indication of the probability of misclassification. The classification threshold can be set to a value in this

region according to the chosen security policy. A high threshold corresponds to a conservative policy, whereas a low value threshold corresponds to a liberal policy.

Finally, the probability density functions can be combined to produce a receiver operator characteristic as shown in Figure 33. This type of graph aims to directly show the tradeoffs involved when adjusting the classification threshold, although it does not usually show the threshold value itself. In this particular case equal FRR and FAR error rates of 13.5% are predicted at a threshold setting of 165.

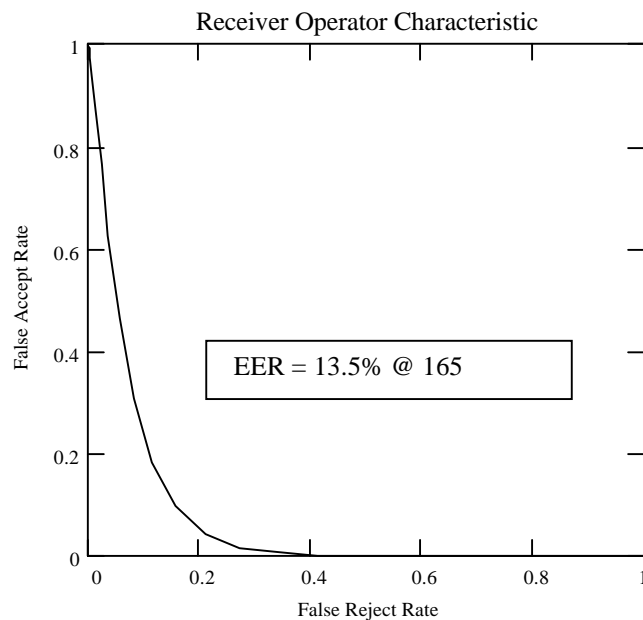


Figure 33 Receiver operator characteristic

4.5 Edge Detection Template Matching

In order to provide a baseline with which to compare the force field results, the same template matching technique has been applied to the ear database, but now using simple edge detection without smoothing, rather than binarised and thresholded convergence maps. The edge detected images have been thresholded by setting to zero all those edge values that are less than two standard deviations. The thresholded results are then binarised before doing template matching in the form of Fourier cross-correlation. Figure 34 shows a sample of the results for the edge detection process. Notice that hair prevents definition of a bounding region, confirming the inherent smoothing advantage in the new approach, which is not troubled by hair.

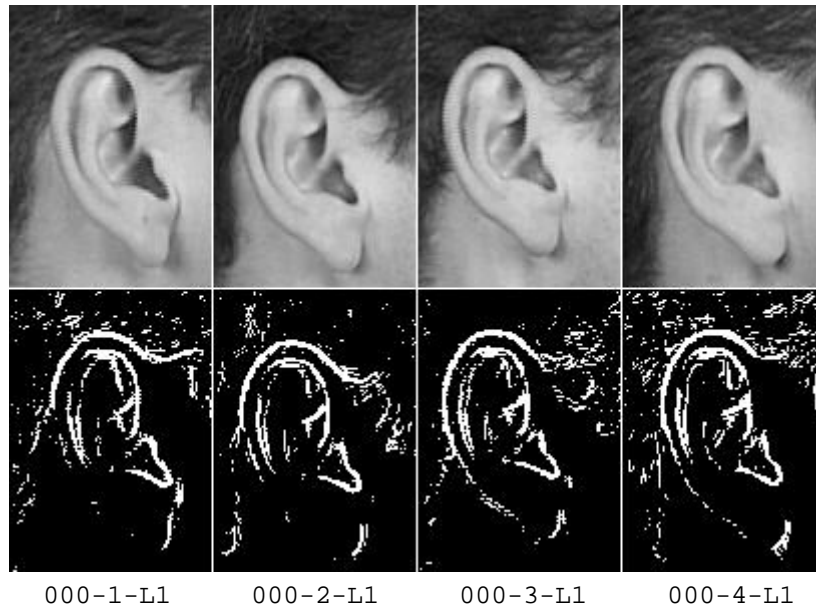


Figure 34 Sample of results of thresholded and binarised edge detection

The mean and standard deviation of the intra-class and inter-class correlations are shown in Table 5 where we notice that the correlation values are roughly three times the magnitude of those obtained with the convergence functions, this being due to the greater area of edge detection templates.

Intra-class		Inter-class	
Mean	Std. deviation	Mean	Std. deviation
611	152	351	75

Table 5 Statistics for intra-class and inter-class edge detection correlations

Using these results, we may also compute the decidability index as before. Using Equation 34 we have,

$$d' = \frac{|611 - 351|}{\sqrt{(152^2 + 75^2)/2}} = 2.2 \quad (36)$$

We see that this result is almost identical to the case for convergence maps, which gave a 2.1 decidability index, thus confirming that the new technique is comparable to the well established recognition technique of template matching using edge detection.

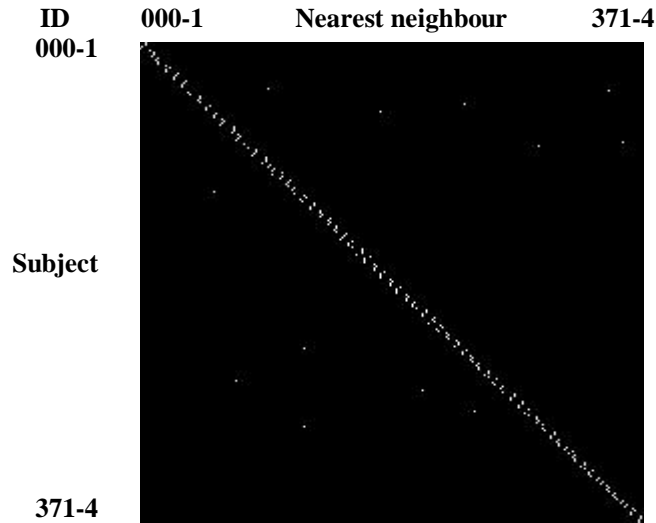


Figure 35 Nearest neighbour classification matrix for edge detection

Figure 35 shows the corresponding nearest neighbour classification matrix for the cross-correlation of the 252 subjects, where we see that there has been a classification rate of 94.4%, which is slightly better than the rate of 91.3% achieved with convergence functions. This confirms the earlier analysis, which also suggested slightly better performance on the edge data. Naturally the performance of either approach could doubtless be improved by better

selection of the parameters involved, or by other distance measures or by other classification strategies. The interest here is that results by basic implementation are of the same order. However, the new force field approach has better invariance characteristics and superior noise performance.

Figure 36 shows normal distributions fitted to the cross-correlation frequency distributions again suggesting that they are normally distributed.

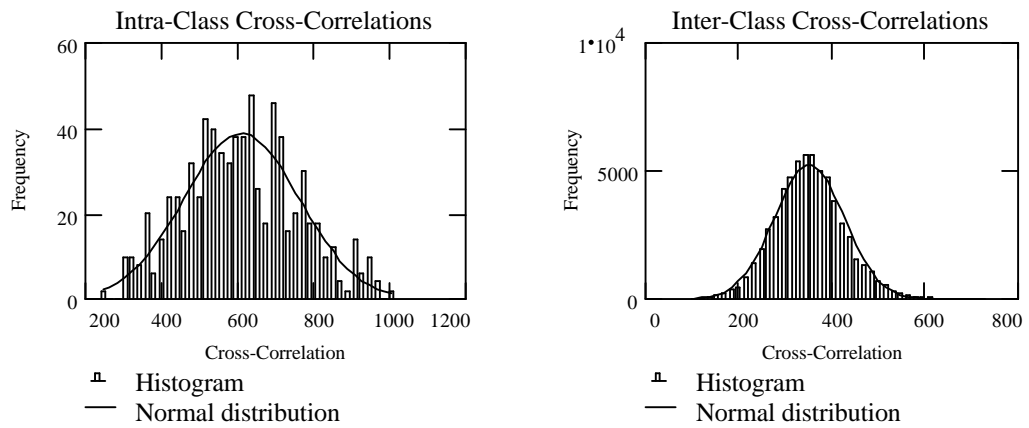


Figure 36 Normal distribution fitted to frequency distribution for edge detection

Because the cross-correlation frequency distributions are normally distributed, we may compute the corresponding probability density functions as shown in Figure 37 where we see that they agree with the decidability index of 2.2 that we calculated above.

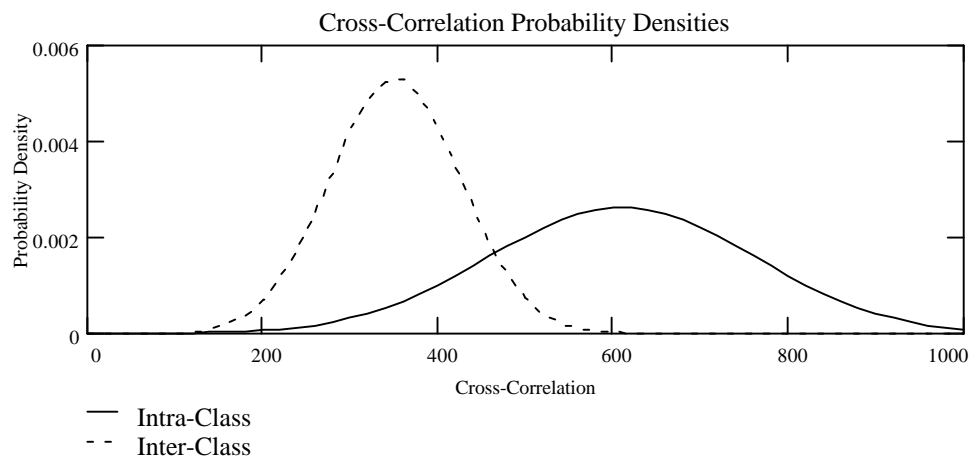


Figure 37 Probability density functions for edge detection

Finally, we may also calculate the receiver operator characteristic for the edge detection case, which is shown in Figure 38.

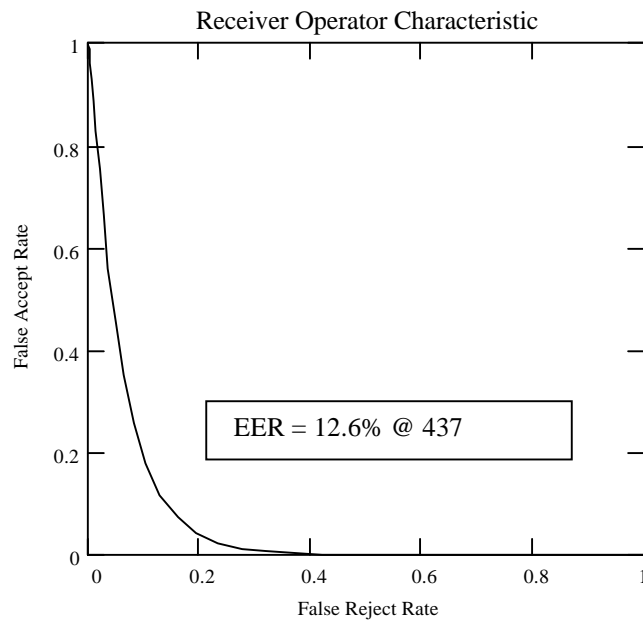


Figure 38 Receiver operator characteristic for edge detection

Again, performance by edge detection is slightly better than that by our new method. However, the edge operators do not enjoy performance attributes offered by the force field transform, including robustness to noise and preference of centrally located objects, notwithstanding the initialisation invariance discussed earlier. As such our contention is that the new approach is certainly a worthy contender to describe ears for recognition.

5 Further Work

This section briefly addresses issues that have not been dealt with in the main body of the thesis or that have been postponed. It also includes a detailed description of polar functions, which have been designed to take advantage of the centroid property of the convergence functions, but which have not been completed.

5.1 Autocentricity

Provided the ear is roughly centred in the image, explicit ear extraction is not required, due to the powerful autocentring property attributable to the steep sides of the dome shaped surface. However, there is a tradeoff involved here; it is possible that valuable information is being discarded unnecessarily. The closer the desired data is to the edges of the surface, the more likely it is to be overlooked, and conversely the more it is concentrated towards the centre, where the slope is shallowest, the more likely it is to be included. Therefore it needs to be investigated to establish how the ratio of the size of the image to the size of the target ear can affect the outcome. The mean value component of the dome shape in Figure 5(b) could be extracted, rescaled, and reinserted, to modify the behaviour. Removing the mean completely may have the effect of completely removing the autocentring property.

5.2 Feature Matching and Indexing

The template matching method used to establish a lower bound on the recognition capability used in Section 4.2 is clearly inadequate for large scale biometrics, if only on the grounds of the sheer number of cross-correlations involved, and also does not do justice to the features. More sophisticated techniques of feature matching will need to be investigated, and also a method of indexing the features to handle very large database searches.

5.3 Generating an Ear Database

In generating a purpose designed ear database it is reasonable to assume that the subjects will be enrolled and verified under carefully controlled conditions, and therefore the bulk of the database samples should be captured under those conditions. However, it would be prudent to anticipate and include samples of as many parameter variations as possible, even if they are not subsequently used. Some of the more important assumptions and considerations are: -

- **subjects reasonably cooperative** – it is assumed that subjects would be prepared to stand in a set place and would be prepared to push their hair to one side but would not be prepared to remove spectacles or jewelry. Therefore some samples of bespectacled subjects should be obtained both with and without their spectacles. Some samples of subjects with and without their jewelry should also be obtained.
- **diffuse lighting at fixed illumination level** – this is a reasonable assumption in biometric applications, but some samples should be obtained under varying illumination and using directed lighting.
- **fixed distance and orientation** – the distance could be selected to be the typical distance one would stand away from a cash machine, one to two feet perhaps, although some close-up and more distant samples should also be obtained.
- **Same left or right ear presented each time** – a subject's left and right ears are generally very similar but by no means identical. Both ears should nevertheless be sampled.
- **variation over time** - time variation should be accommodated by taking samples of the same subjects at regular intervals, perhaps four samples at six-week intervals. This could be arranged to include samples captured both in winter and summer to test for any variation caused by cold ears.
- **natural variation** – the database should include a good mix of ethnic variation of both sexes and should include a good subject age spread.

