Derivation of "natural basis functions" for a group of shapes

Guy Scott

Oxford University, Oxford.

Abstract

A highly efficient means of describing curves, surfaces or other configurations in space (or space-time) is to express the position vector as a sum of functions defined over some some interval in a space of one or more "implicit variables". The most familiar forms of these basis functions are the polynomial, trigonometric and superquadric. There seems no reason, however, why we should be limited to functions having a simple analytic form. In this paper I describe a method of deriving "good" basis functions empirically from observation of a *group* of shapes - or of a single shape showing a high degree of self-similarity.

There is an intimate connection between the search for natural basis functions and the study of transformation between and within shapes. Some light is shed upon the "special case" of the Iterative Function System and also - perhaps unexpectedly - upon the approach to self-organisation in neural networks associated with the name of Kohonen.

1 Introduction

It sometimes seems that computer vision consists of an attempt to answer two, and only two, questions:

- How do you solve the long-range correspondence problem?
- What would you do if you could?

By "long range" I mean to imply the sort of underconstrained, multi-dimensional search that confronts you if you want to match a 3-D model to an image; if you want to mimic human abilities in motion fusion; if you are doing stereo without precise calibration; or if you are looking for non-local structural regularities (such as symmetry or periodicity under distortion). Straightforward search is ruled out by the combinatorics - and hill-climbing by the mögelesque topography of any cost-function you might care to define. The state of the art has an "interactive" aspect - meaning that a human operator must be in the circuit to rescue the program from itself from time to time.

There is nothing very sophisticated about the correspondence techniques used in the experiments reported in this paper. They are sufficiently good to yield the correct result in a number of noise-free and relatively "friendly" cases. The focus is on "natural" shape description and how it might be acquired in a perceiver by experience of a variety of shapes viz. it is the second great question of computer vision that we are addressing.

2

We deal only with shapes in the form of closed, S0 continuous curves in two dimensions and their description in terms of the sum of basis functions defined over a single continuous parameter θ. Open curves, space curves, discontinuous curves, surfaces, solids etc. may be treated in an analogous way; the restrictions are made to keep visualisation simple and the paper short.

$$x = a_0G_0(\theta) + a_1G_1(\theta) + a_2G_2(\theta) + a_3G_3(\theta) + \dots \quad (1)$$

$$y = b_0G_0(\theta) + b_1G_1(\theta) + b_2G_2(\theta) + b_3G_3(\theta) + \dots \quad (2)$$

This is the general parametric form under discussion. We assume that the number of functions G in use is finite. What should the G's be, though? We can, without loss of generality, require that they are orthogonal over θ. We might also hazard that a "natural" choice of G_0 under almost any circumstances is that it be a constant. But can we specify any other desiderata?

In 1987 I introduced the "alternative snake" - a spline in which the G's are Fourier components up to some fairly low order - and showed that it gave concise descriptions of some biological shapes. (The common parametric form for the ellipse is an order 1 alternative snake). The Fourier form handles cusps rather easily but it is not very efficient for rectilinear shapes such as rectangles and triangles. For these we might think of using fractional Fourier components (as with superquadrics) or simply piecewise linear functions. But many things in this world - even if we restrict ourselves to closed 2-D curves - are neither piecewise linear nor, I regret, economically described by the alternative snake. Rather than search for the "golden" set of basis functions (it almost certainly doesn't exist) perhaps we should explore a horses-for-courses policy.

A "good" set of basis functions (or other types of primitive) can only be defined, I would suggest, in terms of of the world in which it is to be used. Given the world of, say, crabs we would require of an optimal system of

AVC 1989 doi:10.5244/C.3.8

shape encoding that it was economical and had good constancy properties. Thus: we would not expect to have to change our description entirely as a young specimen matured into an adult; or as it moved its legs; or as we shifted our attention to other individuals of the same or a related species; or as we shifted our viewpoint....

We would not be concerned if our crab-system proved to be inefficient and inadequate when we turned our attention to the world of large ungulates, say. Things might be different, however, in a universe in which it was the nature of crabs to slowly metamorphose into horses. Then we might have to learn to see things differently.

