THE ALTERNATIVE SNAKE - AND OTHER ANIMALS

G. L. Scott

Robotics Research Group
Department of Engineering Science
University of Oxford
Parks Road
Oxford OX1 3PJ

ABSTRACT

In this paper I introduce a new way of generating and describing shapes which was largely inspired by reading D'Arcy Thompson's classic "On Growth and Form" [1]. There are suggestive parallels between my system and "coupled oscillation" models of handwriting [2] and locomotion [3].

Like many others the system involves a mapping from a space of "implicit parameters" to the space of observables. I limit discussion to the simplest case in which a single parameter θ is mapped into 2-D Euclidean space by two Fourier series $x(\theta), y(\theta)$ whose coefficients may be restricted or cross-constrained in various ways.

One application to computer vision is in defining "snakes" - contour models which are fitted to image data by gradient-climbing techniques [4,5].

I

There are good prima facie reasons for commencing an enquiry into visual perception by way of the problem of encoding biological forms rather than that of encoding ideal (e.g. Euclidean) forms. The most compelling of these reasons is that the visual systems of many animals, including man, have evolved to quickly and accurately apprehend biological forms (and motions) and not ideal forms (or ideal motions). Although in recent times we have come to be surrounded predominantly by artefacts (i.e. things which have been assembled rather than grown) there is much evidence for a strong "animist" strain in their design. Modern motor cars, for example, are not generally styled to *look* as if they are assembled. Their makers strive to endow them with sleek and "spirited" looks and never tire of reminding us that they are alive ("Minis have feelings too"). So insights into biological shape perception might be of importance in understanding the visual system of even the most upwardly mobile of earthpersons.

Many of the systems used in computer vision for describing shapes may be likened to a process of manufacture. Objects are described as composed of more primitive parts. In contrast the system I will discuss is analogous to the process of biological growth wherein an organism arises by a process of successive differentiation working upon an original unshaped "egg".

D'Arcy Thompson drew attention to a number of geometrical properties commonly found in biological forms. These include: an underlying tendency to periodicity in the shape of many organisms; and the ease with which one body-part may be made to resemble another in the same or a different species by "rubber sheet" deformation. Figure 1 shows the outline of the three toes of a tapir within coordinate frameworks constructed by Thompson to show how they may be intertransformed. In figure 2 a more complex set of transformations is shown - this time involving the carapaces of various crabs. The repeated occurrence of cusps around the outlines - with the teasing suggestion of some sort of underlying "system" - provides a particularly clear example of the "periodicity" mentioned above.

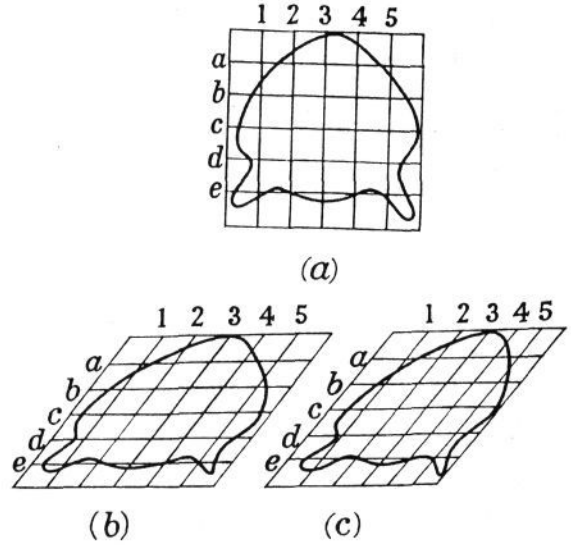

Figure 1: Relations between toes of the tapir (Thompson)

The dynamic processes occurring within the developing organism are very poorly understood (as are, consequently, the ways in which genes may modulate these processes). Alan Turing proposed in 1952 [6] that wave processes arising from the interaction between the concentrations of two chemicals ("Turing waves") could be made to account for many periodic features of organisms. There may be some slight correspondence, therefore, between my method of defining shapes and actual morphogenetic processes.