5.4 *Miscellaneous Further Work*

Other matters that might merit further investigation could include: -

- **Magnitude** – the new feature extraction technique is based solely on force direction; the force magnitude shown in Figure 11 clearly also contains significant information that could be incorporated in the ear description.
- **Wider application** – force field feature extraction has been developed for ear biometrics but there is no reason why it should not have wider application in image processing and this could be further investigated.
- **Proof of invertibility** – this remains incomplete though the transform would appear invertible. If a proof cannot be found then the condition number approach to numerical testing should be further investigated, as determinant calculation is inherently numerically unstable.
- **Corner detection** – the technique would appear to be well suited for detecting curvature. This was not revealed by analysis on ears, more by the analysis on pentominoes as shown in Figure 43 and as corners are important low-level features, this would merit further investigation.
- **Speed of calculation** – even using fast Fourier transform techniques the speed of calculation is not suitable for real time applications so that approximate methods of calculation should be investigated and also whether it may be possible to implement the field transforms directly in hardware using electrostatic mapping.

5.5 Polar Functions

Polar functions have been designed especially to exploit the stability of the convergence function centroid, and are readily adapted to be scale and rotation invariant. The polar functions take the output of the convergence function as input and return closed contours representing the moments [71] of the radial distribution of convergence with respect to the centroid of convergence. The cases of positive and negative convergence can be processed separately and the results combined; or the magnitude of convergence can be taken before processing is carried out.

The centroid \mathbf{m} of the convergence field $C(\mathbf{r})$ given by

$$\mathbf{m} = \frac{\sum_i \mathbf{r}_i C(\mathbf{r}_i)}{\sum_i C(\mathbf{r}_i)} \quad (37)$$

Now if the convergence field is divided into sectors centred about the centroid \mathbf{m} then the mean distance from the centroid for a given sector S is given by

$$md_S = \frac{\sum_{i \in S} |\mathbf{r}_i - \mathbf{m}| C(\mathbf{r}_i)}{\sum_{i \in S} C(\mathbf{r}_i)} \quad (38)$$

If this calculation is repeated for each sector then a sequence of values will result that can be associated with the corresponding sector angular position to obtain a discrete polar function of mean distance versus sector angular displacement. No attempt is made to prescribe the number of sectors to be used or the alignment of the sector boundaries. For example $C(\mathbf{r})$ could be divided into 60 sectors each assigned one of the angles $0, 2\pi/60, \dots, 59 \times 2\pi/60$, with the first sector straddling the x -axis.

In a similar way higher order moments of the convergence distribution can be employed such as variance, skewness, and kurtosis [71]. Since these measures are firmly based on statistical theory they can be expected to

contribute independent information but the effects of noise could be expected to become more troublesome as the order increases.

The *variance* about the mean distance for a given sector S is given by

$$vd_s = \frac{\sum_{i \in S} (|\mathbf{r}_i - \mathbf{m}| - md_s)^2 C(\mathbf{r}_i)}{\sum_{i \in S} C(\mathbf{r}_i)} \quad (39)$$

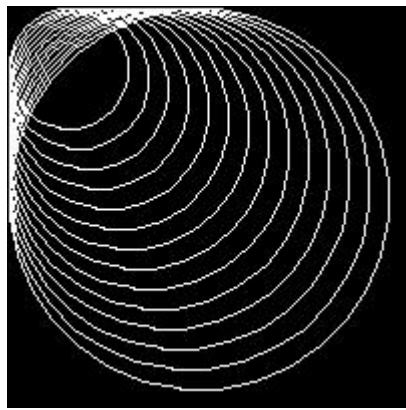
The *skewness* or asymmetry about the mean distance for a given sectors S is given by

$$skewd_s = \frac{\sum_{i \in S} \left(\frac{|\mathbf{r}_i - \mathbf{m}| - md_s}{\sqrt{vd_s}} \right)^3 C(\mathbf{r}_i)}{\sum_{i \in S} C(\mathbf{r}_i)} \quad (40)$$

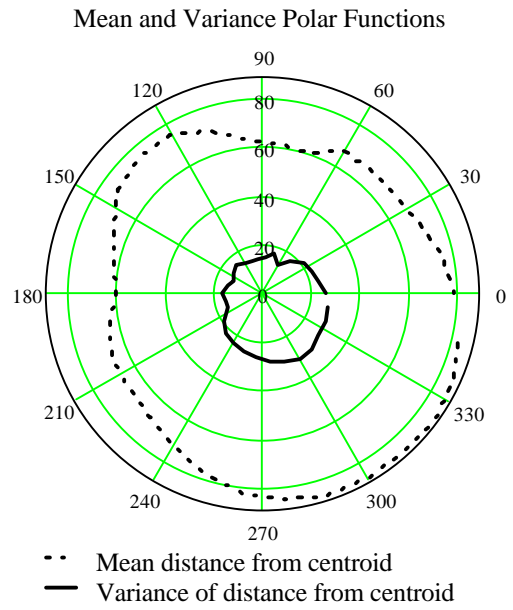
The *kurtosis* or peakiness of the distance distribution about the mean distance for a given sector S is given by

$$kurtd_s = \frac{\sum_{i \in S} \left(\frac{|\mathbf{r}_i - \mathbf{m}| - md_s}{\sqrt{vd_s}} \right)^4 C(\mathbf{r}_i)}{\sum_{i \in S} C(\mathbf{r}_i)} - 3 \quad (41)$$

Figure 39 shows the first two polar functions for a synthetic image designed especially to test the functions. The *mean distance from centroid* shown as the outermost of the two curves, is at a maximum towards the bottom right, which is as expected, and the *variance of distance from centroid*, shown as the innermost curve, is at a minimum in the top left quadrant, which is also as expected since the perimeter segments are much more densely packed there than those of the bottom right quadrant.



Data



(a)

(b)

Figure 39 Mean and variance polar functions in operation

5.6 Synthetic Shape Recognition

As a preliminary check on the viability of the polar function technique, a test was carried out using just the first of the four statistical measures, the *mean distance from centroid* function given by Equation 38. The test involved attempting to recognise synthetic shapes, in the form of a set of pentominoes, at different scales and rotations. There are only 12 distinct shapes that can be formed by placing 5 squares together so that each square always shares at least one side with another square. The 12 shapes are shown in Figure 40 where it can be seen that, with 45° rotations in some cases, they resemble the alphabet letters *f i l p n t u v w x y z*.

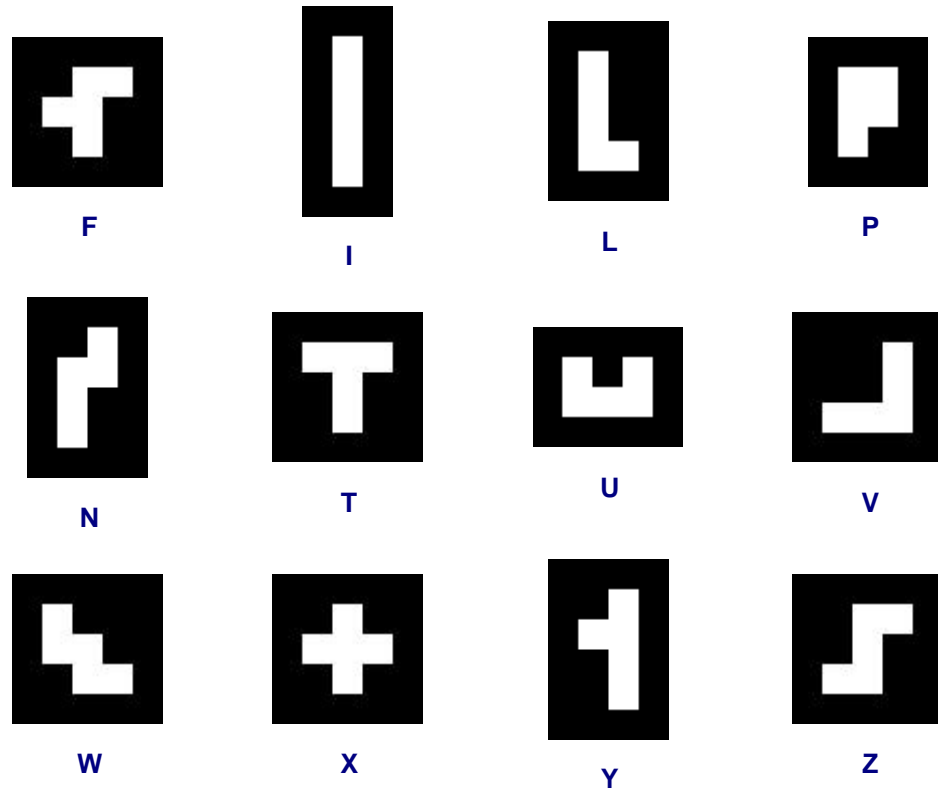


Figure 40 Twelve basic pentominoe shapes

Including the scaled and rotated variations, there are a total of 48 images consisting of 4 versions of each of the 12 basic shapes. Figure 41 shows the four variations of the shape corresponding to the letter 'F'. Figure 41(a) is a 75×75 pixel image with 0° rotation, and this basic template is repeatedly

rotated by 90° , and its scale increased in 25×25 pixel increments, to generate the other 3 images, so that Figure 41(d) has a scale of 150×150 pixels with a rotation of 270° .

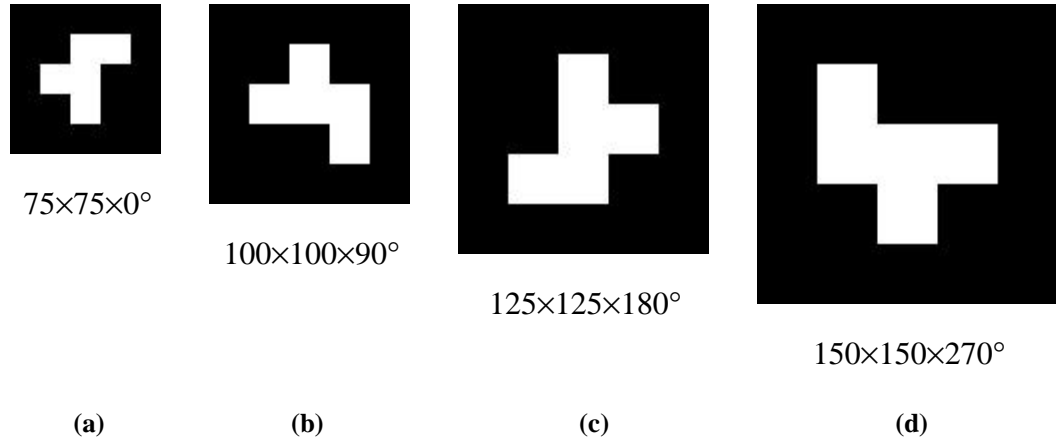


Figure 41 Scaling and rotation of the letter 'F'

The set of 48 shapes are transformed to obtain force and energy fields. The force field magnitudes for the first four, corresponding to the variations of the letter 'F', are shown in Figure 42. The pseudo edge detection effect is very evident in these images; this is because pixel intensity is constant both inside and outside the perimeters of the basic shapes, thus allowing homogeneous cancellation of forces to have maximum effect.

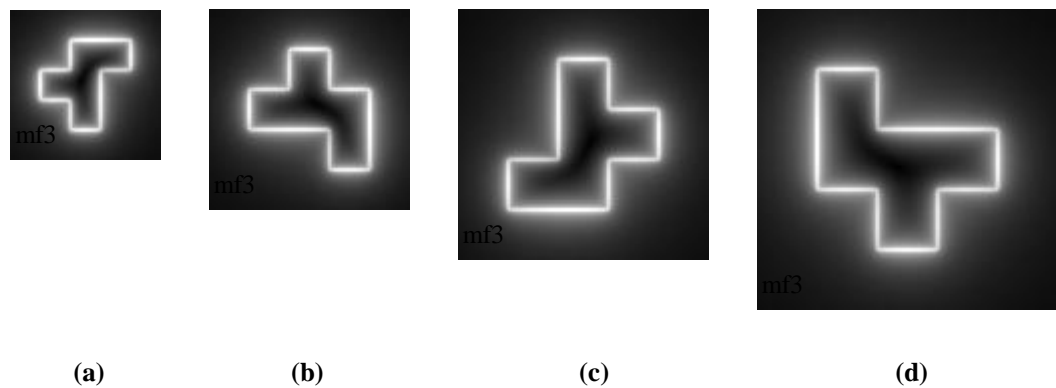


Figure 42 Force field magnitudes corresponding to the letter 'F'

The force fields are finally transformed to obtain their convergence fields as shown in Figure 43. Notice also that the corner detection property of the convergence functions is very evident here. The convex corners appear as isolated peaks of positive convergence, whereas the concave corners appear as isolated peaks of negative convergence. Notice also that the channels, which appear as ridges in the function value, take on the appearance of medial axes.

