Training PDMs on models: The Case of Deformable Superellipses

Maurizio Pilu Andrew W. Fitzgibbon
Robert B. Fisher

Department of Artificial Intelligence
The University of Edinburgh
5 Forrest Hill, Edinburgh EH1 2QL
SCOTLAND

Abstract

This paper addresses the following problem: How can we make a complicated mathematical shape model simpler while keeping a comparable level of representational power? The proposed solution is to use the original model itself – which represents a class of shapes – to train a Point Distribution Model. In this paper the idea is applied to the case of deformable superellipses.

1 Introduction

Superellipses and their 3-D extension superquadrics were introduced by the Danish designer Piet Hein (e.g. [5]); however, although he is referred as the inventor of superellipses, we have found that curves of the form $(x/a)^n + (y/a)^n = 1$, which include superellipses, were presented in 1818 by the French mathematician Gabriel Lamé.

These two representations have been brought into the computer graphics and vision community by Barr [1] and, in particular, Pentland [9], who used superquadrics to model parts of objects in a coarse but very compact way. Either simple or complicated deformations can be applied to extend their modelling capabilities (see, e.g., [13]), obtaining what are normally referred to as deformable superellipses (henceforth DSE) and deformable superquadrics (DSQ);

Although DSEs and DSQs can represent many closed 2-D and 3-D shapes in a straightforward and natural way by using few parameters, they are toys of a rather awkward nature, defined by horribly non-linear equations which make them slow to generate and not very manageable for fitting purposes. To date, in fact, there is no known closed-form error-of-fit (EoF) function and usually some shrewdness is used to compute *approximate* (yet expensive) EoF functions, as in [6] or [13]; practically, however, the fitting is often performed by minimising in parameter space a least squares EoF function which needs a closest-point search on the DSE contour or DSQ surface.

The central problem is whether we are willing to trade the complexity of DSEs and DSQs for the benefits they offer. In fact, in the choice of a model for fitting purposes, the foremost priority should be to find a good balance between ease of fitting and representational power. In this light, it is no wonder that the most used representations in vision are lines and ellipses, since they offer the best trade-off in this respect.

This paper addresses the following problem: How can we render a complicated mathematical shape model easier to fit while keeping a comparable level of representational power? The key idea is to use the mathematical model itself – which represents a class of shapes – to train a Point Distribution Model (PDM) [3, 2]. PDMs are linear, fast to generate and can be fitted both in parameter and image space.

In this papers, the argument is illustrated for the case of DSEs; we call the DSE-trained PDM thus created a *linear deformable superellipse*. The method can be trivially extended to DSQs and, in principle, applied to simplify or parametrise other complicated shape models, such as superellipses with more domain-specific deformations or high-order polynomials.

After a description of the DSE model given in the next section, the statistical PDM is briefly introduced. Then, we show how the PDM training set is built from randomly generated deformable superellipses and give some examples of the parametric shapes thus obtained. Finally, we conclude with a brief discussion and illustrate an application of the linear superellipse model taken from [10].

2 The deformable superellipse model

In this section we describe the deformable superellipse model, which will be subsequently used to train the PDM.

A superellipse can be described in parametric form by:

$$\begin{cases} x = f_x(\theta) = a_x cos(\theta)^\epsilon \\ y = f_y(\theta) = a_y sin(\theta)^\epsilon \end{cases} \qquad -\pi \leq \theta \leq \pi$$

where a_x and a_y are the two semi-axis and $0 \leq \epsilon \leq 1$ is the roundness parameter. By eliminating θ, its implicit equation can be easily obtained:

$$\left(\frac{x}{a_x}\right)^{2/\epsilon} + \left(\frac{y}{a_y}\right)^{2/\epsilon} = 1$$

Either simple or complicated deformations can be applied to the basic superellipse shape. For the sake of the self-containedness of this paper, we give below the mathematical description of the two simple deformations used in this work, linear tapering and circular bending, which have been derived from Solina and Bajcsy's work [13].