Seeing is an activity normally undertaken for some purpose. However interesting a horse may be to the differential topologist, to the more general type of human being it offers a range of affordances that includes: being bitten or kicked by it, eating it, riding it, pulling something with it, racing it, wearing it and worshipping it (the Trojans did all these, including the last). Trainers, jockeys, butchers, hunters... presumably these have horse-perception systems that are "tuned" to elicit various possibilities for action. There seems every good reason why the "geometry of seeing" should be closely coupled to the "geometry of behaviour" both in nature and in robotic systems. An interesting discussion of this coupling in the context of hand tools is to be found in Brady and McConnell's discussion of the "mechanic's mate". Unfortunately there is no room for any more ecological optics in *this* paper!

3

Suppose I am given two closed curves and told that they stand in a relationship of affine transformation. Can I check this information, and can I make some use of it in the future? I must, of course, establish correspondence between the two shapes (which point goes to which). Let us assume for the moment this is easily done. The assertion that the shapes are an affine pair implies that there exist at least two *distinct* linear expressions equating to zero for all sets of corresponding points.

$$\mathbf{a} \cdot \mathbf{p} = 0 \qquad (3)$$

where $\mathbf{p}$ is the vector $[1, x_1, y_1, x_2, y_2]$. The subscripts on x and y refer to the two curves. $\mathbf{a}$ is a non-null vector. It can be easily shown that values of $\mathbf{a}$ may be generated by determining the eigenvalues and eigenvectors of the 5x5 matrix $\sum \mathbf{p}\mathbf{p}^T$, where the sum is over sets of corresponding points. If the two curves are truly affine transforms of each other then there will be two eigenvectors $\mathbf{a}$ satisfying (3) associated with two zero eigenvalues. (Special cases excluded..)

We will call an $\mathbf{a}$ associated with a zero eigenvalue an *invariant* vector. There are three more eigenvectors - in general associated with non-zero eigenvalues - that we will call *basis* vectors. For these it is the case that the dot product $\mathbf{a} \cdot \mathbf{p}$ is not in general zero, but varies from one set of corresponding points to another. Let us refer to them as $G_0(\theta)$,$G_1(\theta)$, $G_2(\theta)$ - in descending order of eigenvalues let us say. What is θ? It is simply a label - a means of distinguishing one cluster of corresponding points, or cluster of values for the three G's, from another. There are numerous practical issues to be addressed but there is no *canonical* mapping between the labelling parameter and position, arclength or whatever. We do not care if our basis functions are analytic - only that the right combination of values is spanned as we travel through the range of θ.

The functions G generated from the basis vectors may serve to expand either shape - or *any* shape related to them by affine transformation - in the form (1). Note that the functions themselves are not necessarily orthogonal over θ, because of its arbitrarily defined density. (I apply Gram-Schmidt orthogonalisation before further use). Note also that there is no reason to expect them to break cleanly into a constant term and varying functions. The procedure just described cannot "learn" that position is a quality independent of deformation unless these are, in some well-defined empirical sense, independent. (If a shape changes its position much more often than it distorts, for example, then there are simple procedures for separation).

4

Suppose you give me three curves and tell me that they constitute an affine trio. To test this I establish (3-way) correspondences and augment the vector $\mathbf{p}$ by the coordinates of the third shape. My 7x7 covariance matrix should now yield *four* zero eigenvalues; they correspond to the mappings between two pairs of shapes, the mapping between the third pair being not independent. I still recover three basis vectors which should define for me three basis functions G which will simply be linear combinations of the G's I would obtain for any pair of the shapes. I note in passing that we may in this case eliminate the constant term from $\mathbf{p}$ since the system now has enough information to eliminate the effects of translation, as well as affine distortion, by linear combination of the three shapes. From the result 6x6 variance-covariance matrix I would obtain three invariants and, again, three basis vectors. I will assume, however, that the constant term is "given".

What if I find that there are fewer than four zero eigenvalues? Then you have lied! But the situation may not be uninteresting. If I have two zero eigenvalues then it may be the case that you have given me two affine pairs and a third, independent shape. Or the three may be inter-related in a more complex way. In any event the number of non-zero eigenvalues tells me precisely how many basis functions I need in order to expand all the shapes - and the corresponding eigenvectors give me the means to construct a set.