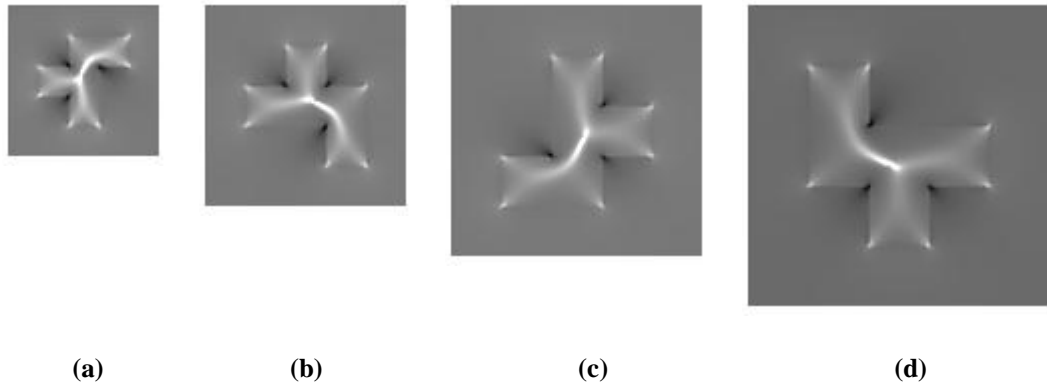


Figure 43 Convergence fields corresponding to the letter 'F'

The convergence fields are thresholded, setting to zero all values more than one standard deviation below the mean value, and converted to 60-point *mean distance from centroid* polar functions. Figure 44 shows the results for the letter ‘F’. Sets of concentric distance rings have been included to assist in verifying that the results have the correct relative scales. We see for example that the maximum mean distance in Figure 44(d) is about 50 pixels compared with a maximum value for Figure 44(a) of about 25 pixels, correctly matching the 2:1 size ratio of the figures shown in Figure 41 (d) and (a). We also see that the graphs in Figure 44 are very similar in form and that they have the correct relative rotations of 90° between successive graphs.

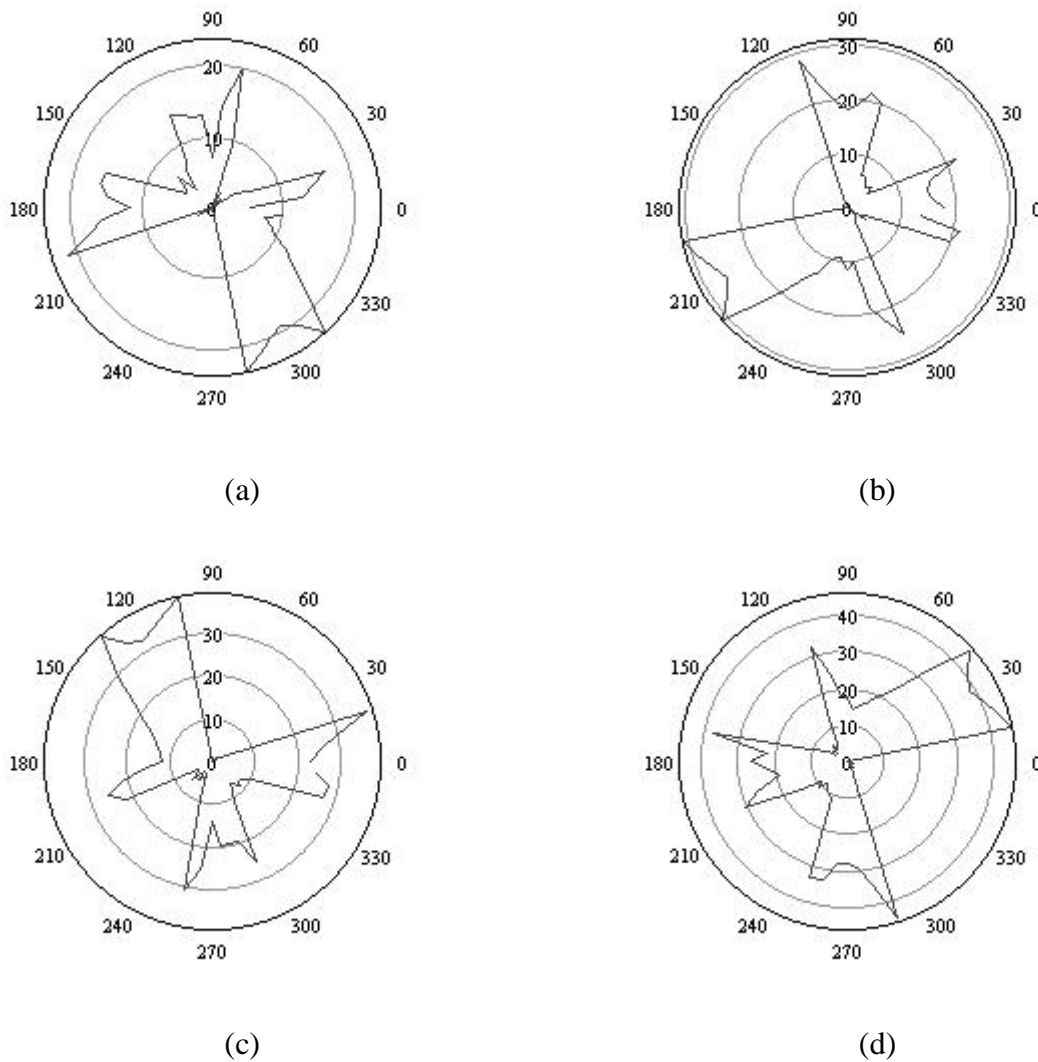


Figure 44 Mean distance polar functions corresponding to the letter ‘F’

The polar functions are then Fourier transformed to introduce a measure of rotation invariance. By taking the magnitude of the spectral coefficients of the transformed data, we take advantage of the inherent shift invariance property of the Fourier Transform.

Finally we seek to compare the results in a scale invariant manner, and we can do this easily by treating the spectral coefficient magnitudes as vectors, and comparing the pseudo angles between them. Using the Cauchy-Schwartz inequality,

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\| \quad (42)$$

the pseudo-angle between any two vectors \mathbf{u}, \mathbf{v} in \mathfrak{R}^n can be defined as

$$\theta = \cos^{-1} \frac{|\mathbf{u} \cdot \mathbf{v}|}{\|\mathbf{u}\| \|\mathbf{v}\|} \quad (43)$$

The argument of the inverse cosine function varies between 0 corresponding to the orthogonal case of maximum vector difference, and 1 for maximum similarity. Scale invariance is ensured since the angle is unaffected by the magnitude of the vector.

The pseudo angles of the resulting vectors are exhaustively cross-compared and the corresponding results for nearest neighbour classification are shown in Figure 45, where we see that 100% correct classification has been achieved. As before, the self-comparisons along the main diagonal are excluded.

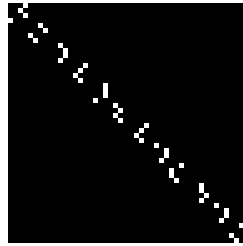


Figure 45 Nearest neighbour classification matrix

Table 6 shows the corresponding intra-class and inter-class mean and standard deviation statistics, based on pseudo angle frequency distributions, where once again the self-comparisons have been excluded from the calculations.

Intra-class		Inter-class	
Mean	Std. Deviation	Mean	Std. Deviation
0.082	0.049	0.521	0.19

Table 6 Statistics for intra-class and inter-class angle comparisons

Using these results, we may also compute the decidability index d' as previously done using Equation 34. Substituting the mean and standard deviation from Table 6 we have

$$d' = \frac{|0.082 - 0.521|}{\sqrt{((0.049^2 + 0.19^2)/2)}} = 3.164 \quad (44)$$

which is a marked improvement on the decidability index of 2.1 obtained for template matching in the ear recognition experiment.

Figure 46 shows the frequency distributions for the intra-class and inter-class pseudo angles. Normal distribution curves corresponding to their respective means and standard deviations have been fitted, where we see that the intra-class distribution is normally distributed, but we see that the normal distribution corresponding to inter-class frequency distribution does not fit quite as well.

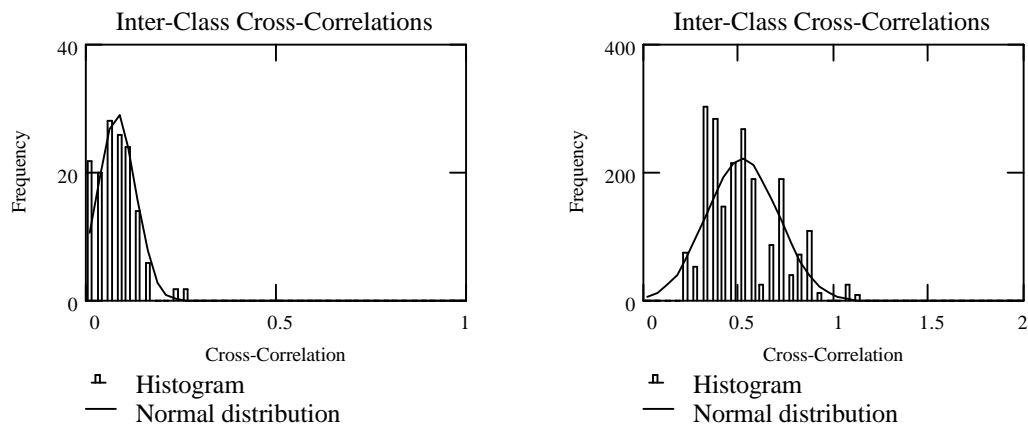


Figure 46 Normal distribution fitted to frequency distributions

However, it is still useful to compute the corresponding probability density functions, and these are shown in Figure 47, where we see that there is good separation and little overlap, confirming the improvement in the decidability index.

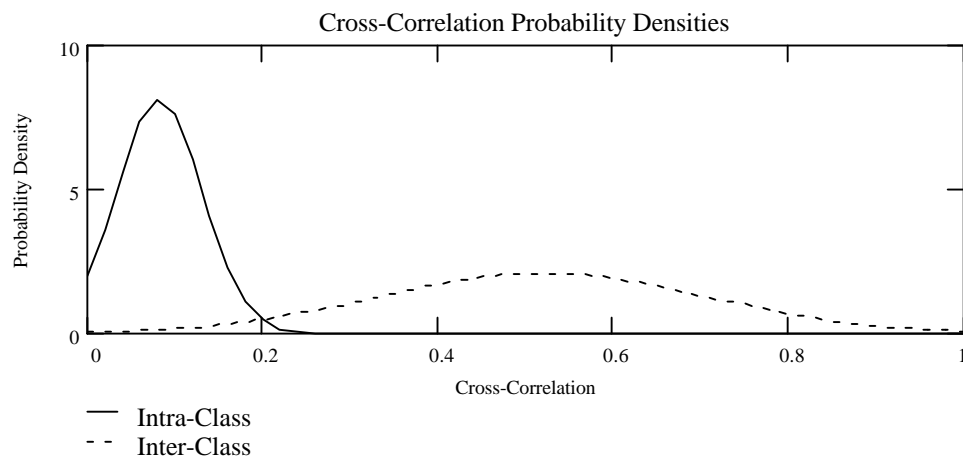


Figure 47 Probability density functions for pentominoe recognition

Finally, we may also compute the receiver operator characteristic for the pentominoe recognition experiment, where we see an improved equal error rate of 3.5% occurring at an angle of 0.17 radians.

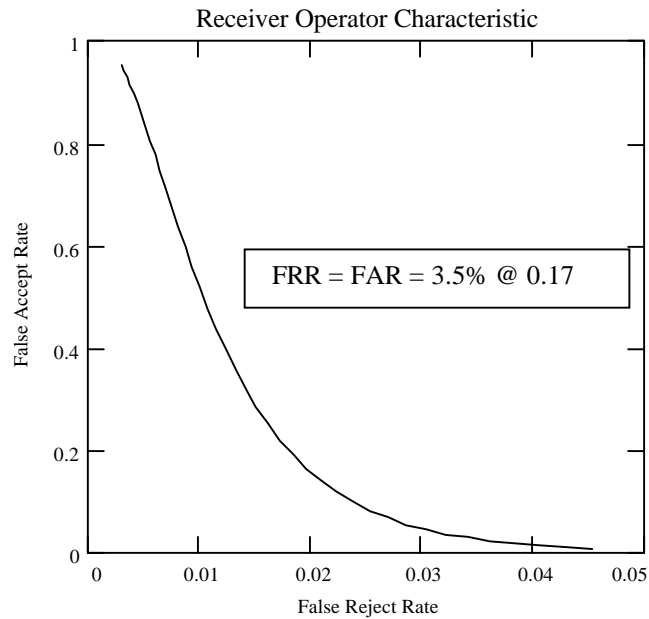


Figure 48 Receiver operator characteristic for pentominoe recognition

When the same experiment, using just the *mean distance from centroid*, was carried out on the ear database, the results were inferior to those obtained with template matching. This probably reflects the fact that real life ears are a lot more difficult to differentiate than synthetic pentominoes. Further work to incorporate the higher moments may lead to an improvement.