II

The equations

$$x = x_0 + a\sin\theta \quad (0 < \theta < 2\pi)$$

$$y = y_0 + b\cos\theta$$

AVC 1987 doi:10.5244/C.1.46

define a circle where $a=b$ and an ellipse aligned with horizontal and vertical otherwise. Equations of the more general form:

$$x=x_0+\sum a_n \sin(n\theta)$$

$$y=y_0+\sum b_n \cos(n\theta)$$

where summation is over n define an extremely wide variety of closed shapes which are symmetrical around the y-axis. If phases are shifted viz

$$x=x_0+\sum a_n \sin(n\theta+\phi_n)$$

$$y=y_0+\sum b_b \cos(n\theta+\psi_n)$$

then bilateral symmetry is in general lost. However, systematic shifts can be defined which have the effect of rotating the axis of symmetry. Conditions for other types of symmetry (tri-radial, S-symmetry etc) can also be defined.

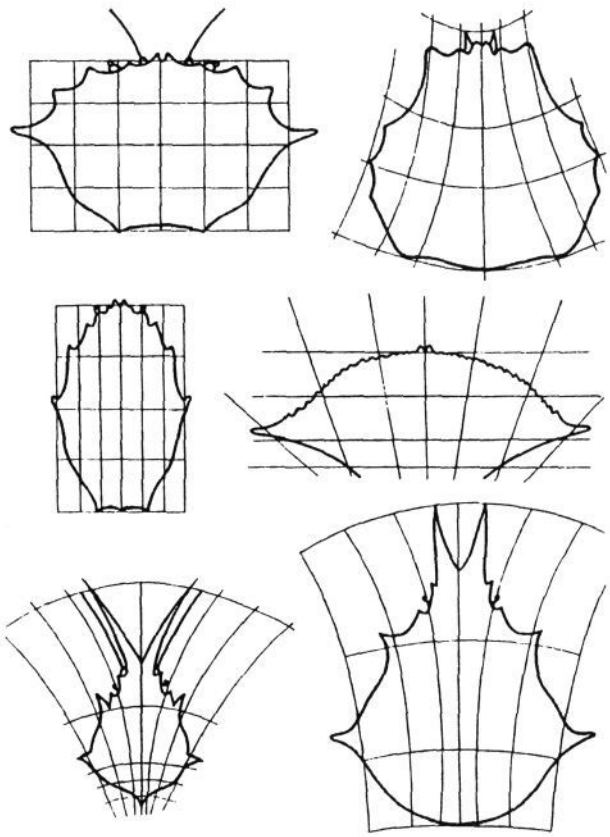

Figure 2: Relations between carapaces of crabs (Thompson)

An open curve can be obtained in a number of ways. One of these is to phase-shift in such a way that the curve backtracks upon itself after half a cycle (e.g. both $x(\theta)$ and $y(\theta)$ are pure sine series). Another method is to append a "sweep" term to one of the functions e.g.

$$x=x_0+v\theta+\sum a_n \sin(n\theta+\phi_n)$$

This produces "handwriting" in a manner similar to that proposed by Hollerbach [1] to account for human performance.

Figure 3 shows a variety of curves of order 6 (involving harmonics up to the sixth). The first three rows are bilaterally symmetric with all harmonics "in phase" as discussed above. The fourth row involves arbitrary relative phase between harmonics. The fifth row consists of "open" curves generated by shifting the phase of all the harmonics in one function by 90 degrees.

Figure 4 shows a "morphogenetic" sequence starting with a circle. Functions of order 4 - with all harmonics initially in phase - are involved. Amplitudes and phase-shifts are then changed by small amounts *randomly* between each "generation". Broadly speaking it is the increasing amplitude of the harmonics which is responsible for the appearance of structure and the increasing phase-differences which are responsible for the distortion of the shape away from perfect symmetry.