Let the superellipse shape $\mathbf{S}$ be expressed in terms of its vectors of coordinates $\mathbf{x}$ and $\mathbf{y}$ and let $\mathbf{X}$ and $\mathbf{Y}$ be the corresponding coordinates after the deformations.

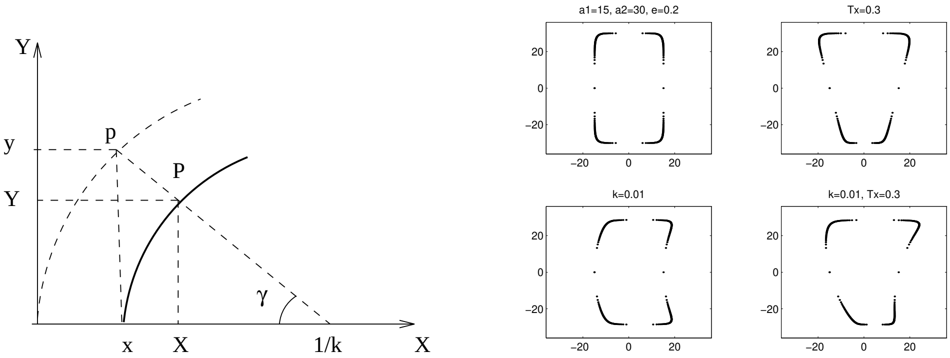


Figure 1: Bending geometry setting (left) and some examples of DSE (right). See text for details.

Linear Tapering: The tapering deformation along the y axis is defined as

$$Taper(\mathbf{S}, K) = \begin{cases} \mathbf{X} = g_x(y)\mathbf{x} \\ \mathbf{Y} = \mathbf{y} \end{cases}$$

If $g_x(y)$ is linear, the tapering will also be linear. By setting $g_x(y) = \frac{K}{a2} + 1$, with $-1 \leq K \leq 1$, we have linear tapering ranging from increasing ($K > 0$), constant ($K = 0$) to decreasing cross-section ($K < 0$).

Circular Bending: Figure 1 (left) sketches the geometry of the circular bending; only one parameter is needed to describe this deformation. As shown in the figure, a positive bending is applied along the y axis in the positive x direction. p is the original point position and P is the position when the deformation is applied. The deformation is given by:

$$Bend(\mathbf{S}, b) = \begin{cases} \mathbf{X} = \mathbf{x} + sign(b) * (\sqrt{\mathbf{y}^2 + r^2} - r) \\ \mathbf{Y} = sin(\gamma) * r \end{cases}$$

where:

$$R = a_y/|b|$$
$$r = R - |\mathbf{x}|$$
$$\gamma = atan(\mathbf{y}/r)$$

and $-1 \leq b \leq 1$ is the bending control parameter. Differently from [13], here the bending parameter is normalised to a_y and bending on both direction has been introduced.

Figure 1 (right) shows four superellipses, without deformation (top-left) with linear tapering (top-right), with bending (bottom-left) and with a combination of them (bottom-right). Note that in the examples, a linear sampling of θ is used in

Eqn. (2), causing a remarkably uneven contour sampling distribution. Solutions to this problem has been proposed in Franklin [4] and also in [11].

A combination of deformations should be carried out by first doing the deformations that are more shape preserving (see, e.g., [8] or [13]). In our case, with just two deformations used, the right order is first tapering and then bending.

3 The Point Distribution Model

Point Distribution Model (PDM) is a term coined by Cootes *et al.* [3, 2] to indicate statistical finite-element models built from a *training set* of labelled contour landmarks of a large number of shape examples. The method has recently received considerable attention because of its flexibility and generality.