6 Conclusions

The purpose of this section is to bring together the main conclusions and results of the research. In a nutshell we may say that in the context of ear biometrics we have developed a new linear transform that transforms an ear image, with very powerful smoothing and without loss of information, into a smooth dome shaped surface whose special shape facilitates a new form of feature extraction that extracts the essential ear signature without the need for explicit ear extraction; and in the process we have verified the recognition potential of the human ear for biometrics. Conclusions about the three distinct aspects of this research will now be considered separately in greater detail

6.1 *Force Field Transform*

The powerful smoothing inherent in the transform is provided by virtue of the fact that each point in the transformed output is a function of all the other points in the original image, which amounts to very large scale weighted averaging, where the weighting is supplied by the inverse $1/r$ function in the case of the potential transform. Yet this powerful smoothing is achieved with no loss of information by virtue of the fact that the potential transform is invertible, implying that the original information can be recovered in principle from the energy surface. The dome shape of the energy surface comes about because points towards the centre of the surface have a shorter mean distance to all their sources of energy than points that are further away. It is this basic dome shape underlying the energy field that gives the transform its autocentring property, which is the reason that feature extraction without the need for explicit ear extraction is possible.

The computation is very time consuming even when carried out using fast Fourier techniques, but it is interesting to speculate that since it is modeled on a natural force field, that it may be possible to implement the transform in hardware by mapping the image to an array of electric charges. Also, because

the formulation is couched in terms of vectors, it may be possible to generalise to higher dimensional problems.

6.2 Feature Extraction

We argue that because the energy surface has preserved all the information in the original image, and because channels and wells modulate its otherwise smooth dome shape, therefore these should make good features. Two equivalent forms of feature extraction have been developed. One is something akin to the popular gradient descent algorithm, where instead of seeking just one extremum, a host of exploratory text pixels seek out all the potential wells on the surface; the channels formed in the process becoming the force field features. This method has the advantage of producing field lines that take the form of integral data objects in the form of a list of points visited, which may have advantages for further processing. For example, a channel might be partly defined by selecting the field line that makes the greatest contribution to the channel.

The convergence method of feature extraction takes the form of a function whose ridges and peaks correspond exactly to the field line channels and wells, but this function provides an extra tier of information in the form of antichannels and antiwells. Another important property is that convergence centroids tend to be quite stable with respect to the features, thus affording a way of aligning the features for comparison, or perhaps for other purposes.

We observe the similarity in form of the convergence operator to the Marr-Hildreth operator, which may also be viewed as a linear shape operator due to its Laplacian of Gaussian nature. However, a very important difference is that the convergence operator is a non-linear operator due to its use of force direction rather than force itself.

6.3 Ear Recognition

A fairly rudimentary method of classification in the form of simple template matching was used just to establish a lower bound on the ear recognition capability. Even without using the template rotation facility we have achieved a recognition rate in excess of 91% on the small database of 252 samples from 63 subjects. This confirms that ears can certainly be used for recognition but does not really give a useful indication of the uniqueness of ears in general. Clearly if ears were to be used as the basis of a large scale biometric based application, then template matching would not be appropriate due to the sheer number of cross-correlations involved, so that some form of indexing scheme based on the features would have to be developed.

As such, a promising new technique has been developed that affords some interesting properties. We look forward to examining in future more of its capabilities, and extensions to it, as outlined in our further work.

Appendices

A1 *Alternative Proof of Transform Linearity*

We now give an alternative proof that the force field transform is a linear transformation by appealing to the basic definition of a linear transformation.

Let V and W be vector spaces.

A function $t: V \rightarrow W$ is a **linear transformation** if it satisfies the following two relations: -

$$\text{Relation 1: } t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2), \text{ for all } \mathbf{v}_1, \mathbf{v}_2 \in V \quad (1)$$

$$\text{Relation 2: } t(\alpha \mathbf{v}) = \alpha t(\mathbf{v}), \text{ for all } \mathbf{v} \in V, \alpha \in \mathfrak{R} \quad (2)$$

In vector space V , let $\mathbf{v}_1 = \mathbf{A}$ and $\mathbf{v}_2 = \mathbf{B}$ where \mathbf{A} and \mathbf{B} are $M \times N$ matrices whose elements $A_{i,j}$ and $B_{i,j}$ are real numbers representing pixel intensities. Also, let W be the vector space of $M \times N$ matrices where the elements are force vectors as defined by the relation

$$[t(\mathbf{A})]_{i,j} \mapsto \begin{cases} 0, & (m,n) = (i,j) \\ \left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{m,n} \frac{\mathbf{r}_{m,n} - \mathbf{r}_{i,j}}{|\mathbf{r}_{m,n} - \mathbf{r}_{i,j}|^3} \right), & (m,n) \neq (i,j) \end{cases} \quad (3)$$

Now to check that relation 1 is satisfied, we need to show that

$$t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2), \text{ for all } \mathbf{v}_1, \mathbf{v}_2 \in V \quad (4)$$

That is, we need to show that the force field of the sum of two images is the same as the sum of the force fields associated with the individual images. Accordingly, for the sum of the separate force fields associated with the two images we have,

$$\begin{aligned} [t(\mathbf{A}) + t(\mathbf{B})]_{i,j} &= \left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{m,n} \frac{\mathbf{r}_{m,n} - \mathbf{r}_{i,j}}{|\mathbf{r}_{m,n} - \mathbf{r}_{i,j}|^3} \right) + \left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} B_{m,n} \frac{\mathbf{r}_{m,n} - \mathbf{r}_{i,j}}{|\mathbf{r}_{m,n} - \mathbf{r}_{i,j}|^3} \right) \\ &= \left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (A_{m,n} + B_{m,n}) \frac{\mathbf{r}_{m,n} - \mathbf{r}_{i,j}}{|\mathbf{r}_{m,n} - \mathbf{r}_{i,j}|^3} \right) \end{aligned} \quad (5)$$

Where the expression above gives the resultant force acting at a single pixel location with position vector $\mathbf{r}_{i,j}$ due to the attraction of all the other pixels at position vectors $\mathbf{r}_{m,n}$ with pixel intensities $(A_{m,n} + B_{m,n})$.

On the other hand, for the force field for the sum of two images we have,

$$[t(\mathbf{A} + \mathbf{B})]_{i,j} = \left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (A_{m,n} + B_{m,n}) \frac{\mathbf{r}_{m,n} - \mathbf{r}_{i,j}}{|\mathbf{r}_{m,n} - \mathbf{r}_{i,j}|^3} \right) \quad (6)$$

As such, the results of Equations (5) and (6) are equal so relation 1 is indeed satisfied.

Now to check relation 2, we require to show that

$$t(\alpha \mathbf{v}) = \alpha t(\mathbf{v}), \text{ for all } \mathbf{v} \in V, \alpha \in \Re \quad (7)$$

That is, we need to show that the force field of an image whose intensity has been scaled is equal to the force field of the original image multiplied by the same scale factor.

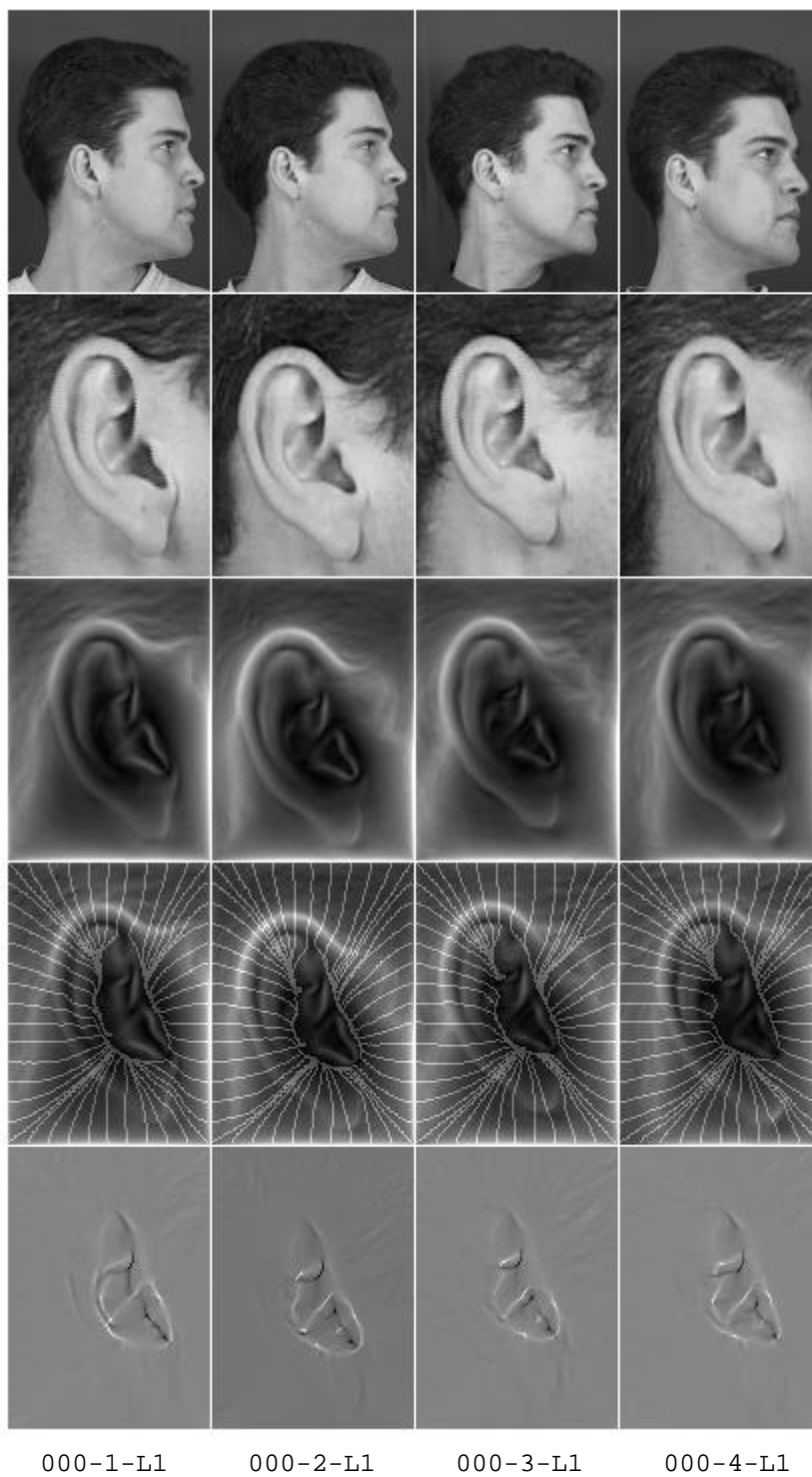
Accordingly we have

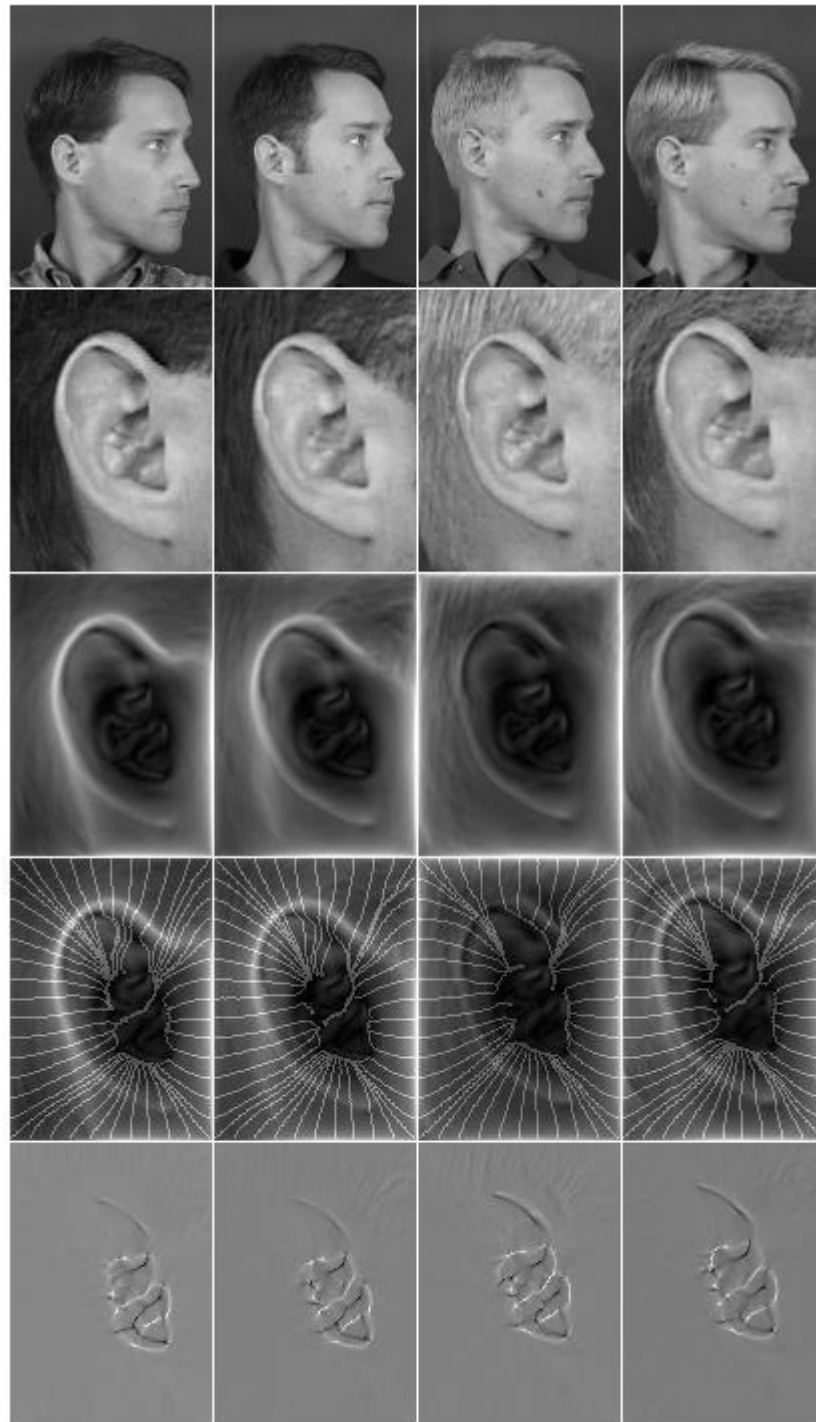
$$[t(\alpha \mathbf{A})]_{i,j} = \left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \alpha A_{m,n} \frac{\mathbf{r}_{m,n} - \mathbf{r}_{i,j}}{|\mathbf{r}_{m,n} - \mathbf{r}_{i,j}|^3} \right) = \alpha \left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{m,n} \frac{\mathbf{r}_{m,n} - \mathbf{r}_{i,j}}{|\mathbf{r}_{m,n} - \mathbf{r}_{i,j}|^3} \right) = [\alpha t(\mathbf{A})]_{i,j} \quad (8)$$