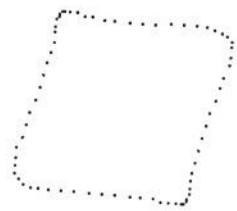

Figure 5: Showing high density of parameter at corners

Although both Fourier series are, being finite, perforce smooth the curve mapped out in xy space may display cusps - points of infinite curvature. These represent values of θ at which both $dx/d\theta$ and $dy/d\theta$ are zero. If we think of θ as time then the point tracing out the curve comes to a standstill for an infinitesimal duration and then accelerates off in a direction which will not in general be that in which it was last headed. Figure 5 shows the position (x,y) at 5° intervals in θ for a figure with cusps and rounded corners. A very marked increase in the density of θ is discernible in the vicinity of the these. Theoretically the density of θ is infinite where the curvature is infinite.

III

In principle any finite continuous open or closed curve can be defined to within some finite accuracy *in a multiple infinitude of ways* by two Fourier series $x(\theta)$ and $y(\theta)$. The reason for the infinitude is that θ is an implicit (hidden) parameter and not some measurable quantity such as distance along the curve or angle subtended at some origin. We may therefore distribute it along the curve in different ways. If we think of θ as time then we may traverse the curve according to any number of different schedules which involve the same total period of travel. Some ways of distributing θ, however, may yield descriptions which are brief and mostly in terms of the low harmonics whereas "arbitrary" distributions will yield descriptions involving high order terms. A simple example is the ellipse in which there is only one way of distributing θ such as to yield a precise description of the curve in terms of the fundamental frequency alone. Any other distribution (for example the "constant speed" parameterisation in which the density of θ is uniform) will involve high-order Fourier expansions.

We are interested in "natural" parameterisation - in those ways of distributing θ that yield simple

