Force Field Energy Functionals for Image Feature Extraction

David J. Hurley, Mark S. Nixon, John N. Carter
Department of Electronics and Computer Science
University of Southampton, Southampton SO17 1BJ, UK
[djh97r|msn|jnc]@ecs.soton.ac.uk

Abstract

Ears are an emergent biometric accruing application advantages including no requirement for subject contact and acquisition without demand. To recognize a subject's ear, we aim to extract a characteristic vector from a human ear image that may subsequently be used to identify or confirm the identity of the owner. Towards this end, a novel force field transformation and potential well extraction technique has been developed which leads to a compact characteristic vector offering immunity to initialization, rotation, scale, and noise. The image is transformed by considering the image to consist of an array of Gaussian attractors, which act as the source of a force field. The directional properties of the force field are exploited to automatically locate a small number of potential energy wells, which form the basis of the characteristic vector. We show how this is extracted for a selection of ears, and demonstrate its advantages. As such, we report a new technique in an exciting new biometric.

1 Introduction

In the context of 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; a burglar was recently convicted of murder in the UK on the basis of ear prints found at the scene of the crime [1]. An ear classification method has been developed for use in forensic science [2]. More recently, an automated system for ear identification has been developed [3].

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 authorized 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 feature vectors 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 vector would be produced every time the

BMVC 1999 doi:10.5244/C.13.60

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 by computer vision. Essentially, we need to find an ear and describe it for recognition. Clearly, there are application constraints, such as occlusion by hair, but here we are concerned with basic technique. Ear extraction could use an active contour [4] but with initialisation problems which can be relieved by a dual active contour [5], but this still requires establishment of inner and outer contours. Techniques derived from fingerprint analysis or texture classification could be used to describe the folds and ridges in a human ear. 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. The two stages 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. Secondly, the directional property of the ensuing force field is exploited to automatically locate a small number of potential wells, which correspond to local energy extrema in the scalar potential energy surface, which underlies the vector force field. It has been found that the potential well location process shows remarkable invariance to the initial choice of starting points and that the force field structure and hence the relative position of the wells is invariant to both scaling and rotation. Further, it appears quite robust in the presence of noise.

Section 2 deals with the synthesis of the force field and the analysis of the transformation. We show that it is in fact a linear transformation. Section 3 describes the potential well extraction process and demonstrates its invariant properties and noise immunity. Section 4 concludes with some observations about further uses of the new technique.

2 Image to Feature Transformation

This section deals with the synthesis of the force field and the analysis of the transformation. We show that it has basic properties: it is a linear transformation with a matrix representation. The concepts underpinning the force field transformation and the mathematics used to describe it can be found in various introductory works on physics [6,7] and electromagnetics [8,9]. Mathematical modeling techniques used in physics have recently attracted the attention of researchers in computer vision; for example [10] describes the use of vector potential to extract corners by treating the Canny edge map of the image as a current density. A recent approach [11] has used a potential field model in a medial axis transform.

2.1 Force Field Transformation

The image is transformed by considering the image to consist of an array of Gaussian attractors, which act as the source of a force field. Each pixel is assumed to generate a spherically symmetrical force field so that the force $\mathbf{F}_i(\mathbf{r})$ exerted on a pixel of unit intensity at the pixel location with position vector $\mathbf{r}$ by a remote pixel with position vector $\mathbf{r}_i$ and pixel intensity $P(\mathbf{r}_i)$ is given by

$$\mathbf{F}_i(\mathbf{r}) = P(\mathbf{r}_i)\frac{\mathbf{r}_i - \mathbf{r}}{\left|\mathbf{r}_i - \mathbf{r}\right|^3} \tag{1}$$