Let us indicate by Σ_2^n the set of shapes defined by a labelled set of n two-dimensional points $P_i = (x_i, y_i)$, also called *landmarks*. We desire to model a certain class of similar shapes belonging to Σ_2^n in order to identify and parametrise their significant degrees of freedom.

A well known tool for achieving this dimensionality reduction is the Karhunen-Loeve transform, or Principal Components Analysis (PCA) [7], by which a relatively large set of examples is used to infer global statistical properties of the whole set.

From a set of examples, n landmark points are chosen, labelled and put in correspondence across the whole training set. Let[1] $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_s}$ be the N_s aligned shape examples, each represented by $2n$-long vectors of landmark coordinates:

$$\mathbf{x}_i = [x_{i,1} \quad y_{i,1} \quad x_{i,2} \quad y_{i,2} \ \cdots \ x_{i,n} \quad y_{i,n}]^T.$$

The mean shape is calculated by averaging each coordinate points, that is

$$\overline{\mathbf{x}} = \frac{1}{N_s} \sum_{i=1}^{N_s} \mathbf{x}_i$$

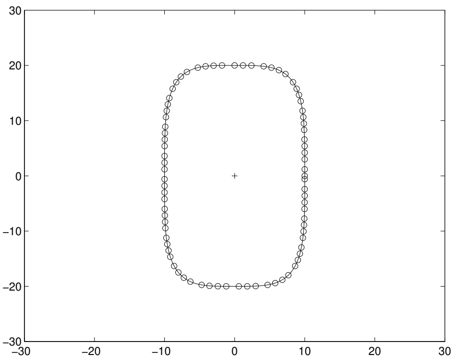
and the $2n \times 2n$ (positive definite) covariance matrix of the points is given by

$$\Lambda = \frac{1}{N_s} \sum_{i=1}^{N_s} (\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^T$$

Let $(\lambda_1, \mathbf{p}_1), (\lambda_2, \mathbf{p}_2), \ldots, (\lambda_{2n}, \mathbf{p}_{2n})$ be the eigenvalue-eigenvector pairs of Λ sorted such that $\lambda_i \geq \lambda_{i+1}$. As well known from statistics, the physical meaning of the eigenvector of a covariance matrix is a hyper-direction ($2n-$dimensional in our case) along which normal the variance of the point distribution equals the corresponding eigenvalue. Therefore the eigenvectors corresponding to the largest eigenvalues most describe the statistics of the point distribution.

This property of the eigenvalue decomposition is the key that has been cleverly used by Cootes *et al.*[3] for *approximating* any shape $\mathbf{x}$ in the training set by a

[1] In the following we shall use a similar notation as in Cootes *et al.* [3].



	λ_i	η_i
Mode 1	5692.0	56.9 %
Mode 2	1933.3	19.8 %
Mode 3	1509.9	15.3 %
Mode 4	615.2	6.2 %
Mode 5	91.5	0.9 %
Mode 6	50.1	0.5 %
Mode 7	48.3	0.4 %

Figure 2: Landmarks of the natural (undeformed) superellipse model (left) and contribution of each mode to the overall point variance over the training set (right).

weighted sum of displacements in the direction of the t most significant eigenvectors with respect to the mean shape, that is:

$$\mathbf{x} = \overline{\mathbf{x}} + \mathbf{P}\mathbf{b}, \tag{1}$$

where $\mathbf{P} = [\mathbf{p}_1|\mathbf{p}_2|\cdots|\mathbf{p}_t]$ and the weights $\mathbf{b} = [b_1 \; b_2 \; \ldots \; b_t]$ are called *modes of variations*.

Equation (1) allows not only to represent the training set but also to generate new shapes, thus being *de facto* a parametric model of the class of training shapes, provided that the b_i are kept within proper ranges.

By the nature of the decomposition used, each λ_i is the variance of the corresponding b_i over the training set and therefore the ranges for the b_i should fall within ± 2 or $3\sqrt{\lambda_i}$ [3].