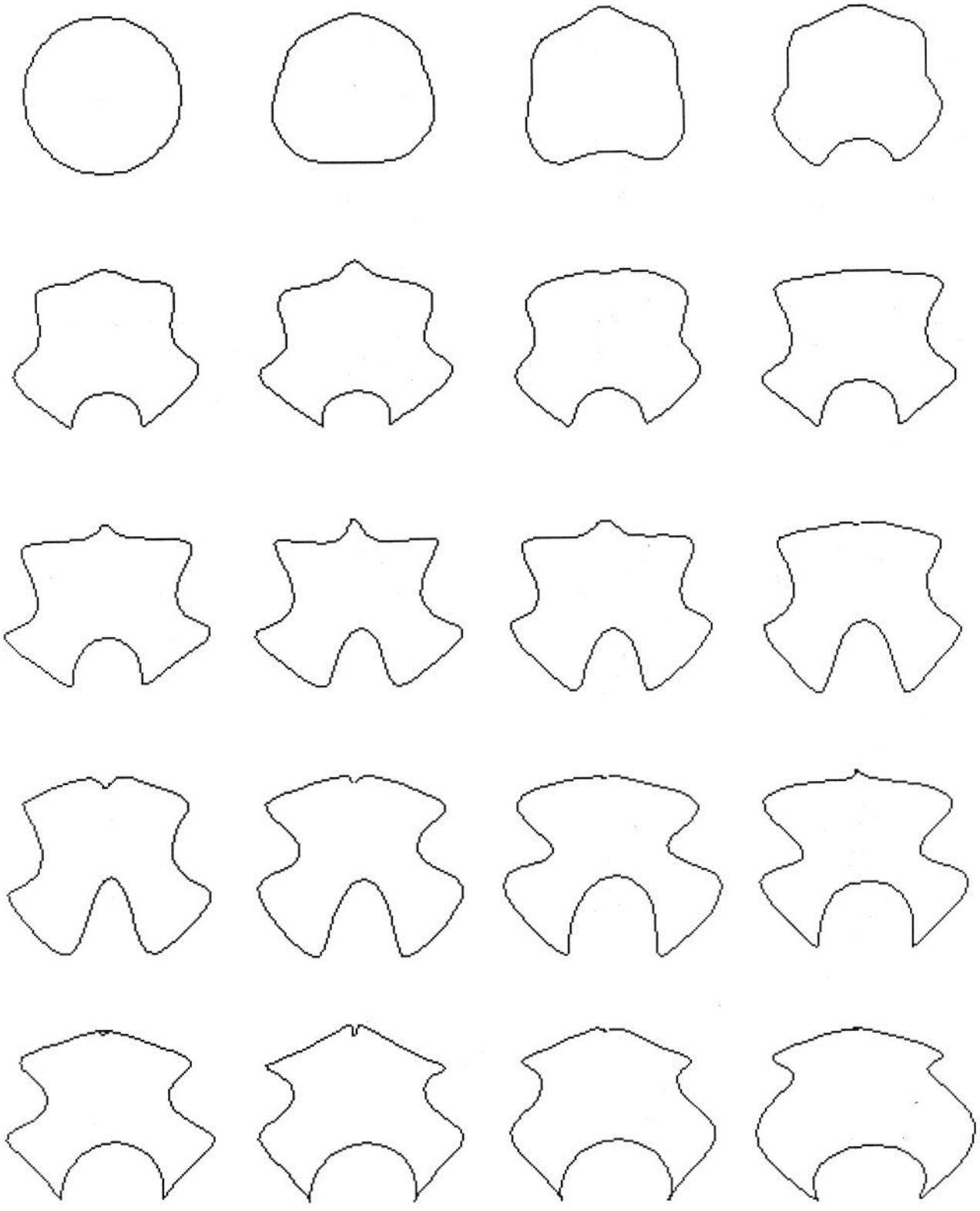

Figure 4: Random "morphogenetic" sequence

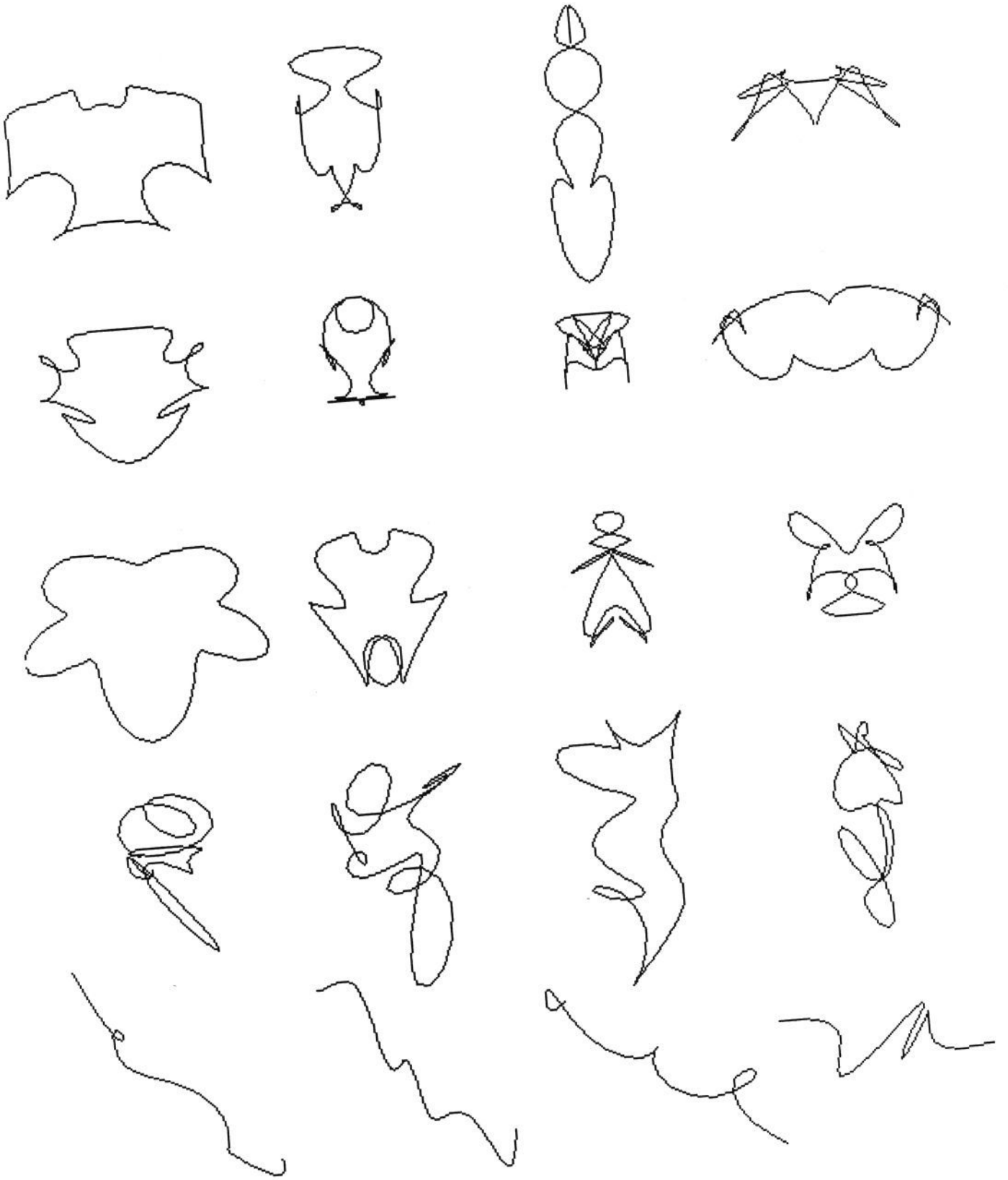

Figure 5: Unselected variety of figures generated at random

descriptions of the curve. There are several reasons for this interest of which one of the more subtle is the following. Where the *minimal* encoding of a curve is of order N the minimal encoding of its orthographic projection is also of order N (barring a violation of "General Viewpoint" - in which case the order will be less than N). The optimal parameterisation is projectively invariant viz. if two points on the original curve have a separation in θ-space of T then their projections will have the same separation T. (We should talk of separation rather than absolute value because θ-space may be arbitrarily rotated or reflected without increasing the order of the description). The circle, projecting to an ellipse, provides the simplest example. The natural way of parameterising a circle (uniformly around the circumference) "projects to" the natural way of parameterising an ellipse (which is non-uniformly around the perimeter).

Isaac Weiss [7] has explored a variational approach to the distribution of a non-observable "natural parameter" along a curve. His primary concern was to show that if a limited quantity of his parameter α was distributed along or around a curve in such a way as to minimise a certain cost function then in certain cases: the α-distance between two points in image-space corresponded to the *real* linear or angular distance between those two points in 3-space. Our system has similar properties in respect of similarly simple shapes under orthographic projection. The example of the circle/ellipse discussed above being a case in point.