Given a family of shapes generated from a finite set of basis functions I can determine those functions (or linear combinations thereof) from observation of a sufficient number and variety of members of the family. This is provided, of course, I can establish the correct correspondences.

5

The correspondence problem is reflexive: the global results follow from the correct matches; the correct matches are the ones that give the best results. Iterative swapping back and forth between local and global processes seems to be the order of the day.

To explore the ideas of this paper in a quantitative and visual way I have written a program that takes a set of curves as data and then attempts to find a limited set of basis functions with which it can expand them all. The multiple correspondence is mediated through a single parameter θ. Each input shape is assigned a "snake" - a spline defined as a mapping from θ to image space. The snakes are alternately:

1. Subjected to a spell of "data drive" during which randomly chosen data points identify the nearest point on their snake and pull it towards them. Gradient-based methods might serve as well but I was interested in studying the stochastic method - which is essentially that used by Kohonen in establishing compact mappings between parameter spaces and vector spaces of observables. Some smoothing of the spline is necessary in the early stages to avoid points being "left behind".

2. "Basis function limited". Given a set of snakes in a variety of configurations we can construct a set of basis functions in which they can be expanded. We are interested in *limiting* the allowable variation, so we take only the "best" N basis functions and fit them in a least squares sense to each of the snakes. They may be further changed by "data drive" from their new shapes.

The procedure is iterated until no significant change occurs throughout a cycle. Figure 1 shows the outlines of four human fibroblast cells. Figure 2 shows the succession of snake positions as they "lock on" to these outlines. The starting position in each case is the circle roughly centered over each shape. The number of basis functions was here restricted to five - shown in figure 3. Note the (almost) constant function and the "quadrature pairing" of the other four. Although they resemble sinusoids they are significantly different from them. They represent something of a compromise between "rounded" functions and piecewise linear ones - as a result of their having to cope with both smoothly curved and angular structure.

Figure 4 shows the resulting "fit" to the fibroblasts when the number of basis functions is limited to three. This implies that the fitted curves must all be affine transforms of each other - as careful inspection reveals them to be. Figures 5, 6 and 7 show the results for three families of synthetic shapes. Note that the basis functions all have characteristics "appropriate" to the family. Figure 7 consists of "alternative snakes" generated from a constant term and the first two pairs of harmonics. Note that the basis functions appear to be simply (mixed) sinusoids.

6

What is a shape but an assemblage of "sub-shapes" that have a tendency to resemble each other and - sometimes - their parent? There are many types of "self-similarity", using the term in a loose sense. There are symmetries - both "discrete" and "differential" - in which the whole shape maps back into itself. There are varying degrees of resemblance between sub-parts, such as that between my two hands, or my hands and my feet. There are certain recursive-type structures in which the parts are copies of the whole and to which the term self-similar is applied in its strict "fractal" sense.

Each thing - save the most trivial perhaps - is a world of smaller things between which we can seek similarity, and thence obtain invariants and basis functions.

[some empirical results here]

An interesting question concerns the conditions under which a set of internal invariants is "complete" in the sense that they completely define the shape. A well-understood case is the Iterative Function System in which a structure is defined by "tiling" it with affine copies of itself. Only the transformation parameters need be stored - since they define the shape uniquely! The standard example is the spleenwort (?) fern shown in figure 8 which can be seen as composed of four affine copies of the whole thing. The fern may be regenerated simply by moving a point around by applying the affine mappings randomly and repeatedly.

Our analysis suggests there may be useful possibilities of lifting the affine restriction on the IFS by "upping the dimenisonality" with hidden variables. The self-similar maps would then have the general form

$$\mathbf{r}' = \mathbf{A}\mathbf{r} \qquad (4)$$

where $\mathbf{r}$ is a position vector augmented with hidden variables (which might include the actual Euclidean third dimension).