4 Using random DSEs as training set

A properly built PDM can well represent the kind of variability needed to model shapes like DSEs in terms of dimension, bending and tapering, squareness and also shearing.

For doing this, however, a method for efficiently building a large set of samples has to be devised and, obviously, the most natural one is to *use the DSE model to train the PDM.* To this end, a number ($N_s = 2000$) of random superellipses were generated, their contours sub-sampled at equal distance (by the method proposed in [11]) and from the set of landmark points of all the training examples the PDM was built as in the previous section.

Figure 2-left shows a natural (undeformed) superellipse in canonical position with the landmark points. We used $n = 80$ (i.e. 20 landmarks in each DSE quadrant). By using the DSE construction as given in Sec. 2, the range of parameters used to generated the random training set was $-5 \leq a_x \leq 5$, $-10 \leq a_y \leq 10$, $-0.7 \leq K \leq 0.7$, $-0.7 \leq b \leq 0.7$ and, finally, $0.1 \leq \epsilon \leq 0.9$. Note that the absolute values of a_x and a_y is irrelevant, since the PDM will be rescaled during the

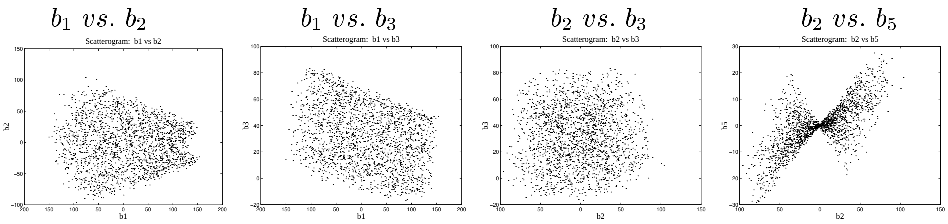


Figure 3: Four examples of scattergrams of the modes of variation. Low-order modes are relatively uncorrelated to each other whereas more and more correlation crops up for higher order modes.

fitting stage [2]; their ratio, however, expresses the the eccentricity of the shape, which is usually assumed elongated along the y axis.

The table in Fig. 2-right shows the contribution in percentage of the first ten modes to the total variance of the training set [7] given by:

$$\eta_i = \frac{\lambda_i}{\sum_{j=1}^{2n} \lambda_j}.$$

It can be seen that most of the shape variation is covered by the first 4 modes which, as we shall see in a moment, are strictly related to the actual height, width, tapering and bending parameters of the DSE; however, the first seven modes will be taken into account in the rest of the paper.

The principal component analysis method teaches us to use scattergrams to check correlation between the modes over the training set: a scattergrams of two modes should look like a cloud of random points if they are uncorrelated [7].

Figure 3 shows four scattergrams of various modes of variation computed over the training set. In our case the first 3 modes (b_1 vs. b_2, b_1 vs. b_3 and b_2 vs. b_3 in Figure 3) look relatively uncorrelated but not for higher-order modes such as, e.g., b_2 vs. b_5 (right).

An interesting experiment that has been carried out, and reported here, was to relate the original deformable superellipse parameters – used for creating the training set – to the modes of variation in order to assess their reciprocal correlations. Figure 4 shows the scattergrams of the first seven modes $b_1, ..., b_7$ (rows) with respect to the five deformable superellipse parameters a_1, a_2, ϵ, K and b, which are represented in the columns; a conspicuous line-like pattern of points in the scattergram indicates strong correlation.

It can be seen that modes b_1, b_2, b_3 and b_4, chiefly correlates to a_2, b, a_1 and K, respectively, whereas they are pretty much uncorrelated to other parameters. This is a very welcomed behaviour, because it allows easy and natural classification of the shapes by just using modes straight away. As expected, the strong non-linearity of the roundness deformation – controlled by ϵ, which actually does not involve major structural change in the shape – is not clearly correlated with any mode, although slightly with b_7. The roundness deformation is strongly non-linear and therefore this behaviour was somewhat expected.