Weiss's paper is somewhat technical and a metaphor may be deployed to explain his minimisation problem. We have a vehicle which we have to drive around a closed track in a certain amount of time. The rate at which we turn the vehicle (through the points of the compass) costs us - in tyre-wear shall we say. The speed at which we travel also enters into the cost calculation. In minimising cost therefore we cannot travel infinitely fast or change heading infinitely fast. A little reflection which reveal that: we must be stationary while turning at points of infinite curvature, we will travel more slowly the higher the curvature and that we will travel fastest - but not infinitely fast - along the straight bits. Our precise timetable depends on the relative weighting of cornering-cost and speeding-cost (though this unfortunate fact is obscured in Weiss's paper).

It does not appear straightforward to relate Weiss's variational principle to properties of the Fourier series that we would use to encode the curve. There is another principle which can be simply related however - the minimisation of the quantity:

$$\int_0^{2\pi} \left[\frac{d^2x}{d\theta^2}\right]^2 d\theta + \int_0^{2\pi} \left[\frac{d^2y}{d\theta^2}\right]^2 d\theta$$

This is the integral over "time" of the square of "acceleration". It is simply the sum of the second-order energies of $x(\theta)$ and $y(\theta)$ which, in Fourier terms, is:

$a_1^2 + 16a_2^2 + 81a_3^2 \ldots$

+ corresponding energy term for second series

Minimising second-order energy thus implies a strong bias *against high order harmonics* in describing the shape.