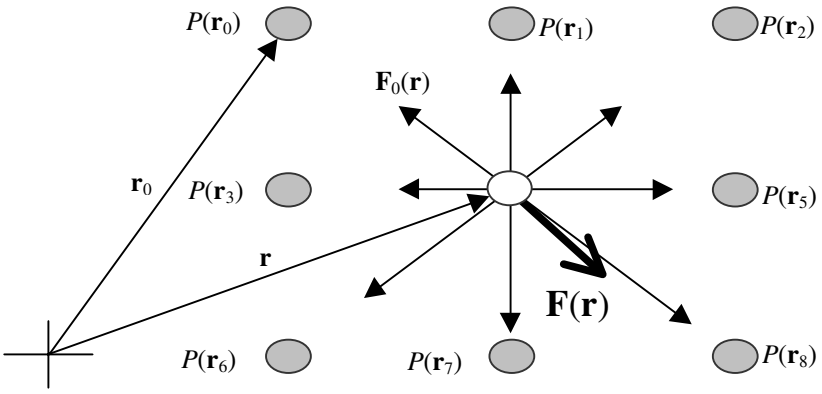


Figure 1 Force field geometry

This calculation is illustrated graphically for the total force acting at a typical pixel position in Figure 1. The units of pixel intensity, force, and distance are arbitrary, as are the coordinates of the origin of the vector field. The total force acting on a unit pixel at a given position is the vector sum of all the forces due to the other pixels in the image and is given by,

$$\mathbf{F}(\mathbf{r}) = \sum_i \mathbf{F}_i(\mathbf{r}) = \sum_i P(\mathbf{r}_i)\frac{\mathbf{r}_i - \mathbf{r}}{\left|\mathbf{r}_i - \mathbf{r}\right|^3} \tag{2}$$

In order to calculate the force field for the entire image, this equation should be applied at every pixel position in the image.

2.2 Potential Energy Field

Associated with the force field generated by each pixel there is a spherically symmetrical scalar potential energy field, where $E_i(\mathbf{r})$ is the potential energy imparted to a pixel of unit intensity at the pixel location with position vector $\mathbf{r}$ 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

$$E_i(\mathbf{r}) = \frac{P(\mathbf{r}_i)}{\left|\mathbf{r}_i - \mathbf{r}\right|} \tag{3}$$

The potential energy function of a single isolated pixel appears as an inverted vortex as shown in Figure 2. 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

$$E(\mathbf{r}) = \sum_i E_i(\mathbf{r}) = \sum_i \frac{P(\mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{r}|} \tag{4}$$

This summation is then carried out at each pixel location to generate a potential energy surface, which is a smoothly varying surface due to the fact that the underlying inverted vortices have smooth surfaces.

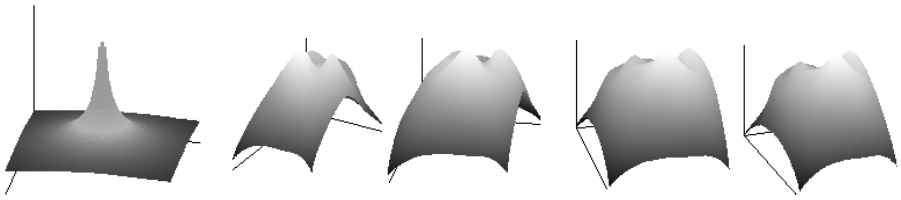


Figure 2 Potential energy surface for an ear (four perspectives on the right) formed by summing many thousands of potential energy functions of individual pixels (left).

The vector force field and scalar potential energy 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,

$$\mathbf{F}(\mathbf{r}) = -grad\big(E(\mathbf{r})\big) = -\nabla E(\mathbf{r}) \tag{5}$$

2.3 Field Lines, Channels, and Wells

We introduce the concept of a unit value exploratory text pixel to assist in describing field lines. When such a test pixel is placed in a force field and allowed to follow the field direction, its trajectory is called a *field line*. If this process is carried out at many different starting points a set of field lines will be generated that capture the general flow of the vector field.

An important property of field lines is that they never cross over for the simple reason that the field vector at a point is unique. So if two trajectories should happen to arrive at the same pixel location they will follow the same path from that point onwards. If other trajectories join this path then they will also follow it, thus forming channels. We refer to these as *potential energy channels*.

The potential energy surface may undulate in such a manner that it forms local extrema called *potential energy wells*. The potential surface of the ear shown in Figure 2 clearly shows two such wells. Notice that the most prominent of these also shows a potential channel leading into it from the left in the form of a gently sloping ridge. Field lines which flow into these wells become trapped because they follow the gradient towards the extremum where the force is zero.

2.4 Field Line Generation

No attempt is made here to prescribe exactly how field lines should be generated. In the examples that follow, typically 50 test pixels are initialized at equally spaced intervals in the form of an ellipse. The coordinates of each test pixel are maintained as real numbers

rather than as integers. This means that each point moves in the direction indicated by the force field sample rather than just in one of eight directions, which results in a much smoother trajectory. No attempt has been made to interpolate between samples, although this should be possible if more accuracy is required. Each test pixel's position was updated in increments of one pixel width.