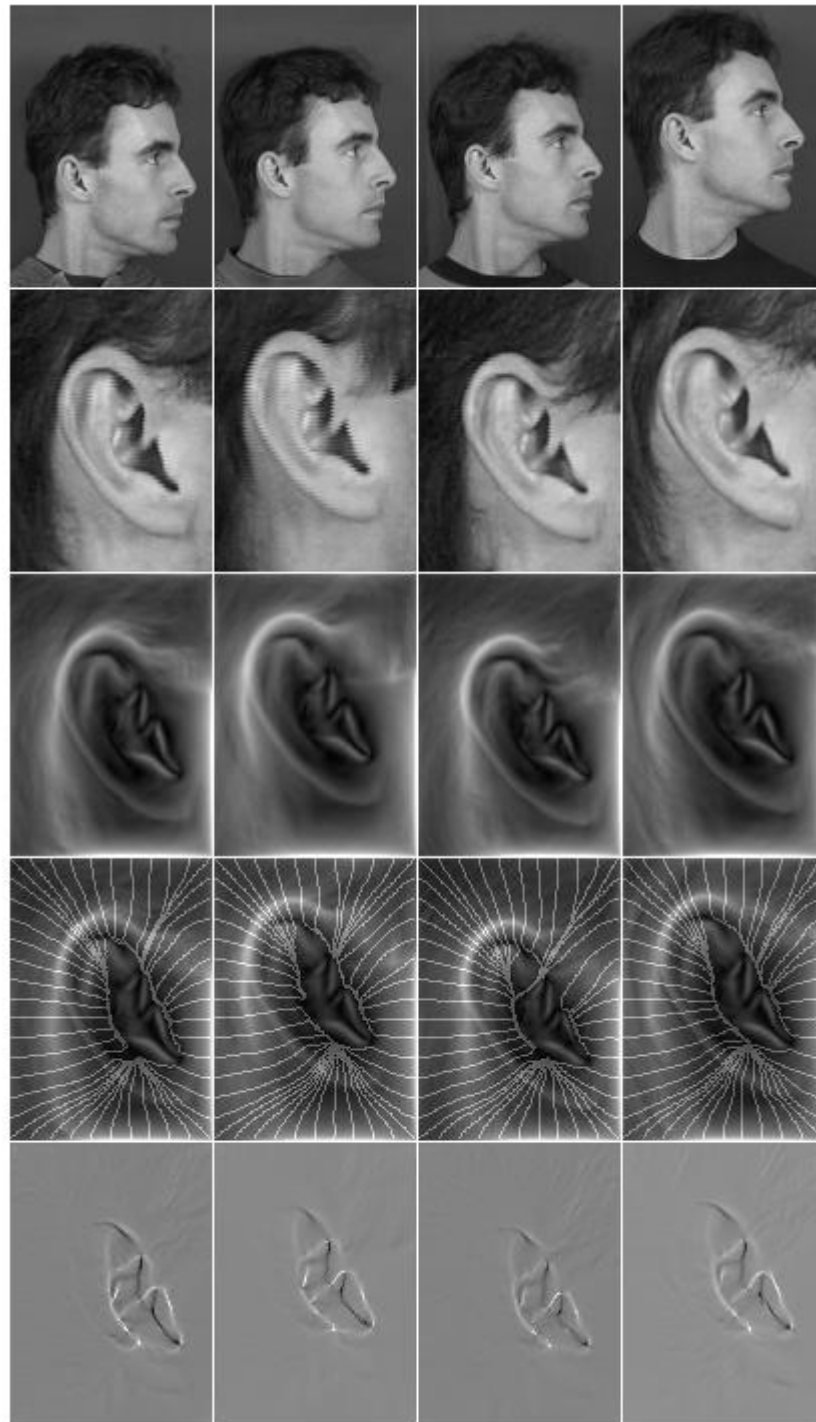
so relation 2 is satisfied.

Hence, since both relations 1 and 2 are satisfied, the force field transform is indeed a linear transformation.