IV

Kass et al [4] describe their snake as "an energy-minimising spline guided by external constraint forces and influenced by image forces that pull it towards features such as lines and edges". The reason for their using a "distributed" model of a spline is that "the geometric coverage...can be significantly broader than the *lumped* parametric families of shapes, including the superquadric geometric models...." [5]. Superquadrics [8] are in effect used as (3-dimensional) snakes in the geometrical modelling/fitting of Bajcsy and Solina [9] and the limitations of this parametric form are indeed evident. A rather nondescript "blob" (in two or in more dimensions) might be impossible to describe in the superquadric formalism without segmentation - whereas the the Kass snake and its higher-dimensioned variants are infinitely deformable. But the cost of distributed descriptions is that they do not efficiently encode or "summarise" the data. There are one or two global quantities to be derived from a given conformation of the Kass snake (overall energies etc.) but these do not determine its shape. This is encoded as a list of points.

My Fourier system shares with the Kass snake the feature that it is infinitely deformable and with superquadrics the feature that a global description - often a very compact one - is to hand. This suggests that we might study its properties *as* a snake.

Kass et al include first-order energy (which disposes towards constant speed parameterisation) and sometimes second-order energy (which disposes towards high parameter densities at points of high curvature) in their total energy functional. In what follows I will not use an explicit energy term but simply restrict the model to a particular form (e.g. symmetrical about y-axis) and order (e.g. no terms above 3rd harmonic). Truncating a potentially infinite function at some sensible point is in fact the "regularisation" measure [10] best known to, loved by, and practiced by us all and it remains to be seen whether more sophisticated regularisation will bring many benefits.

Let me emphasise at the outset that I do not believe that gradient-based methods are an effective and generalisable means of matching an image with a model - or even with another quite similar image. The problematic history of optic flow methods [11][12] - in which one image is effectively used as a "snake" to be fitted to the other - have surely taught us this. I deploy my figures as gradient-climbing snakes so as to provide a simple framework for the discussion of the relationship between my method of explicit *global* shape description and the sort of "variational" approach employed by Kass et al. (Their snake can be intuitively understood as having a mixture of the properties of plasticine, rubber and high-strain steel wire.)

Figure 6 features a shell-like outline generated by a bilaterally symmetric function of order 4. The off-centred circle superimposed on the first outline is the starting position of a snake - itself described by a symmetric function of order 4 in which, initially, the coefficients are set so that it is is a circle. A field of "distance to the nearest point" has been

set up by the brush-fire technique. In this all the edge pixels are "lit" and the fire propagated iteratively both inside and outside the shape according to the simple formula (at each iteration): if a pixel was "lit" at the previous iteration then light up any neighbouring unburnt pixel. The iteration number at which any pixel is ignited provides a crude metric (bearing in mind problems to do with the rectangularity of the pixel configuration) of that pixel's distance from the nearest edge-pixel.

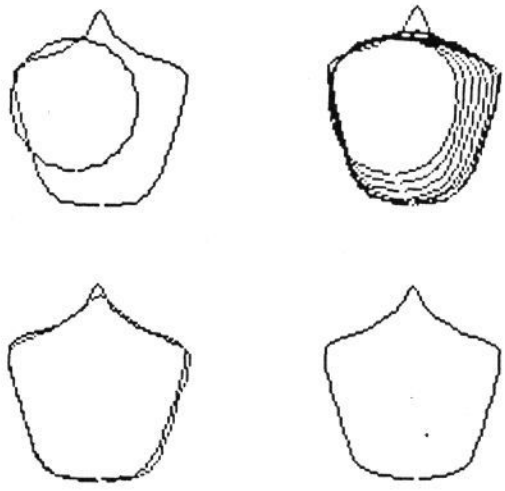

Figure 6: Snake conforming to target outline

The distance-field is then convolved with two Sobel filters to provide the slope, at each pixel, in x- and y-directions. Now the improvement in the "fit" of the snake which results from an infinitesimal change in any of the coefficients a in $x(\theta)$ is:

$$\delta(FIT)/\delta a = \sum -(\delta x(\theta)/\delta a)g_x$$

where summation is over θ on the range $0 \leq \theta \leq 2\pi$ and where g_x is the slope in the x direction at the point $x(\theta)$. Similarly for the coefficients in $y(\theta)$.

At each iteration of the fitting procedure we compute the partial improvement in FIT with respect to each Fourier coefficient in both series and have as a result the direction of steepest descent in coefficient space. The coefficients of the series are moved by some small amount in this direction - the snake is moved accordingly - and the procedure repeated.

