A New Force Field Transform for Ear and Face Recognition

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

Abstract

The objective in defining feature space is to reduce the dimension of the original pattern space yet maintaining discriminatory power for classification [1]. To meet this objective in the context of ear and face biometrics a novel force field transformation has been developed in which the image is treated as an array of Gaussian attractors that 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 and channels that form the basis of a characteristic feature vector. Here, we generalise the analysis, and the stock of applications.

1 Introduction

In vision-based biometrics we aim to extract a compact description from an image which may subsequently be used to identify or confirm the identity of the owner [2]. Now in order to meet this objective we have developed an invertible linear transformation which we call the *force field transform* [3]. Our primary concern has been automatic ear recognition [3]; we now extend our analysis and show that the transformation appears to have similar potential for automatic face recognition.

The entire ear image is converted into a force field by pretending that each pixel exerts an isotropic force on all the other pixels that is proportional to pixel intensity and inversely proportional to the square of the distance. There is a potential energy surface associated with this force field, which in the case of an ear can be likened to a small mountain with a few peaks joined by ridges. We call these peaks potential energy wells and the ridges joining them potential energy channels. The directional property of the force field is exploited to automatically locate these potential wells and channels, which then form the basis of the ear's signature.

An array of unit value exploratory mobile test pixels is arranged in a closed loop formation surrounding the target ear. Each test pixel is then allowed to follow the pull of the force field so that its trajectory forms a field line and it will continue moving until it reaches the center of a well where no force is exerted and no further movement is possible. Since the force field at a point is unique all field lines which arrive at a given point will follow the same path from that point onwards thus forming channels. This process is illustrated in Figure 3 where an elliptic array of 50 test pixels is placed in the force field and iterated to produce the field lines shown in the center. The most striking example of the channel formation process is seen at the top of the ear where 14 field lines combine to form a channel which flows rightwards following the contour of the ear-rim to finally terminate in a well. The locations of the wells are extracted by simply noting the coordinates of the clusters of test pixels that eventually form. These locations are shown on the right, superimposed on force field magnitude.

The structure of the force field as described by the field lines shows remarkable initialization invariance [3] in the sense that if the radius of the ellipse is altered or if its center is translated, the same channel and well description will result. Also if the image is scaled the force field structure scales with the image. We have also found that the process is very tolerant of noise, due to its inherent averaging. In this paper we generalise our earlier presentation to show more basic properties of the technique and how it can be applied to other pattern recognition problems.

2 Force Field Approach

The image is transformed by pretending that it consists of an array of N Gaussian attractors, which act as the source of a force field. Each pixel is considered 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 any other pixel with position vector $\mathbf{r}_i$ and pixel intensity $P(\mathbf{r}_i)$ is given by

$$\mathbf{F}_i(\mathbf{r}_j) = P(\mathbf{r}_i)\frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3} \quad (1)$$

The units of pixel intensity, force, and distance are arbitrary, as are the co-ordinates of the origin of the vector field. 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,

$$F(\mathbf{r}_j) = \sum_{i=0,\neq j}^{N-1} \mathbf{F}_i(\mathbf{r}_j) = \sum_{i=0,\neq j}^{N-1} \left(P(\mathbf{r}_i) \frac{\mathbf{r}_i - \mathbf{r}_j}{\left| \mathbf{r}_i - \mathbf{r}_j \right|^3} \right) \quad (2)$$

In order to calculate the force field for the entire image, this equation should be applied at every pixel position in the image. Since this procedure is quadratic in the number of pixels N, a more efficient approach is to exploit the speed of the *Fast Fourier Transform* by viewing the process as a convolution of the image with the field associated with a unit value test pixel. A nine-fold memory penalty is incurred since zero padding is required for anti-aliasing. However the reward is a computation complexity of order $9N\log(9N)$.

Associated with the force field generated by each pixel there is a spherically symmetrical scalar potential energy field 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 any other pixel with position vector $\mathbf{r}_i$ and pixel intensity $P(\mathbf{r}_i)$, and is given by

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

The defining equation is simpler than the force field equation but the concept is less intuitive. If an exploratory unit test pixel is moved around in the force 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 potential energy between the two rings. The potential energy function of a single isolated pixel looks like an inverted vortex as shown in Figure 1
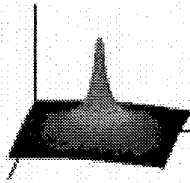

Figure 1 Isolated pixel potential energy function

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

$$E(\mathbf{r}_j) = \sum_{i=0,\neq j}^{N-1} E_i(\mathbf{r}_j) = \sum_{i=0,\neq j}^{N-1} \frac{P(\mathbf{r}_i)}{\left| \mathbf{r}_i - \mathbf{r}_j \right|} \quad (4)$$

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.
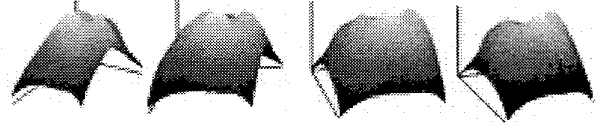
Figure 2 Potential energy surface for an ear

3 Invertible Linear Transform

In the appendix we show that the force field transform is a linear transformation and this is confirmed here by giving its matrix representation [4]. The form of the matrix 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 a column vector of pixel intensities (P_i) by a matrix of inverse square displacement vectors $\mathbf{d}_{ij}$ to give a column vector of forces $(\mathbf{F}_i)$. We have,

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

This is a skew-symmetric matrix: 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.