A2 *Representative Sample of Ear Results*





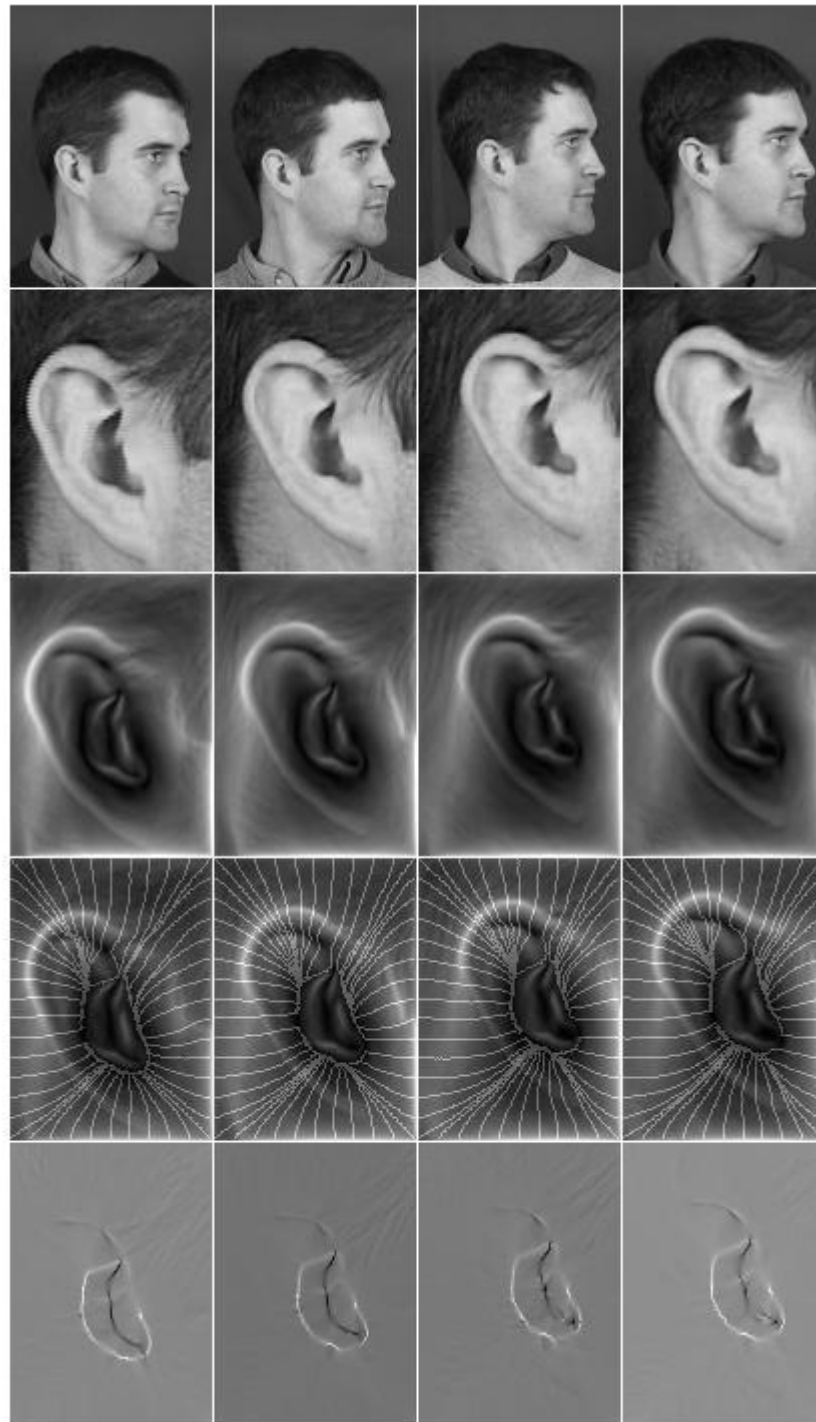


004-1-L1

004-2-L1

004-3-L1

004-4-L1

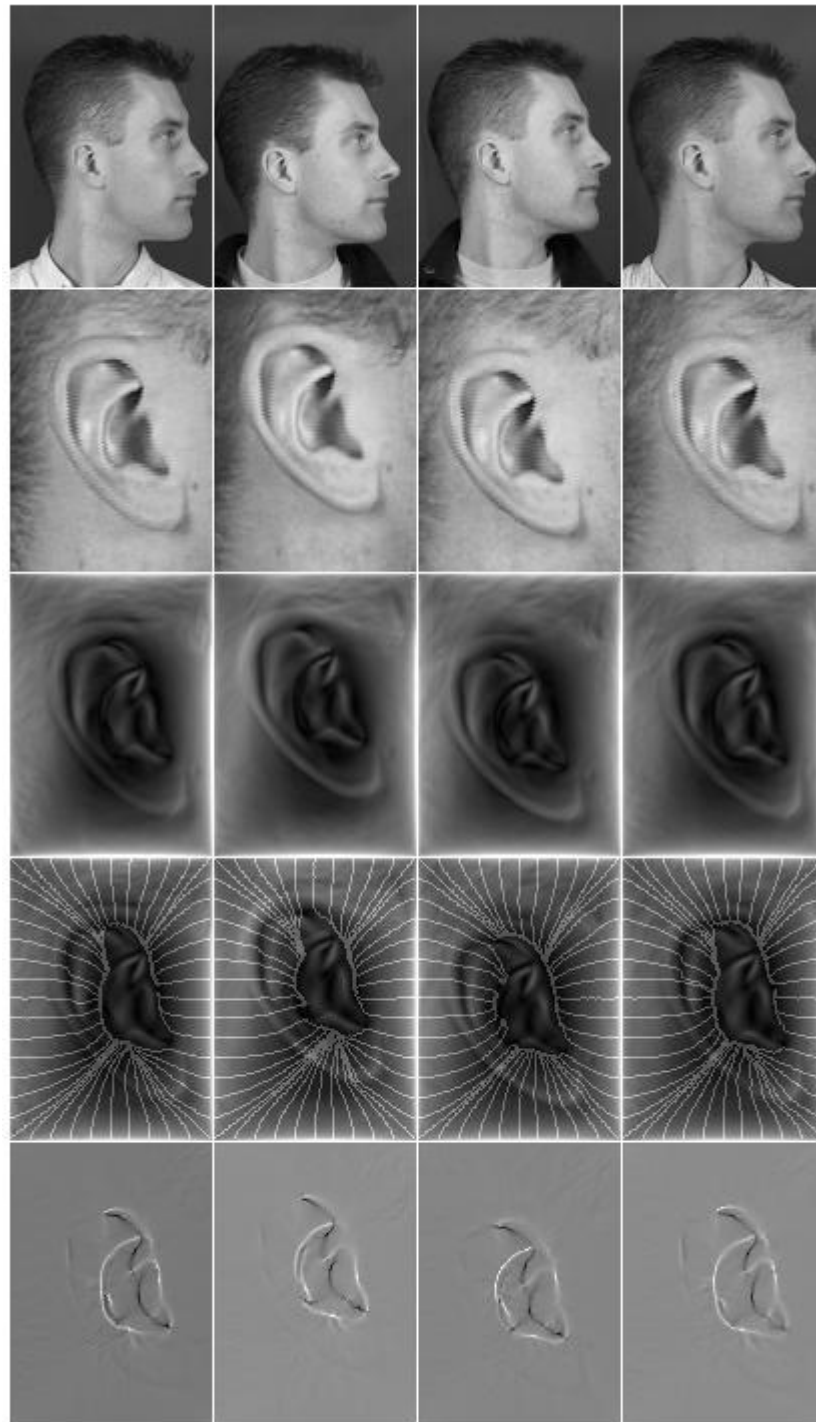


008-1-L1

008-2-L1

008-3-L1

008-4-L1

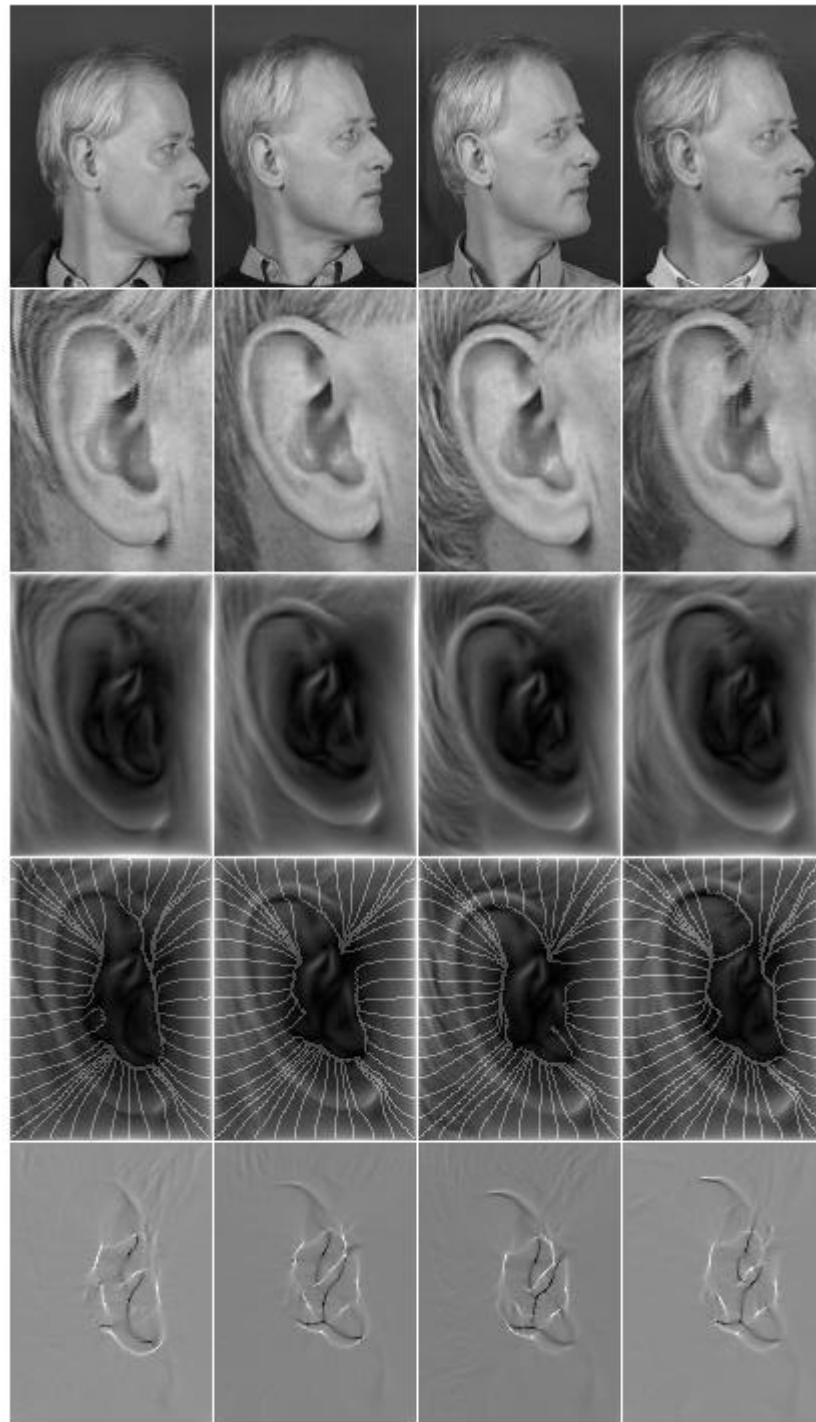


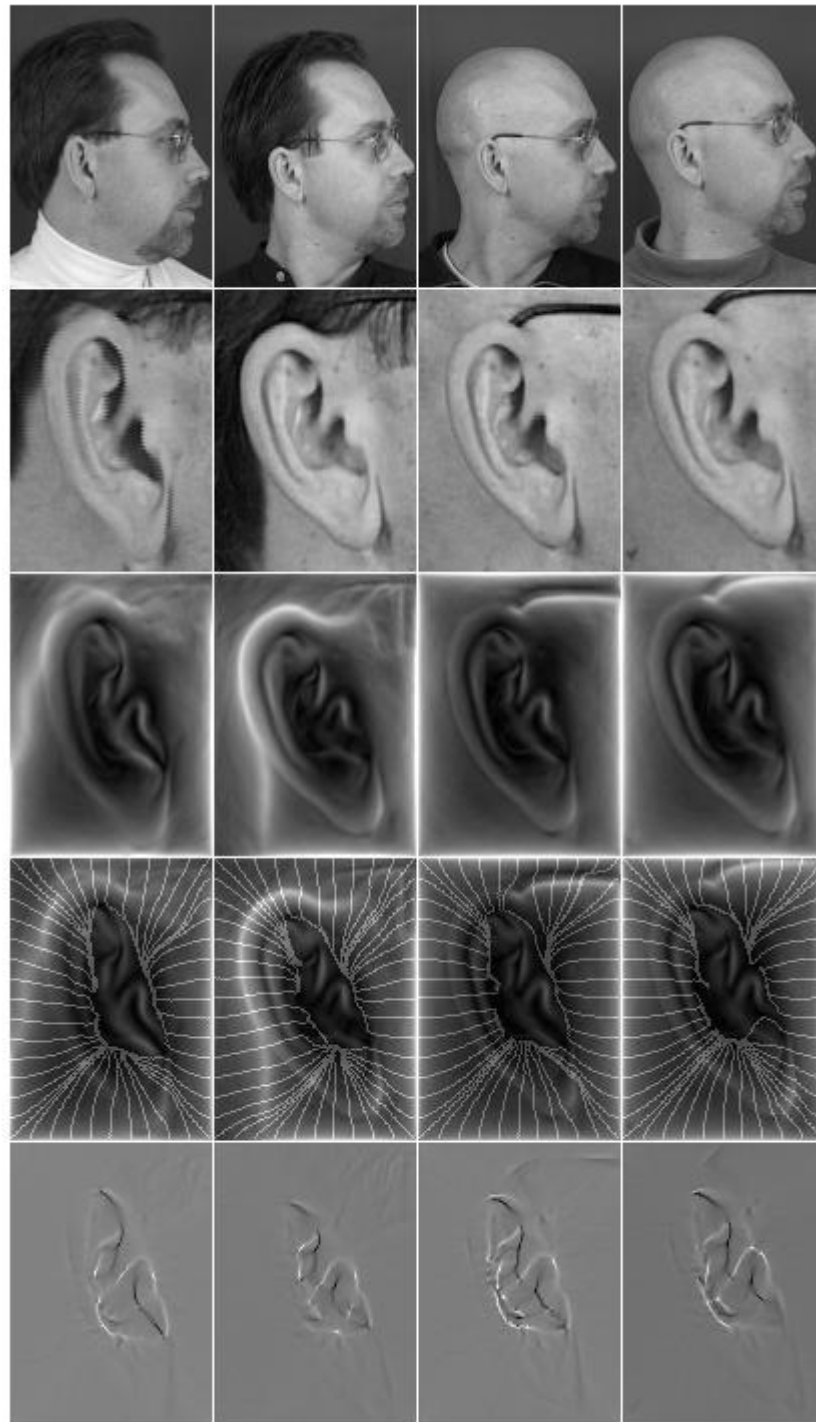
010-1-L1

010-2-L1

010-3-L1

010-4-L1



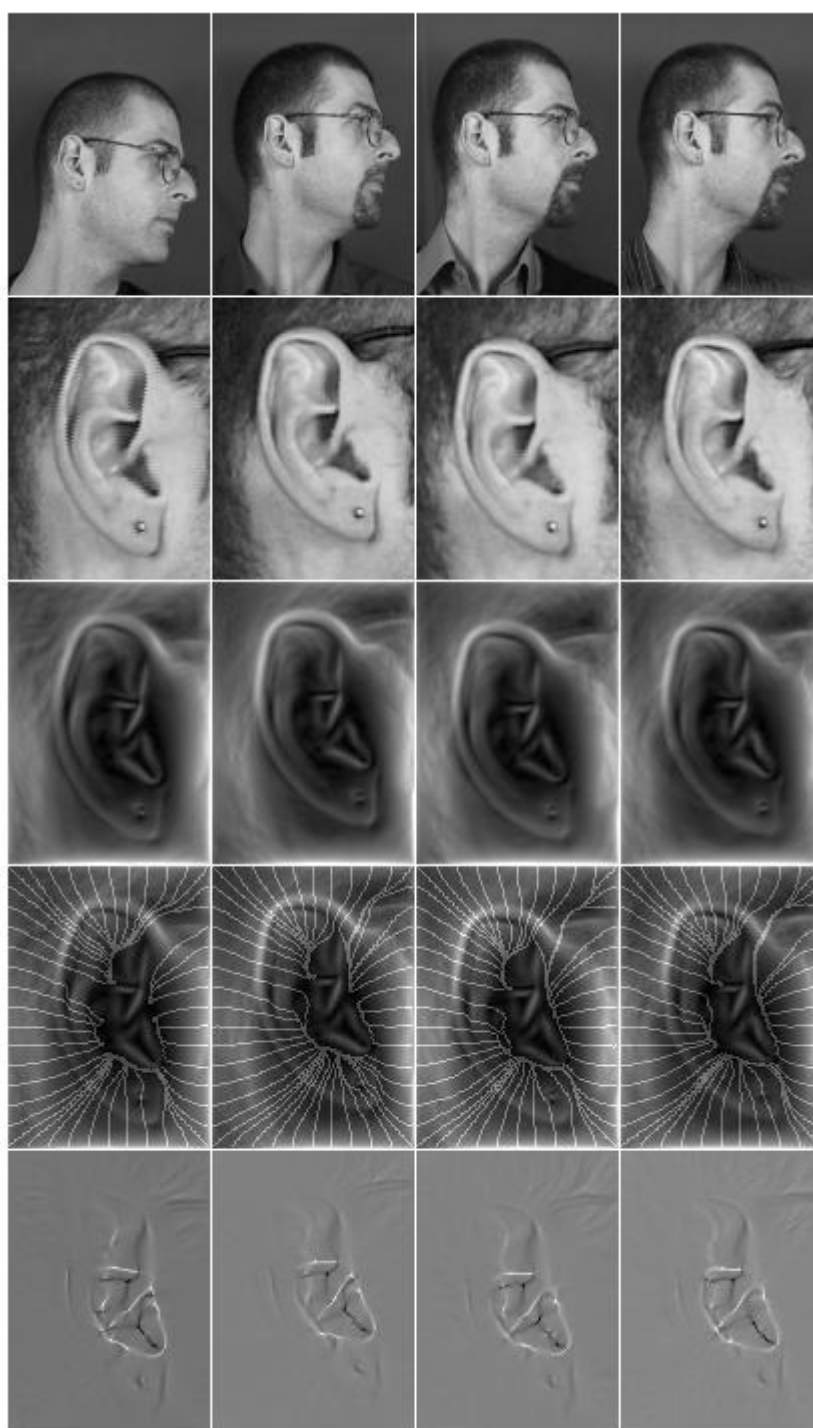


016-1-L1

016-2-L1

016-3-L1

016-4-L1

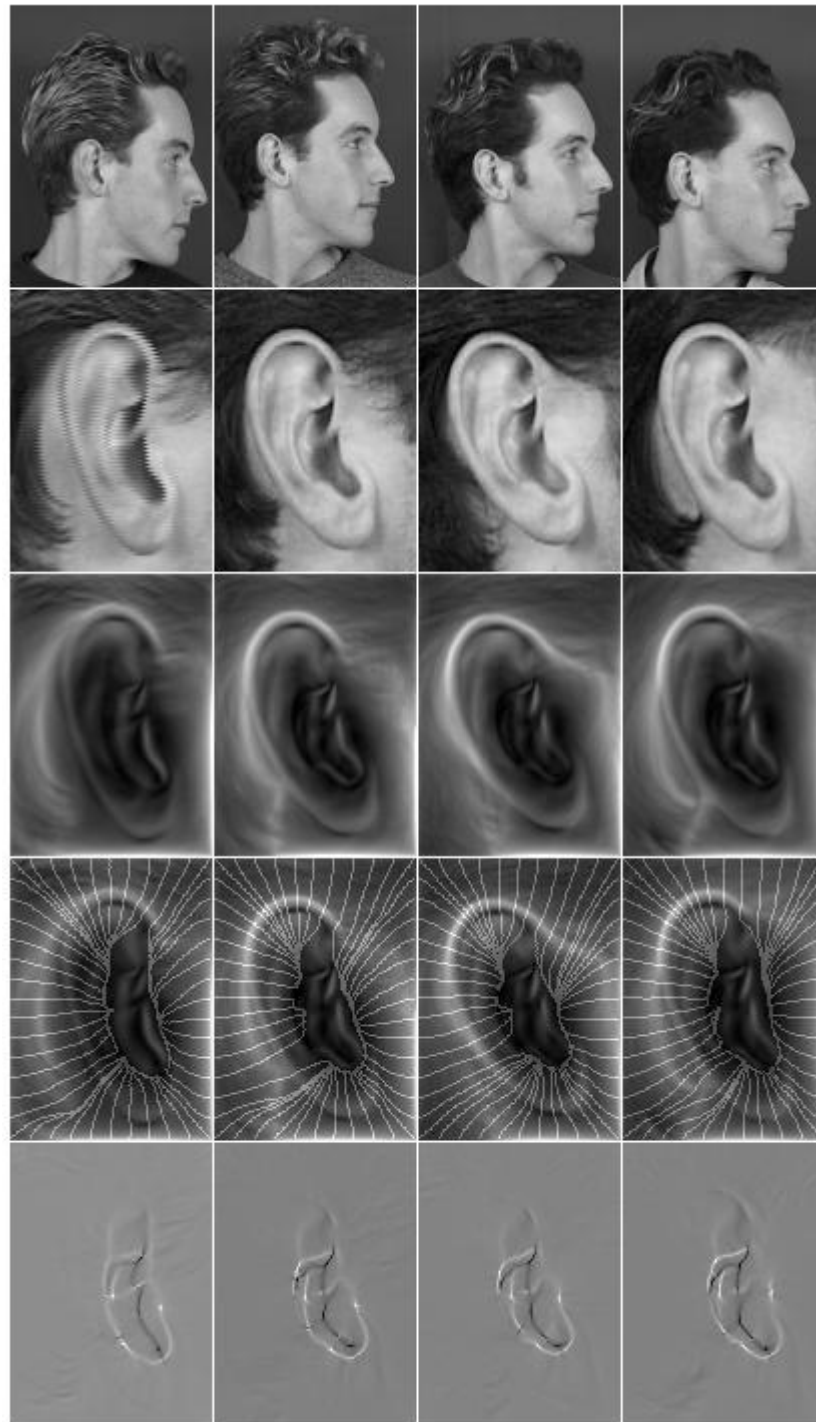


019-1-L1

019-2-L1

019-3-L1

019-4-L1

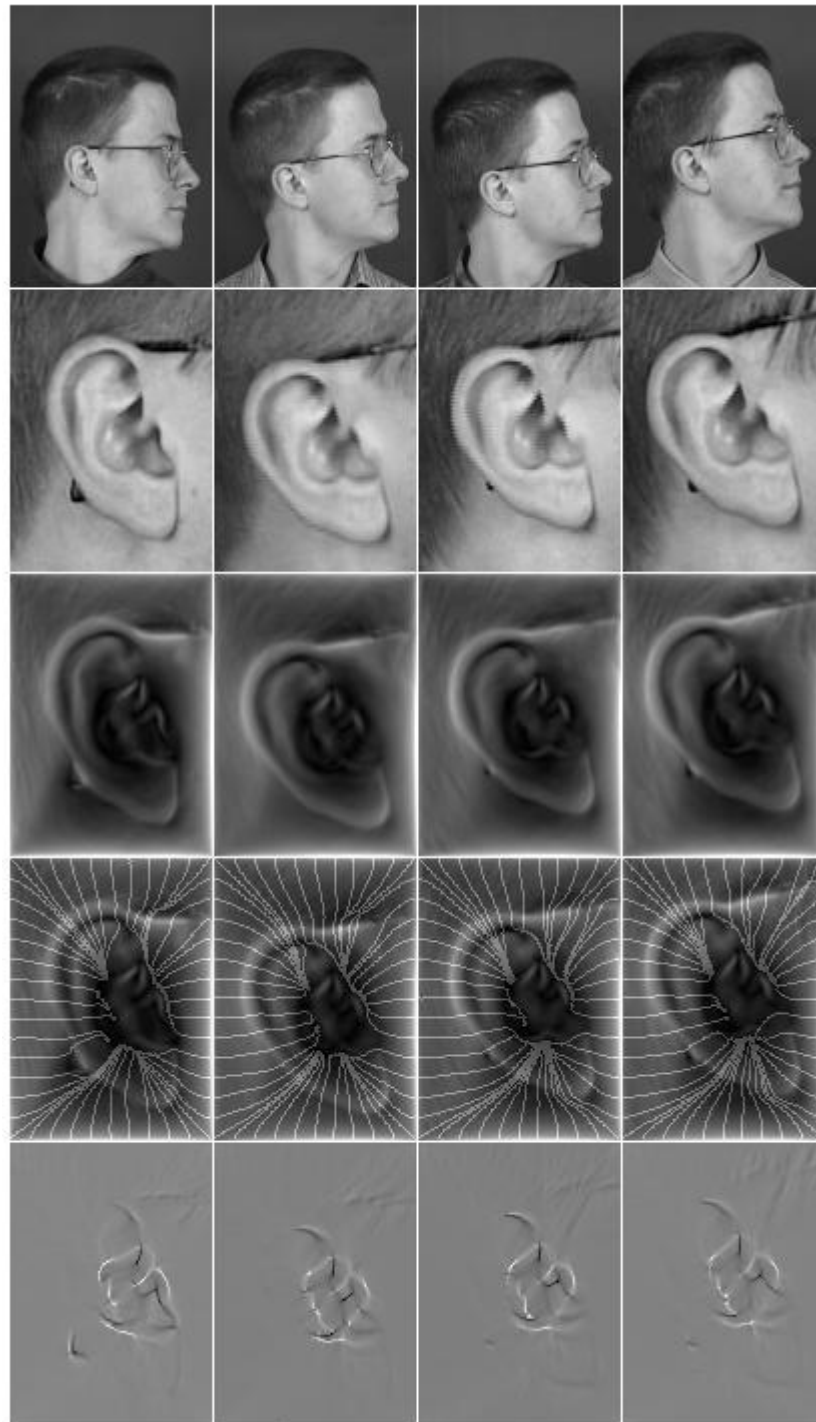


021-1-L1

021-2-L1

021-3-L1

021-4-L1

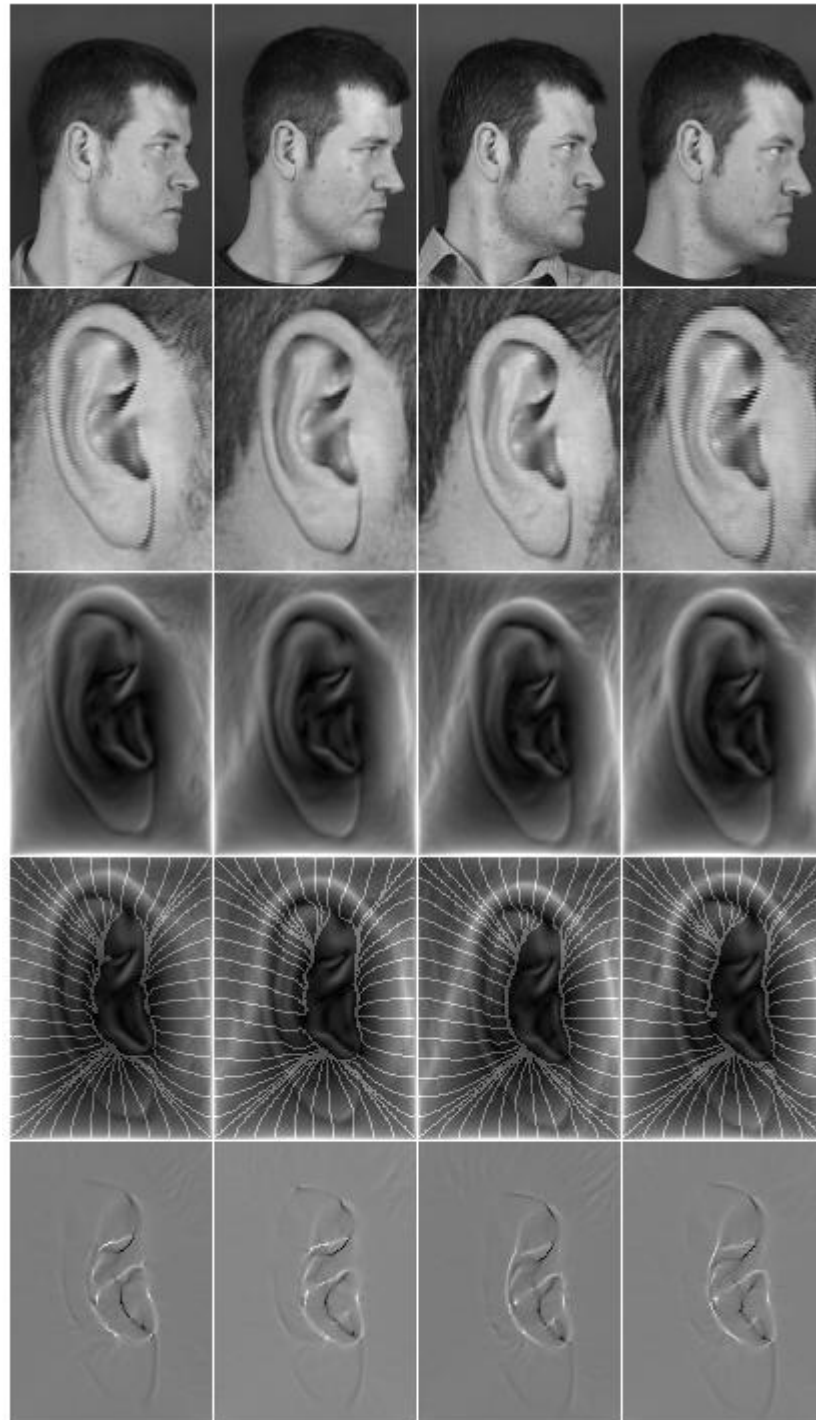


026-1-L1

026-2-L1

026-3-L1

026-4-L1



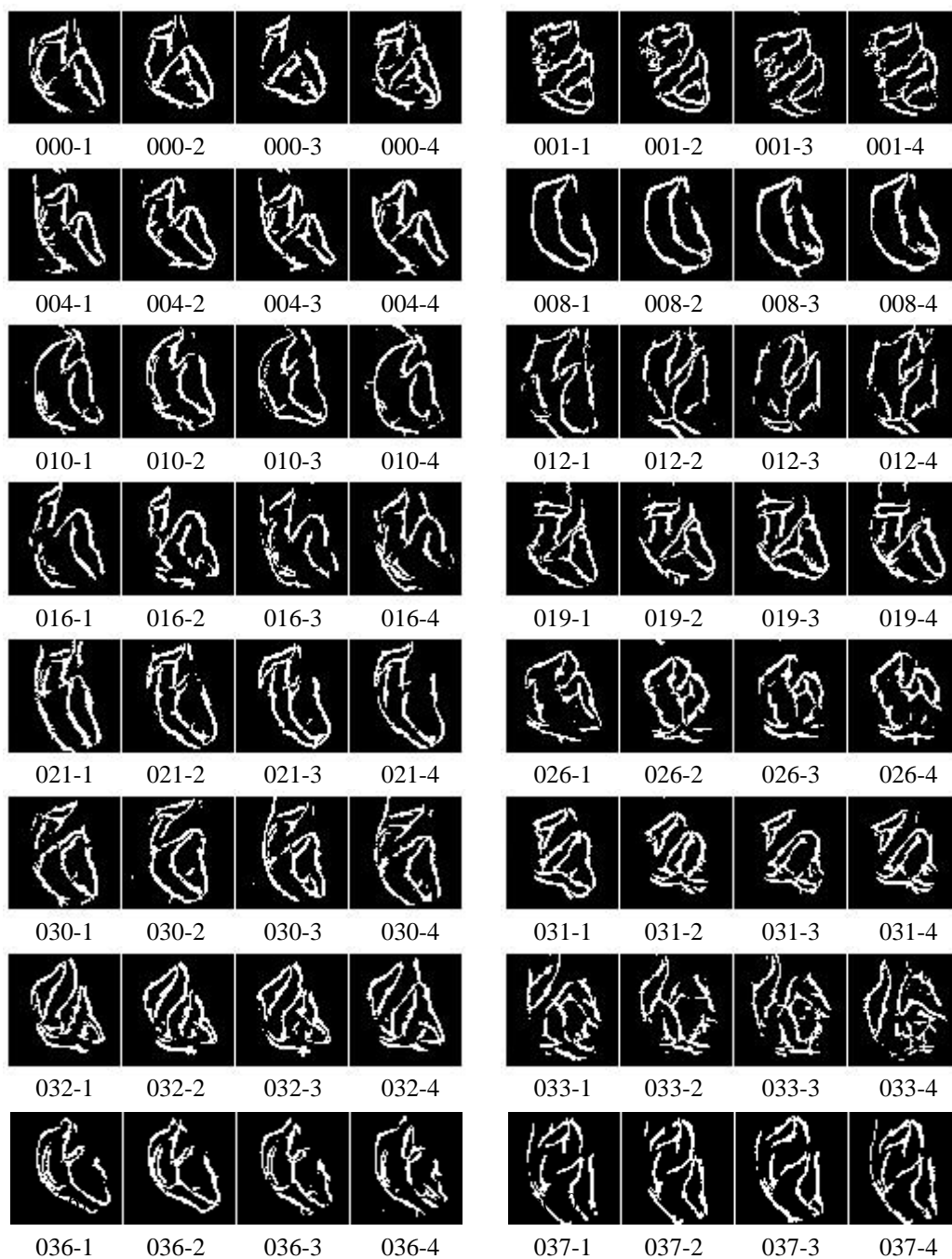
030-1-L1

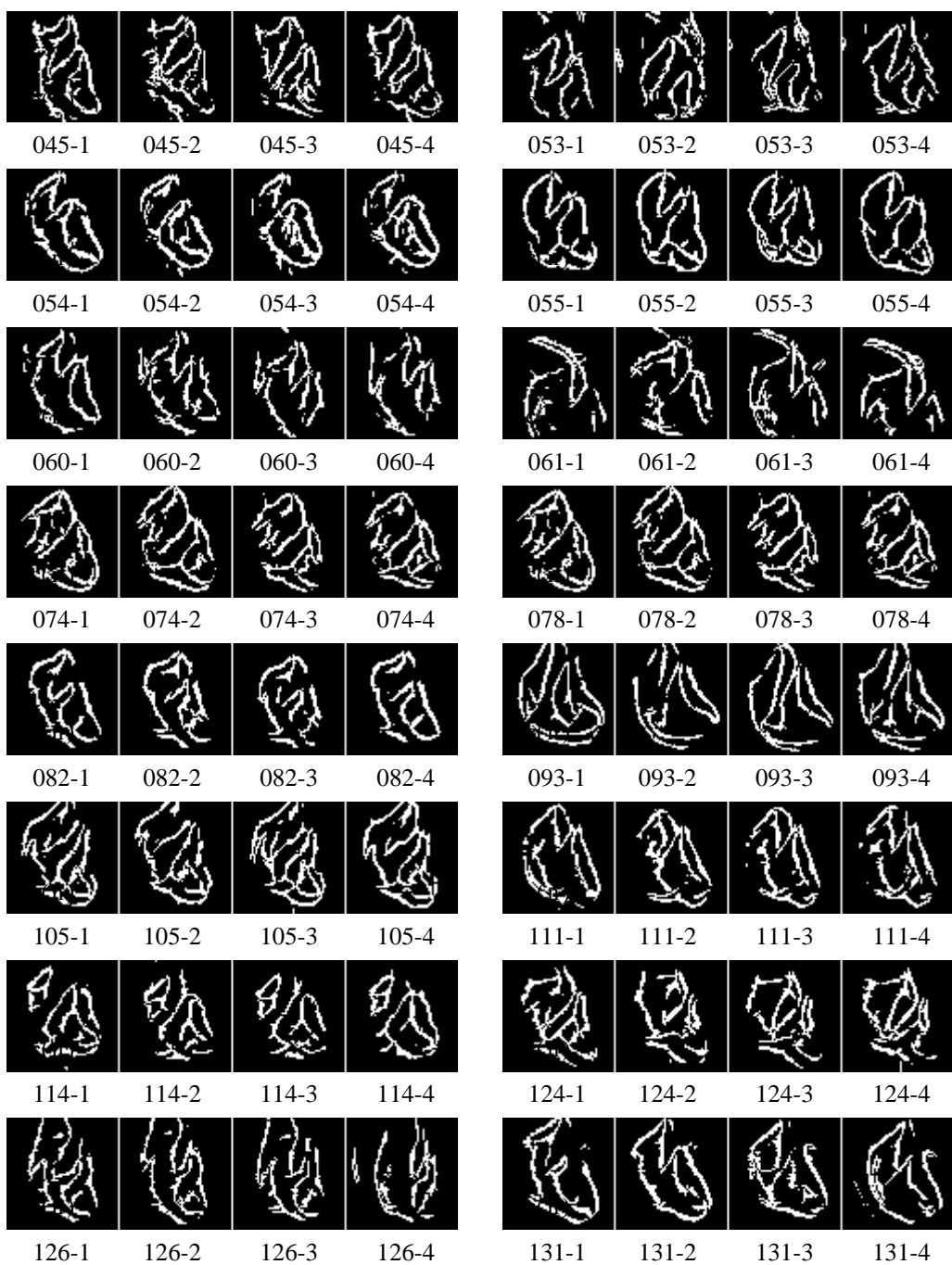
030-2-L1

030-3-L1

030-4-L1

A3 *Binarised Thresholded Convergence Functions*









231-1 231-2 231-3 231-4



246-1 246-2 246-3 246-4



259-1 259-2 259-3 259-4



285-1 285-2 285-3 285-4



288-1 288-2 288-3 288-4



310-1 310-2 310-3 310-4



324-1 324-2 324-3 324-4



371-1 371-2 371-3 371-4



233-1 233-2 233-3 233-4



248-1 248-2 248-3 248-4



274-1 274-2 274-3 274-4



287-1 287-2 287-3 287-4



293-1 293-2 293-3 293-4



314-1 314-2 314-3 314-4



360-1 360-2 360-3 360-4

A4 XM2VTS Identification Numbers of Ear Subjects

000	030	060	126	174	246	324
001	031	061	131	175	248	360
004	032	074	132	180	259	371
008	033	078	142	188	274	
010	036	082	143	201	285	
012	037	093	145	206	287	
016	045	105	147	216	288	
019	053	111	153	221	293	
021	054	114	166	231	310	
026	055	124	171	233	314	

Table 7 XM2VTS identification numbers of ear subjects

References

- [1] Andrews, H. C., *Introduction to Mathematical Techniques in Pattern Recognition*, John Wiley & Sons Inc., p. 9, 1972
- [2] Lynch, C., *Ear-prints provide evidence in court*, Glasgow University News, November 22, 2000
- [3] STATE v. David Wayne KUNZE, Court of Appeals of Washington, Division 2. 97 Wash.App. 832, 988 P.2d 977, 1999
- [4] Iannarelli, A., *Ear Identification*, Paramount Publishing Company, Freemont, California, 1989
- [5] BBC Online News, *Ear print catches murderer*, Tuesday, December 15, 1998
- [6] BBC Online News, *'Ere son - you're nicked!*, Friday, February 20, 1998
- [7] Immigration and Naturalisation Service, U.S. Department of Justice, *Application for Naturalisation Form N-400*
- [8] Allen, D. E., *Anna or Anastasia? Part 2*, Department of Biological Sciences, University of Delaware, <http://ublib.buffalo.edu/libraries/projects/cases/Anna.html>
- [9] Burge, M., Burger, W., Ear Biometrics, in: Jain, A., Bolle, R., Pankanti, S.: *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, pp. 273-286, 1999
- [10] Cole S. A., *Suspect Identities: A history of fingerprinting and criminal identification*, Harvard University Press, 2001

- [11] BBC Online News, *Fingerprint experts 'making mistakes'*, Monday, 23 October, 2000
- [12] Moenssens, A. A., *Ear Identification Research Proposal*, Forensic-Evidence.com, http://forensic-evidence.com/site/ID/ID00004_4.html
- [13] Bertillon, A., *La photographie judiciaire, avec un appendice sur la classification et l'identification anthropométriques*, Gauthier-Villars, Paris, 1890.
- [14] Kass, M., Witkin, A., Terzopoulos, D., Snakes: Active Contour Models, *International Journal of Computer Vision*, **1**, pp. 321-331, 1988
- [15] Blake, A., Isard, M., *Active Contours*, Springer-Verlag, London, 1998
- [16] Gunn, S. R., and Nixon, M. S., A Robust Snake Implementation; A Dual Active Contour, *IEEE Transactions on PAMI*, **19**(1), pp. 63-67, 1997