Figure 6 shows a succession of snake positions as it "locks on" the the broad outline of the figure, slowly accomodates the small cusp on the target figure and, eventually, merges itself with it.

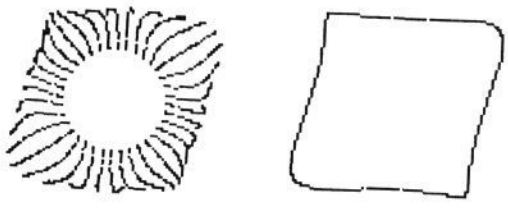

Figure 7: "Squaring the circle"

Figure 7 gives another example of a fit to a synthetic image - this time to the "parallelogram" of figure 5. In this case bilateral symmetry is broken (the steepest descent process described above is easily generalised to the non-symmetrical case by writing the Fourier series in the more conventional mixed sine/cosine form). The "potential well" in this case has been set up by a blurring process rather than by the brushfire technique. The initial values in the image are 1 at an edge-point and zero elsewhere. Gaussian smoothing is applied at a suitable scale and the resulting value I at any pixel converted to "potential" as $-\log(\varepsilon+I)$ where ε is a small constant introduced to avoid arithmetic errors. In the case of a single target point this gives rise to a well in which potential varies approximately as x^2+y^2 - as in the harmonic oscillator of quantum physics.

Note how the initially uniform distribution of θ is adjusted to the appropriate "dense at corners" distribution. Note also that the trajectories of some points "overshoot" before settling down on the edge of the target figure. Both the target and the snake are of order 3 without symmetry restrictions.

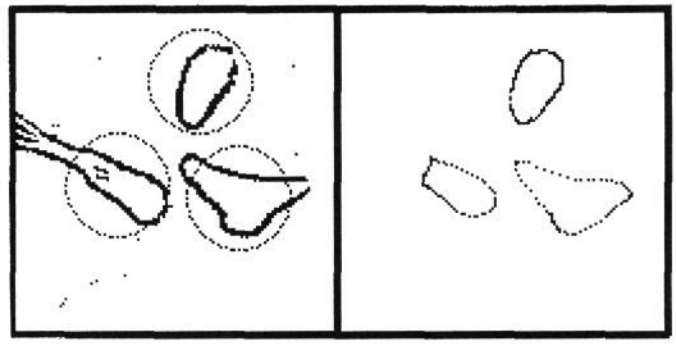

Figure 8: Snakes modelling real biological cells

Figure 8 shows the results of running three snakes on crude edge data derived from a phase-contrast micrograph of human cells in culture (kindly provided by the Imperial Cancer Research Fund). The snakes are of order 3 (except in the case of the triangular cell where it was raised to order 4 in an unsuccessful attempt to obtain a very precise fit).

The elogated cell has not been "captured" - rather it has been segmented. This is a typical consequence of gradient-based fitting. Due to the isotropic nature of brushfires and other diffusion type processes information concerning distance edge-points is drowned by information from nearer ones. (The potential is effectively determined by the *nearest* edge point). The snake cannot "smell" the distant part of the cell since it is everywhere rather close to some "edge point".

V

This problem of "obscuration" arising in potential fields derived by diffusion processes is extremely difficult to overcome! It is closely related to problems encountered in locating multiple symmetry axes by diffusion processes (see the discussion in Brady and Scott [15]). Figure 9 shows a snake "stuck" in a potential well generated by the brushfire method. The target figure is synthetic and

of the same order as the snake - so the impediment to progress is not that the snake lacks the required variability. It is that the potential gradient *along* the neck of the figure *in* the neck of the figure is zero. In the case of a well generated by the Gaussian method it would not be quite zero but would be down "in the noise" in the case of a real image.

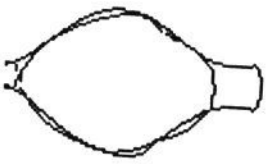

Figure 9: The blind snake

The important difference between my approach and that of Kass has nothing to do with the technicalities of generating attractor fields or of starting (and rescuing!) snakes. It is the fact that I have a compact global description of the spline. Not only is this something we *must* have if we are to proceed with higher-order processing but it gives us more direct control during the fitting process.