2.5 An Invertible Linear Transformation

We now show that the force field transformation is a linear transformation. It is sufficient to show that there the force field transformation has a corresponding matrix representation, since linear transformations between finite-dimensional vector spaces are precisely those transformations that have matrix representations. The form of the matrix representation is illustrated for a trivial 2 x 2 pixel image. It is easily verified that this represents the application of Equation 2 at each of the four pixel locations. This equation multiplies the vector $\mathbf{p}$ of pixel intensities (P_i) by the representation matrix $\mathbf{A}$ (comprised of inverse square displacement vectors, $\mathbf{d}_{ij}$) to give the force vector $\mathbf{F}$. We have,

$$\mathbf{Ap} = \mathbf{F}$$

$$\begin{matrix} 0 & \mathbf{d}_{01} & \mathbf{d}_{02} & \mathbf{d}_{03} & P_0 \\ \mathbf{d}_{10} & 0 & \mathbf{d}_{12} & \mathbf{d}_{13} & P_1 \\ \mathbf{d}_{20} & \mathbf{d}_{21} & 0 & \mathbf{d}_{23} & P_2 \\ \mathbf{d}_{30} & \mathbf{d}_{31} & \mathbf{d}_{32} & 0 & P_3 \end{matrix} = \begin{matrix} \mathbf{F}_0 \\ \mathbf{F}_1 \\ \mathbf{F}_2 \\ \mathbf{F}_3 \end{matrix} \quad \text{where} \quad \mathbf{d}_{ij} = \frac{\mathbf{r}_j - \mathbf{r}_i}{\left|\mathbf{r}_j - \mathbf{r}_i\right|^3} \tag{6}$$

This is a skew-symmetric matrix of the form, $\mathbf{A}^T = -\mathbf{A}$. The leading diagonal of zeros reflects the fact that no pixel attracts itself and the skew symmetry is accounted for by the fact that we are dealing with a fully connected network but with a pair of directed edges connecting every pair of nodes.

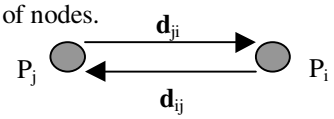


Figure 3 Skew symmetry

There is a corresponding representation for the potential energy transformation. Since the representation matrix is square we naturally ask whether it is invertible. We find that the force field matrix is singular if the number of image pixels is odd but that it is invertible if the number is even. The potential energy matrix is invertible in either case. This result is important because it means that the original image is in principle recoverable from the potential energy surface, and therefore all the information contained in the image is preserved in the transformation.

3 Results for Human Ears

In this section we show how the foregoing theory can be applied to the problem of ear biometrics. We begin by taking our first look at a force field that has been generated

from a 160 x 100 pixel ear image. It is not possible to see a force field directly because it consists of vectors so we convert it to a scalar field by taking the magnitude of each vector. Figure 4 shows the result, where we see that the ear is still clearly recognizable. We notice that the transformation appears to provide a remarkable degree of intrinsic smoothing and also that there appears to be something akin to edge detection. The former can be explained by viewing Equation 2 as a giant smoothing kernel with an inverse square profile and whose domain is the entire image. The latter we attribute to a process which we call homogeneous cancellation; 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. These and other effects are under investigation and will be reported in due course.

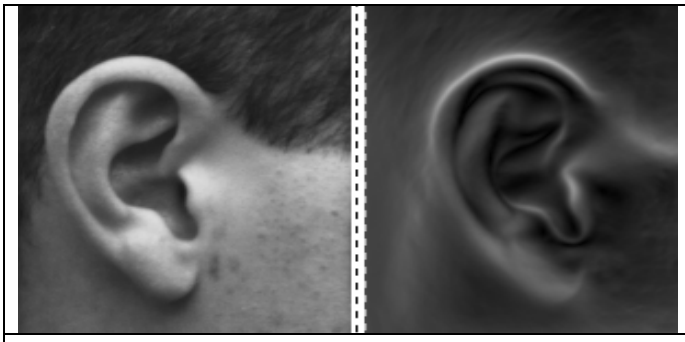


Figure 4 Magnitude of force field

Having looked at the global description offered by the potential energy surface and the finer detail contained in the force field magnitude plot we will now see how field lines can be used to extract information that lies outside the sensitivity span of the human eye.