- [17] Marr D., Hildreth E. C., Theory of edge detection. *Proceedings of the Royal Society of London B*, **207**, pp. 187-217, 1980
- [18] Canny, J., A computational approach to edge detection, *IEEE Transactions on PAMI*, **8**(6), pp. 679-698, 1986
- [19] Spacek, L. A., Edge Detection and Motion Detection, *Image and Vision Computing*, **4**(1), pp. 43-56, 1986
- [20] Morrone, M. C., Burr, D. C., Feature detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society of London B*, **235**, pp. 221-245, 1988.
- [21] Morrone, M. C., Owens, R. A., Feature detection from local energy, *Pattern Recognition Letters*, **6**, pp. 303-313, 1987.

- [22] O'Neill, B., *Elementary Differential Geometry* – 2nd ed, Academic Press, London and San Diego, CA, 1997
- [23] Prokop R. J., Reeves, A. P., A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP: Graphical Models and Image Processing*, **54**, pp. 438-460, 1992
- [24] Horn B. K. P., *Robot Vision*. Graw-Hill, New York, 1986.
- [25] Luo, B., Cross, A. D., Hancock, E. R., Corner Detection Via Topographic Analysis of Vector Potential, *Pattern Recognition Letters* **20**(6), pp. 635-650, 1999
- [26] Ahuja, N., Chuang, J. H., Shape Representation Using a Generalised Potential Field Model, *IEEE Transactions on PAMI*, **19**(2), pp. 169-176, 1997
- [27] Xu, C., Prince, J. L., Gradient Vector Flow: A New External Force for Snakes, *Proc. IEEE Conf. on Comp. Vis. Patt. Recog. (CVPR)*, pp. 66-71, 1997
- [28] Messer, K., Matas, J., Kittler, J., Luetttin, J., and Maitre, G., XM2VTSDB: The Extended M2VTS Database, *Proc. AVBPA '99* Washington D.C., 1999
- [29] Faulds, H., On The Skin-Furrows Of The Hand, *Nature*, **22**(574), p. 605, 1880
- [30] Quetelet, A., *Anthropometrie ou Mesure des differentes facultes de l' homme*, C. Muquardt, Bruxelles, 1871.
- [31] Herschel, W. J., Skin Furrows of the Hand. *Nature*, **23**(578), p.76, 1880

[32] Bertillon, A., *Identification anthropometrique; instructions signaletiques*. Impr. administrative, Melun, 1893.

[33] Bertillon, A., *Ethnographie moderne*, G. Masson, Paris, 1883.

[34] Bertillon, A., and McClaughry, R.W., *Signaletic instructions including the theory and practice of anthropometrical identification*, The Werner company, Chicago, New York, 1896.

[35] Galton, F., *Fingerprints*, Macmillan, London, 1892

[36] Galton, F., Personal Identification and Description, *Nature*, June 21, pp. 173-177 & pp. 201-202, 1888

[37] Galton, F., Human Variety, *Nature*, January 24, pp. 296-300, 1889

[38] Cherrill, F. R., *The Fingerprint System at Scotland Yard. a Practical Treatise on Finger Print Identification*, HMSO, London, 1954

[39] Bible: Hebrew, 1 Kings. 3:25-37.

[40] Rudin, N., Inman, K., Stolovitzky, G., Rigoutsos, I., DNA-Based Identification, in: Jain, A., Bolle, R., Pankanti, S. Eds.: *Biometrics: Personal Identification in Networked Society*, Kluwer Acad. Publishers, 1999, pp. 287-310.

[41] De Coras, J., *Arrest Memorable, du Parlement de Tolose*, Antoine Vincent, Lyon, 1561.

[42] Cho, A., *An Identity Crisis*, *Simon Cole Interview*, New Scientist, June 16, 2001

[43] Daubert v. Merrell Dow Pharmaceuticals (92-102), 509 U.S. 579 (1993).

[44] (USA) Electronic Signatures in Global and National Commerce Act, June 2001

[45] Directive 1999/93/EC of the European Parliament and of The Council of December 12th 1999 on a Community Framework for Electronic Signatures.

[46] (UK) Electronic Communications Act, HMSO, 2000

[47] Daugman, J., United States Patent No. 5,291,560 (issued in March 1994). *Biometric Personal Identification System Based on Iris Analysis*, Washington DC: U.S. Government Printing Office, 1994