There is a corresponding representation for the potential energy transformation since 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,

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

Since the representation matrices are square it is of theoretical interest whether they are invertible or not. If they are invertible then the original image can be recovered for example from the potential energy surface. This implies that all the information in the original image is conserved by the transformation, which is an important result. In practice the representation matrices for images of even modest size are very large, for example a 10x10 image has a matrix with 10,000 elements. However we

have tested the potential energy representation matrices for all square images up to 32x32 pixels and all non-square images up to 7x8 pixels and have found them to be invertible. These results suggest that the potential energy transform is indeed invertible for most image sizes and aspect ratios.
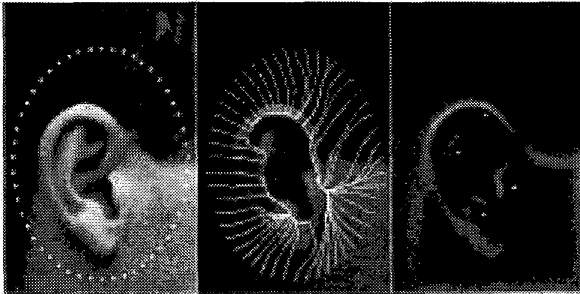

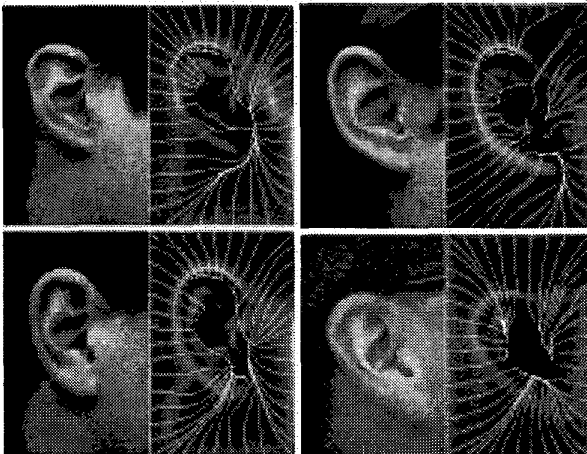**Figure 3 Force field feature extraction for an ear**

Figure 4 Different descriptions for different ears

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 all of the image information is conserved by the transformation.

In Figure 4 we show results of the same extraction process for different ears. Clearly the new transformation leads not only to a different channel description for each ear, but also to different wells

4 Force Field Faces

Here we demonstrate that the technique may also be employed in face recognition. We find generally that there are fewer potential wells in the results for faces when compared with those for ears, so that greater emphasis and reliance must be placed on potential
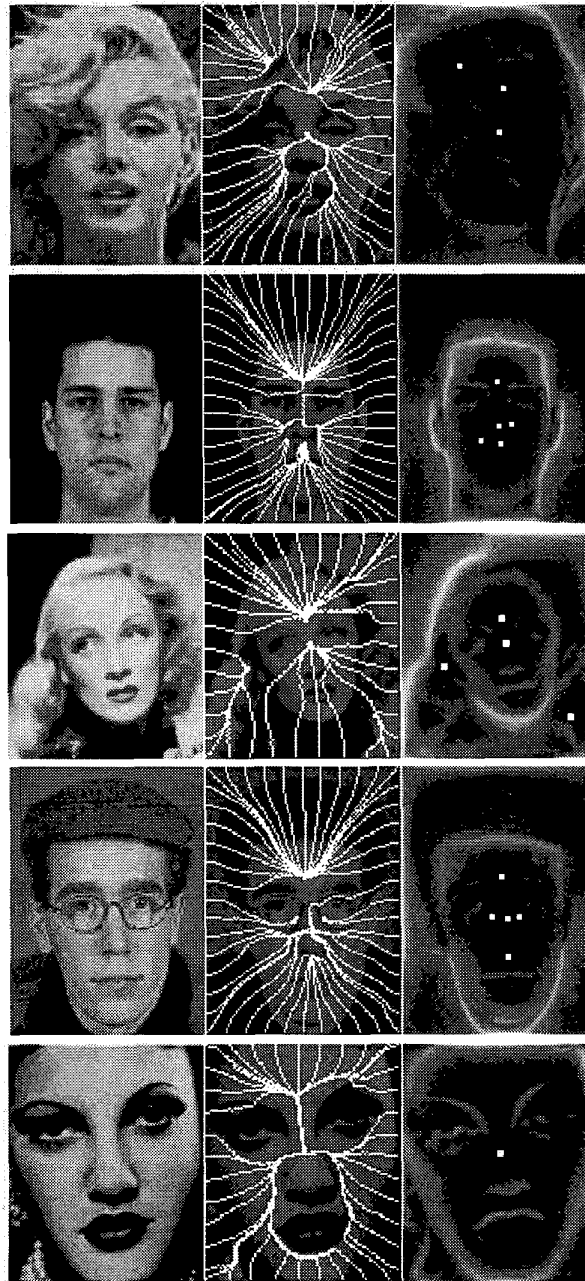