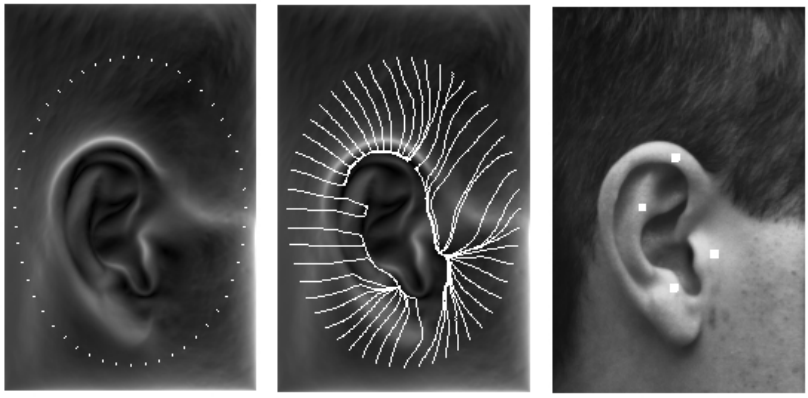
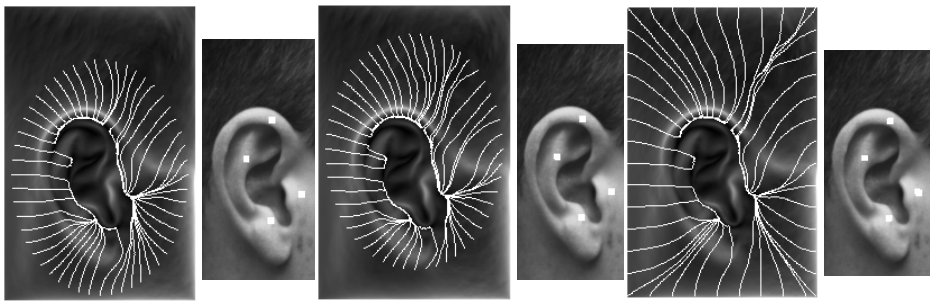


Figure 5 Potential well extraction

Figure 5 shows how 50 test pixels have been arranged in an ellipse shaped array and then iterated to generate a set of field lines. Even though the potential surface indicates the presence of two potential wells and one potential channel, we see that field lines are much more sensitive and have extracted four potential wells whose positions are automatically indicated on the rightmost image. We see how field lines flow into potential channels and continue onwards until they terminate in potential wells. For example, notice how fourteen field lines cross the upper rim of the ear, each joining a common channel that follows the curvature of the rim rightwards and finally terminates in a potential well.

Having demonstrated the remarkable ability of field lines to extract potential wells and depict channels, we now need to check that the outcome remains the same when we alter the initial conditions such as the initial arrangement of the array of test pixels. Perhaps more importantly we need to confirm that an image of the same ear at a smaller pixel resolution produces the same force field. Does noise completely alter the result? What about different illumination conditions? Does rotating the image have any effect? Are the results for different ears sufficiently different to act as a discriminant? We report the results of addressing some of these issues here.

Figure 6 demonstrates **initialization invariance**. The location of potential wells is the same in the left and center images with two quite different initializations. This is hardly surprising since the force field is not altered merely because we choose to enter it at different locations. Clearly the density of field lines needs to be sufficiently high to ensure adequate coverage of the image. The rightmost image shows an initialization along the edges of the image at intervals of 20 pixels. We see that it is a matter of chance whether the ellipse starting points happen to coincide with one of these field lines. Again, the wells are in the same position.



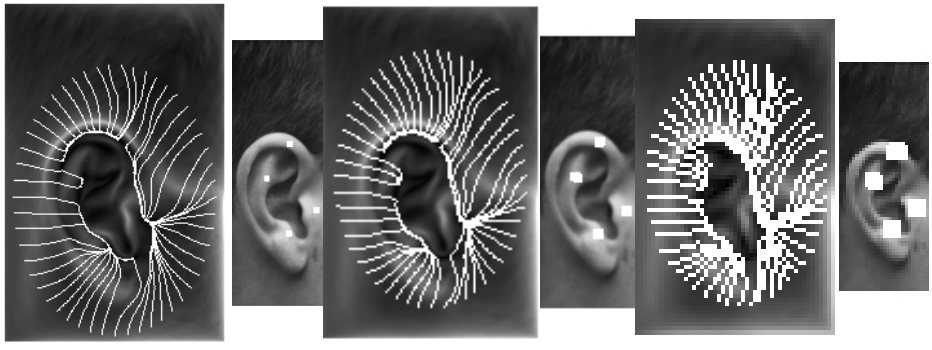
anwd = 1.16, *wdir* = 4.19	*anwd* = 1.16, *wdir* = 4.20	*anwd* = 1.15, *wdir* = 4.20