[48] Daugman, J. D., High Confidence Visual Recognition of Persons by a Test of Statistical Independence, *IEEE Transactions on PAMI*, **15**(11), pp. 1148-1161, 1993

[49] Weng, J. and Swets, D. L., Face Recognition, in: Jain, A., Bolle, R., Pankanti, S. Eds.: *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, 1999, pp. 65-86.

[50] O’Gorman, L., Fingerprint Verification, in: Jain, A., Bolle, R., Pankanti, S. Eds.: *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, 1999, pp. 43-64.

[51] Campbell, J. P., Speaker Recognition, in: Jain, A., Bolle, R., Pankanti, S. Eds.: *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, 1999, pp. 165-190.

[52] Zunkel, R. L., Hand Geometry Based Identification, in: Jain, A., Bolle, R., Pankanti, S. Eds.: *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, 1999, pp. 87-101.

- [53] Jain, A., Bolle, R., Pankanti, S., Introduction to biometrics, in: Jain, A., Bolle, R., Pankanti, S. Eds.: *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, 1999, pp. 1-42.
- [54] Hill, R. B., Retina Identification, in: Jain, A., Bolle, R., Pankanti, S. Eds.: *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, 1999, pp. 123-141.
- [55] Daugman, J. D, Downing, C. A., Demodulation: a new approach to texture vision In: Harris, L., Jenkin, M., Eds.: *Spatial Vision in Humans and Robots*, Cambridge University Press, 1993
- [56] Daugman, J., Downing, C., Epigenetic randomness, complexity and singularity of human iris patterns. *Proceedings of the Royal Society of London B*, **468**, pp. 1737-1740, 2001.
- [57] Nixon, M. S. and Carter, J. N., Automatic Gait Recognition, in: Jain, A., Bolle, R. and Pankanti, S. Eds.: *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, 1999, pp. 231-250.
- [58] Persaud, K. C., Lee. D-H., Byun, H-G., Objective Odour Measurements, in: Jain, A., Bolle, R., Pankanti, S., Eds.: *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, 1999, pp. 251-272.
- [59] Sommerville, B. A., McCormac, J. P., Broom, D. M., Analysis of human sweat volatiles : an example of pattern recognition in the analysis and interpretation of gas chromatograms, *Pestic. Sci.*, **41**, pp. 365-368, 1994
- [60] Nalwa, V. S., Automatic On-Line Signature Verification, in: Jain, A., Bolle, R. and Pankanti, S., Eds.: *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, 1999, pp. 143-164.

- [61] Halliday, D., and Resnick, R., *Physics Part I*, Wiley, New York, Third Edition, 1977
- [62] Halliday, D., and Resnick, R., *Physics Part II*, Wiley, New York, 1962
- [63] Grant, I. S. and Phillips, W. R., *Electromagnetism*, Wiley, New York, 1990.
- [64] Sadiku, M. N. O., *Elements of Electromagnetics*, Saunders College Publishing, Second Ed., 1989.
- [65] Open University, *M203: Introduction to Pure Mathematics*, Handbook, Unit 4, Section 2: Matrices of Linear Transformations, Open University Press, 1999
- [66] Fraleigh, J. B., Beauregard, R. A., *Linear Algebra*, 2nd Ed. Addison-Wesley
- [67] Strang, G., *Linear Algebra and its Applications*, 3rd Ed. Saunders HBJ, 1988
- [68] Van Oterloo, P. J., *A Contour-oriented Approach to Shape Analysis*, Prentice-Hall, 1991
- [69] Loncaric S., A Survey of Shape Analysis Techniques, *Pattern Recognition*, **31**(8), pp. 983-1001, 1998
- [70] Daugman, J., Biometric decision landscapes, *Technical Report TR482*, University of Cambridge Computer Laboratory, 1999
- [71] Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., *Numerical Recipes in C: The Art of Scientific Computation*, Cambridge University Press, 1997