figure 1

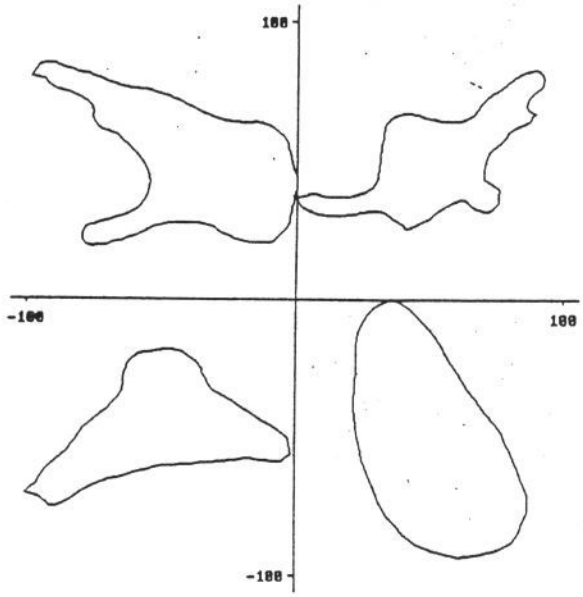

figure 2

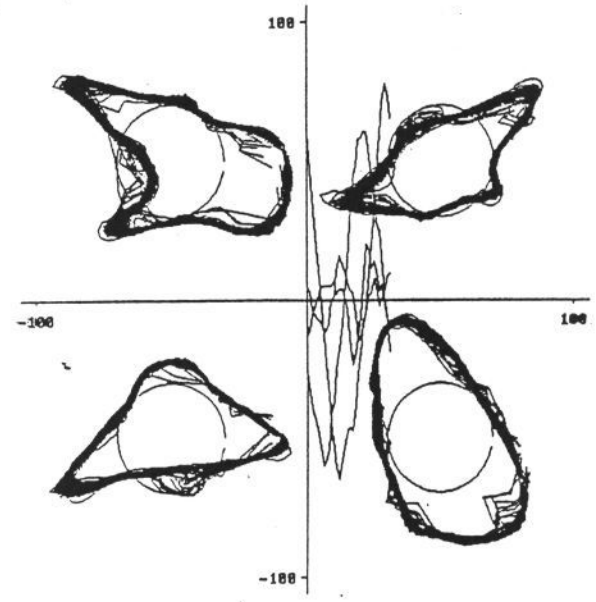

figure 3