Figure 6 Initialization Invariance

To assess difference between the results, we shall use a measure of the average normalized distance of the well positions *anwd*, together with the accumulated direction to the position of each well-point, *wdir*. For *W* wells at points *w*, these measures are:

$$anwd = \frac{\sum_{i=1}^{W} |w_i|}{\max(|w_i|) \cdot W} \quad \text{and} \quad wdir = \sum_{i=1}^{W} \left(\langle w_i \right) \tag{7}$$

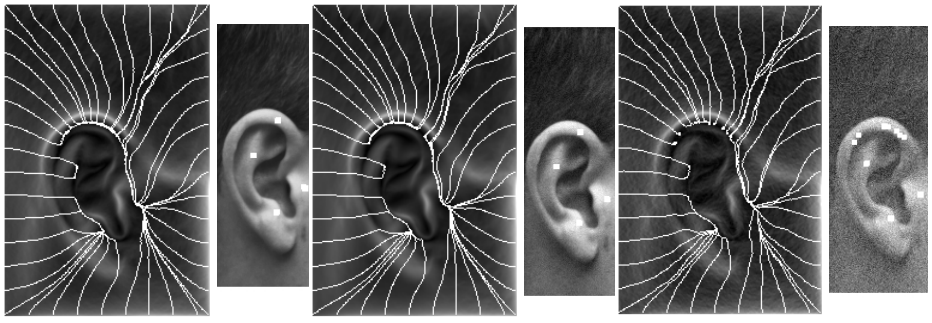
The measures are shown in Figure 6 for each different starting point, and show that very similar measures are achieved, reflecting visual analysis of these results. Figure 7 demonstrates **scale invariance.** We see that the structure of the force field is essentially preserved when an image is at lower resolution. This is an important result because it means 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. The earlier measures are again tabulated in Figure 7 and again show invariance to scale, and are very close to the measures in Figure 6.



170 x 261 pixels	100 x 154 pixels	77 x 50 pixels
anwd = 1.16, *wdir* = 4.19	*anwd* = 1.15, *wdir* = 4.20	*anwd* = 1.17, *wdir* = 4.21

Figure 7 Scale Invariance

Noise tolerance is demonstrated in Figure 8 where it is seen that the force field structure is essentially preserved in the presence of Gaussian noise. Notice that the channels begin to disintegrate into individual wells in the presence of severe noise. The channel outline is still clearly present so optimal estimation techniques could be employed in the presence of severe noise.