VI

Obvious requirements of complete visual processing include: inferring in some way the "real coordinates" of points which project to the image; and "unifying" a single phenomenon into a single cohesive perceptual structure. It is conventional in computer vision to leave these two nasty problems to mysterious "high level processes" - as if these could somehow be endowed with powers logically denied to "low level" processes!

It may be the case that image interpretation must be dependent upon the application of specialised "knowledge" of a propositional and procedural sort. But the need for such expert intervention is magnified many times if the primary processes are equipped with weak, brittle systems of shape encoding and weak, brittle methods of instantiating descriptions.

There are a number of directions in which the approach to shape "apprehension" outlined in this paper is being developed. These include:

1) Extending the functional form to higher dimensions in both "parameter" space and "observable" space e.g. $\theta,\phi \rightarrow x,y,z,t$. The one-parameter, two-observable figure is of restricted practical interest in computer vision (or graphics).

2) "Back-projection" criteria - from image space into 3-space - form a rich area of investigation. We have seen that natural parametrisation may yield an estimate of "true separation" in some instances. Further, existing back-projection variational principles for contours all involve *increase of symmetry* in some sense (e.g. Kanade [13], Brady and Yuille [14]). Symmetry is something of a built-in feature of harmonic functions - though it can nonetheless be quite difficult to enforce in a numerical computation.

As I have indicated the "snake" approach to image interpretation has inherent limitations in that the snake must be started close to the target (relative to the spatial frequency of the image) if there is to be any chance of success. My own suspicion is that the snake is not a creature that can be harnessed into doing hard work. It may, though, contain within it the germ of a really good idea.

REFERENCES

[1] Thompson, D'Arcy *On Growth and Form*, Cambridge University Press 1971

[2] Hollerbach, J.M. (1980), "An Oscillation Theory of Handwriting" *MIT AI-TR 534*

[3] Raibert, M. (1986), *Legged Robots That Balance*, MIT Press Series in Artificial Intelligence

[4] Kass, M., Witkin, A. and Terzopoulos, D. (1987), "Snakes: active contour models", *Proc. First ICCV p. 259*

[5] Terzopoulos, D., Witkin, A. and Kass, M. (1987), "Symmetry-seeking models for 3D object reconstruction", *Proc. First ICCV p. 269*

[6] Turing, A.M. (1952), "The chemical basis of morphogenesis", *Phil. Trans. Roy. Soc. B* Vol 237 p 37

[7] Weiss, I. (1985), "3-D shape representation by contours", *Proc. 9th IJCAI* pp 969-972

[8] Pentland, A.P., "Parts: structured descriptions of shape", *Proc. Nat Conf on Artificial Intelligence, AAAI-86*, Philadelphia, PA, 695-701

[9] Bajcsy, R. and Solina, F. (1987), "Three dimensional object representation revisited", Proc. First ICCV p. 231

[10] Tikhonov, A.N. and Arsenin, V.Y. (1977), *Solutions of Ill-Posed Problems*" Winston & Sons, Washington, D.C.

[11] Horn, B.K.P. and Schunck B.G. (1981), "Determining optical flow", *Artificial Intelligence 17* pp 185-203

[12] Hildreth, E.C. (1984), *The Measurement of Visual Motion*, MIT Press, Cambridge, Mass

[13] Kender, J and Kanade, T. (1980), "Mapping image properties into shape constraints: skew symmetry, affine transformations, and the shape-from- texture paradigm", *Proc. Am. Assoc. Artif. Intell. (August) pp 4-7*

[14] Brady, J.M. and Yuille, A. (1984), "An extremum principle for shape from contour" *IEEE Trans. Patt. Anal & Mach. Int.*, PAMI-6

[15] Brady, J.M. and Scott, G.L. (1987), "Parallel algorithms for shape representation", in *Parallel Architectures and Computer Vision* (Ian Page ed.) Oxford University Press (in Press)