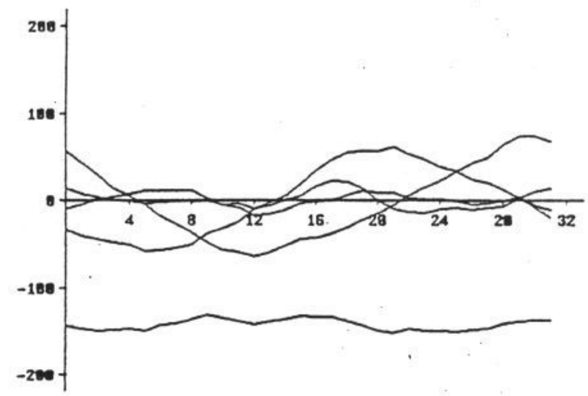

figure 4

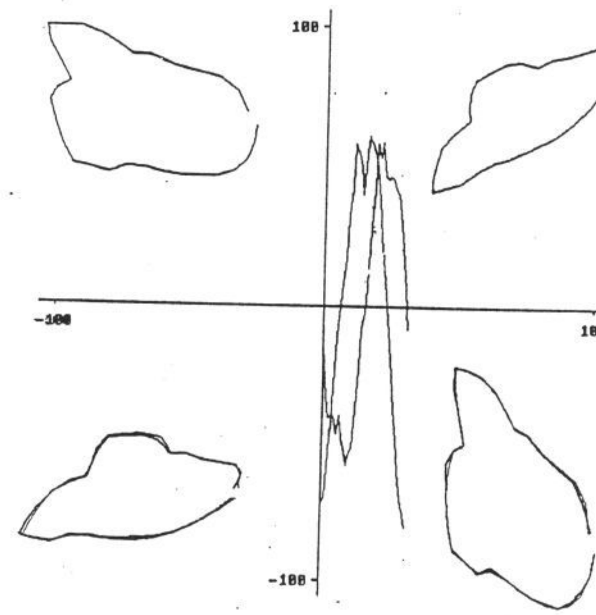

figure 5

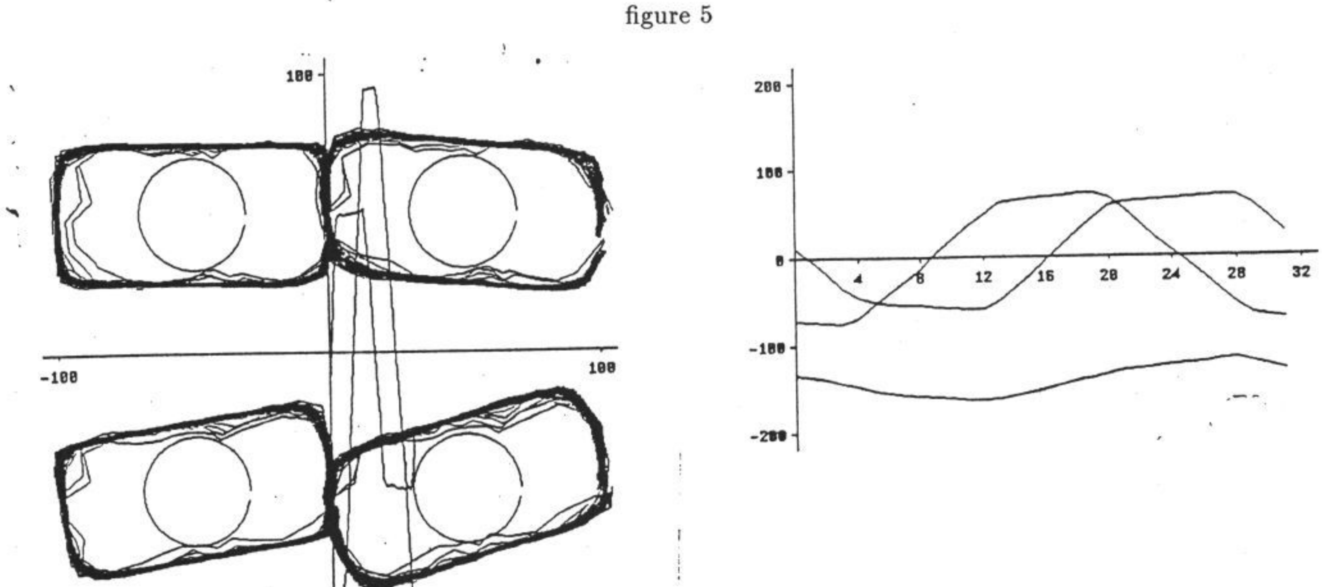