No Noise	Medium Noise	Severe Noise

Figure 8 Noise tolerance

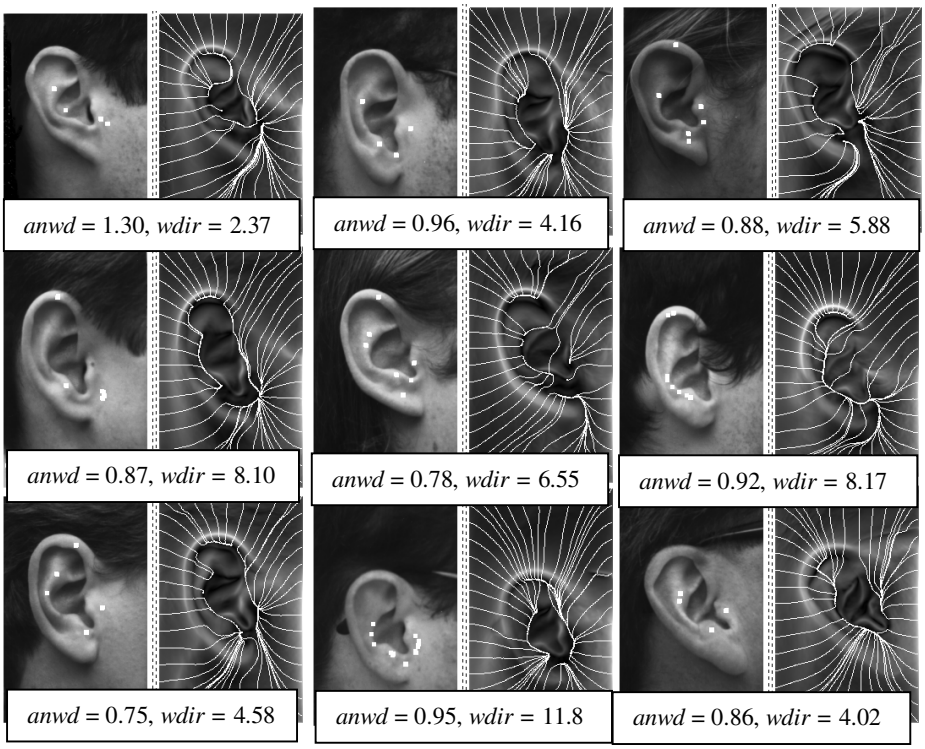


Figure 9 Different ears produce different features

Figure 9 demonstrates that a variety of different ears produce quite different feature vectors and that potential channels and well locations are unique to each image. Further, the measures are quite different to those for Figures 6 to 8 (which are those for a different ear), which show that the within-class variation appears less than the between-class variation. Clearly, a richer selection of measures will emphasize this effect. Future work, on a large database, will aim to confirm the potential for this technique in ear recognition.

4 Conclusions

We have developed a new feature extraction technique, targeted primarily at ear biometrics, with remarkable invariance to initialization, scale, and rotation and which demonstrates good noise tolerance. The beauty of this technique is that an explicit description of the ear topology is not necessary and extracting the ear biometric is simplicity itself – merely follow the force field lines and observe eventual clustering of coordinates. It is anticipated that in order to achieve greater discrimination with larger ear populations that more information will need to be extracted. This information is readily available in the topology of the potential channels.

Whilst the force field transformation has been demonstrated in the context of ear biometrics, it is an important new development in its own right. Preliminary investigations suggest that the transformation may be used in face recognition. In fact it

may well be that it provides a general technique of converting complex natural images into compact signatures in a robust and reliable manner. A very important aspect of the transformation is the fact that it simulates a natural process, namely the formation of electric fields in the vicinity of electric charge distributions. For example the image formed on a charge-coupled device will result in a charge distribution which will have an associated electric force field. This holds out the prospect of a solid state device with direct image to force field conversion in real time. Even more interestingly, it may well be that the image formed on the human retina has a charge distribution and an associated electric force field. It may well turn out that the nervous system in the human eye is sensitive to such force fields and exploits them to convert complex images into compact signatures that assist with pattern recognition.

References

[1] BBC News, BBC Online Network, 15 Dec.1998. http://news.bbc.co.uk/hi/english/uk/newsid_235721.stm

[2] A. Iannarelli, *Ear Identification*, Paramount 235000/ Publishing Company, Freemont, California, 1989

[3] M. Burge, and W. Burger, Ear Recognition, *In:* A. K. Jain, R. Bolle and S. Pankanti Eds., *Biometrics: Personal Identification in Networked Society*, pp. 273-286, Kluwer Academic Publishing, 1998

[4] M. Kass, A. Witkin, D. Terzopoulos, *Snakes: Active Contour Models*, International Journal of Computer Vision, *1*:321-331, 1988

[5] S. R. Gunn and M. S. Nixon, A Robust Snake Implementation; A Dual Active Contour, *IEEE Transactions PAMI*, **19**(1), pp 63-67, 1997

[6] D. Halliday, and R. Resnick, *Physics Part I*, John Wiley & Sons, Third Edition 1977

[7] D. Halliday, and R. Resnick, *Physics Part II*, Wiley International Edition, 1962

[8] I. S. Grant, and W. R. Phillips,. *Electromagnetism*, John Wiley & Sons, Second Ed., 1990.

[9] M. N. O. Sadiku, *Elements of Electromagnetics*, Saunders College Publishing, Second Ed., 1989.

[10] B. Luo, A. D. Cross, E. R. Hancock, Corner Detection via Topographic Analysis of Vector Potential, *Proceedings of the 9th British Machine Vision Conference,* 1998

[11] N. Ahuja, J. H. Chuang, Shape Representation Using a Generalized Potential Field Model, *IEEE Transactions PAMI*, **19**(2), pp 169-176, 1997

[12] The Open University, *M203 Introduction to Pure Mathematics, Linear Algebra Block: Unit 4 Linear Transformations*, Open University Press, 1999