Figure 5 Force field feature extraction applied to faces

channel descriptions. As shown in Figure 5 the potential channels are quite unique for each of the four faces. Notice that only one potential well is located for the last face in the set, roughly in the middle of the nose. We see however that a rich set of distinctive channels leads into this single well that can be used to provide a reliable

27

description. Notice how two channels grow downwards to surround the mouth and a single channel runs up along the length of the nose which then splits into two further channels running parallel to the eyebrows. In all cases there are also channels forming along the cheekbones.

5 Conclusions

We have developed a new feature extraction technique, targeted primarily at ear biometrics but which readily extends to include face recognition. The technique is robust and reliable with remarkable invariance to initialization and possessing excellent noise tolerance. The beauty of this technique is that an explicit description of the target topology is not necessary. Extracting the potential well description merely involves following the force field lines and observing eventual clustering of coordinates. Taking account of the channel shape and ultimately the underlying shape of the energy surface can increase the level of detail in the description to meet any demand.

We are currently investigating how linearity can be exploited to perform force field tracking in dynamic images. A dynamic force field can be updated by simply adding the force field associated with the difference between sequential images, often a sparse array and therefore easily computed. We are also investigating how information conservation, which is a consequence of the transform invertiblilty, can be exploited to perform image compression.

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.

.

References

[1] H. C. Andrews, *Introduction to Mathematical Techniques in Pattern Recognition*, John Wiley & Sons Inc., p 9, 1972

[2] 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

[3] D.J. Hurley, M.S. Nixon, J.N. Carter, "Force Field Energy Functionals for Image Feature Extraction", *Proc. British Machine Vision Conf.* BMVA Press, **II**, pp. 604-613, 1999

[4] M203 Introduction to Pure Mathematics, Linear Algebra Block: Unit 4 Linear Transformations, Section 2: Matrices of linear transformations, M203 Handbook, p29, Open University Press, 1999

Appendix

Proof of Linearity

Let A and B be elements of the vector space V of MxN matrices whose elements $A_{i,j}$ and $B_{i,j}$ are real numbers representing pixel intensities. Also let W be the vector space of MxN matrices where the elements are forces as defined by the relation:

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

Now, to check that the defining relation for a linear transform is satisfied, we need to show that

$$t(\alpha_1\mathbf{A} + \alpha_2\mathbf{B}) = \alpha_1 t(\mathbf{A}) + \alpha_2 t(\mathbf{B}),$$
$$\text{for all } \alpha_1, \alpha_1 \in \Re, \mathbf{A}, \mathbf{B} \in V \tag{6}$$

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

$$[t(\alpha_1\mathbf{A})]_{i,j} = \left(\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \alpha_1 A_{m,n} \frac{\mathbf{r}_{m,n} - \mathbf{r}_{i,j}}{\left|\mathbf{r}_{m,n} - \mathbf{r}_{i,j}\right|^3} \right) \tag{7}$$

$$[t(\alpha_2\mathbf{B})]_{i,j} = \left(\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \alpha_2 B_{m,n} \frac{\mathbf{r}_{m,n} - \mathbf{r}_{i,j}}{\left|\mathbf{r}_{m,n} - \mathbf{r}_{i,j}\right|^3} \right) \tag{8}$$

$$[t(\alpha_1\mathbf{A}) + t(\alpha_2\mathbf{B})]_{i,j} =$$

$$\left(\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} (\alpha_1 A_{m,n} + \alpha_2 B_{m,n}) \frac{\mathbf{r}_{m,n} - \mathbf{r}_{i,j}}{\left|\mathbf{r}_{m,n} - \mathbf{r}_{i,j}\right|^3} \right)$$

$$= [t(\alpha_1\mathbf{A} + \alpha_2\mathbf{B})]_{i,j} \tag{9}$$

Hence, since we have shown that the defining relation is satisfied, the force field transform is indeed a linear transformation.