figure 6

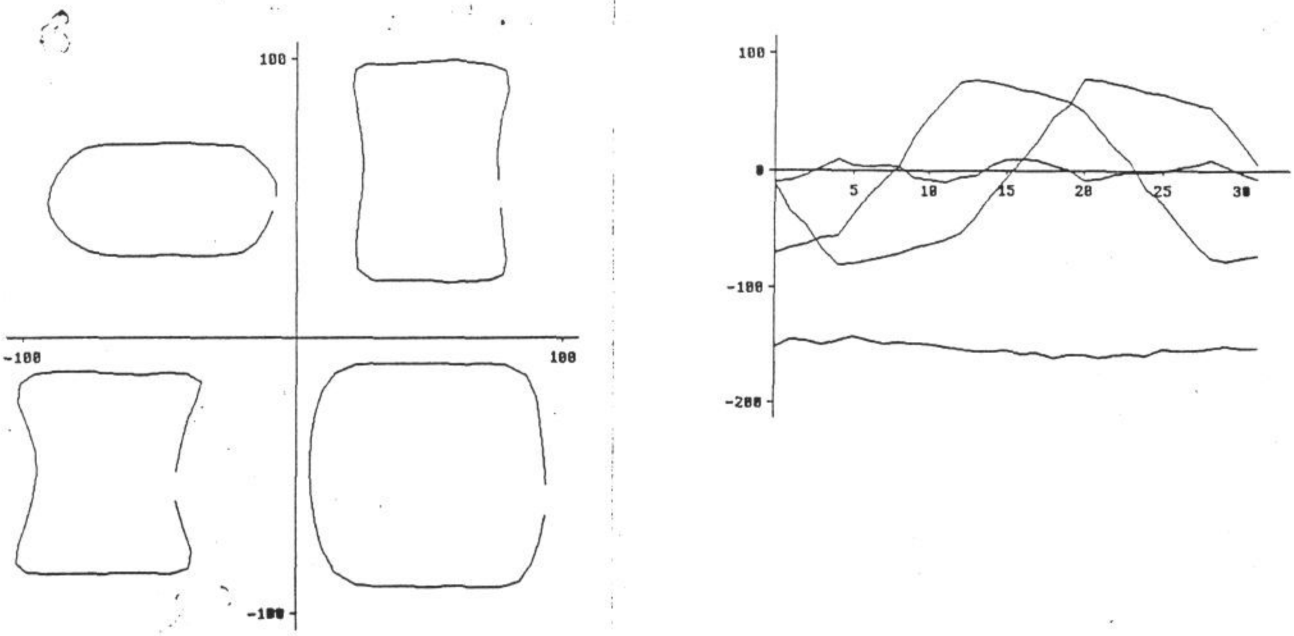

figure 7

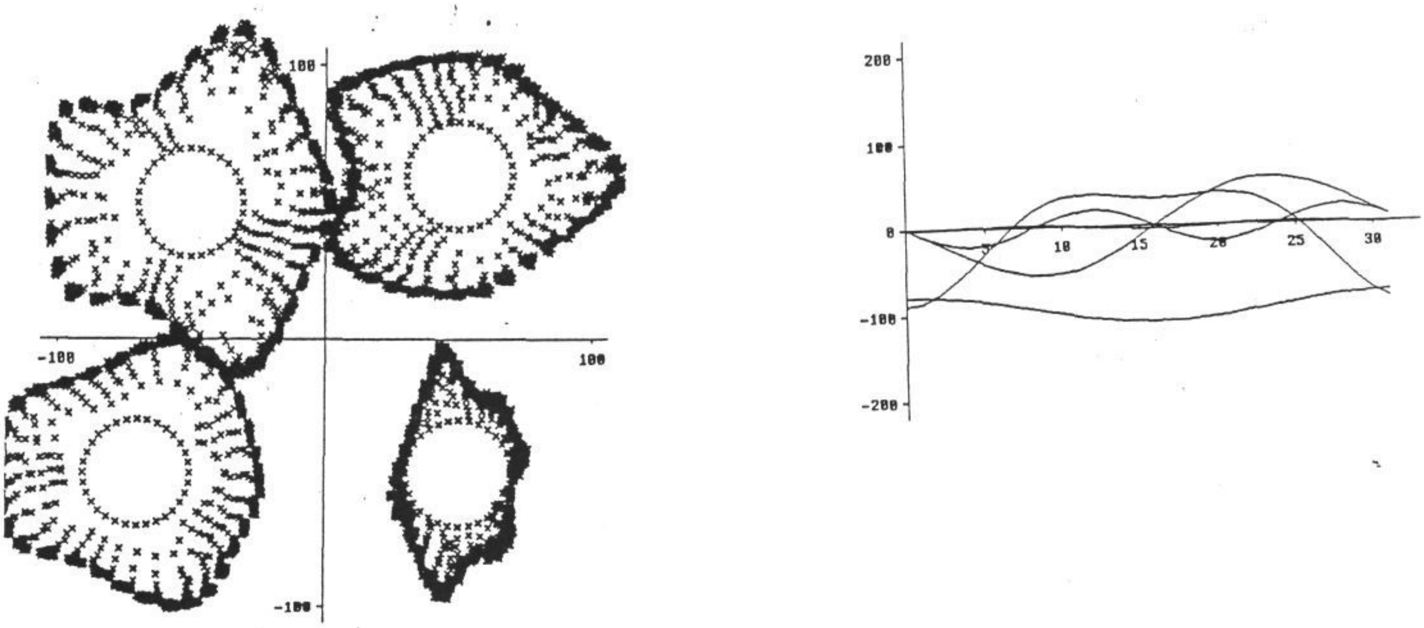

figure 8