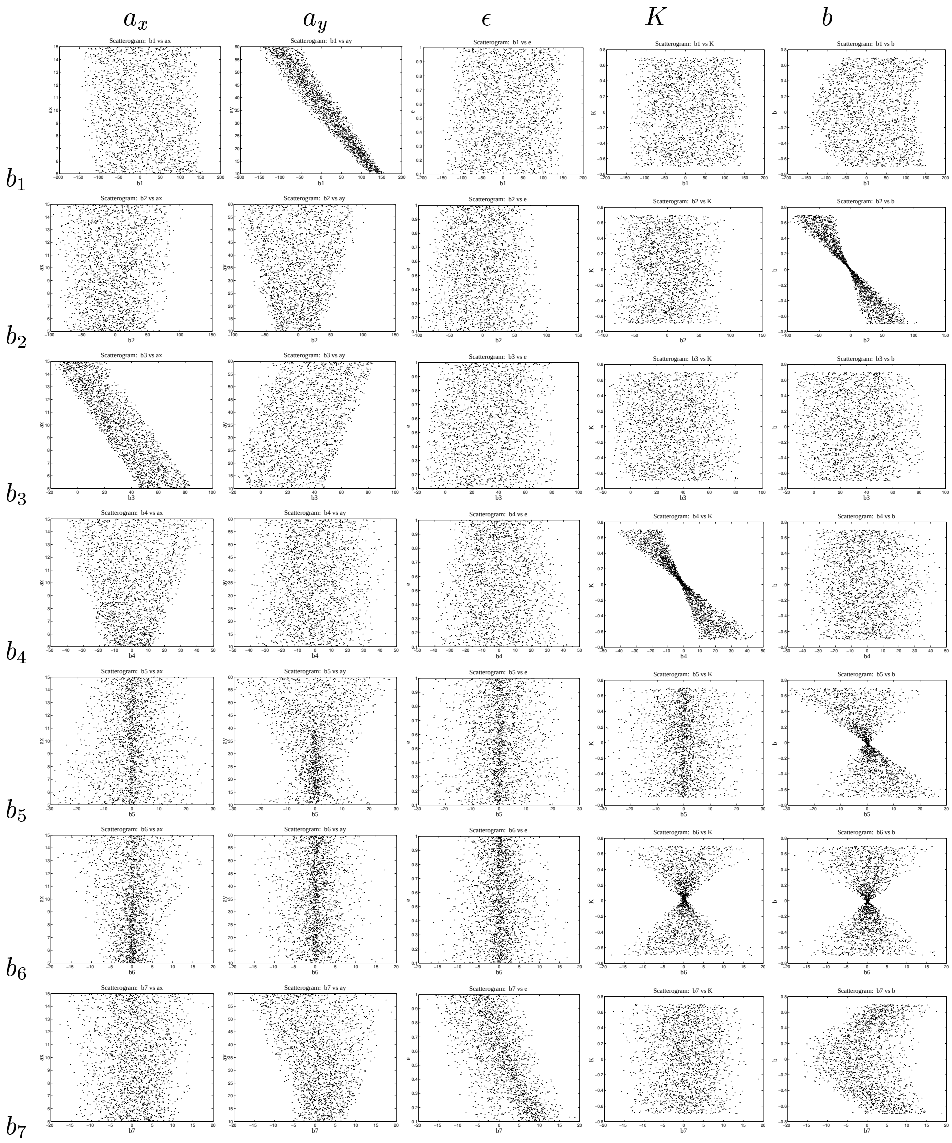


Figure 4: Scattergrams that relates the modes of variation to the original deformable superellipse parameters over the training set. High concentration of points around a line indicate high correlation. It can be seen that modes b_1, b_2, b_3 and b_4, chiefly correlates to a_2, b, a_1 and K, respectively, and pretty much uncorrelated to other parameters.

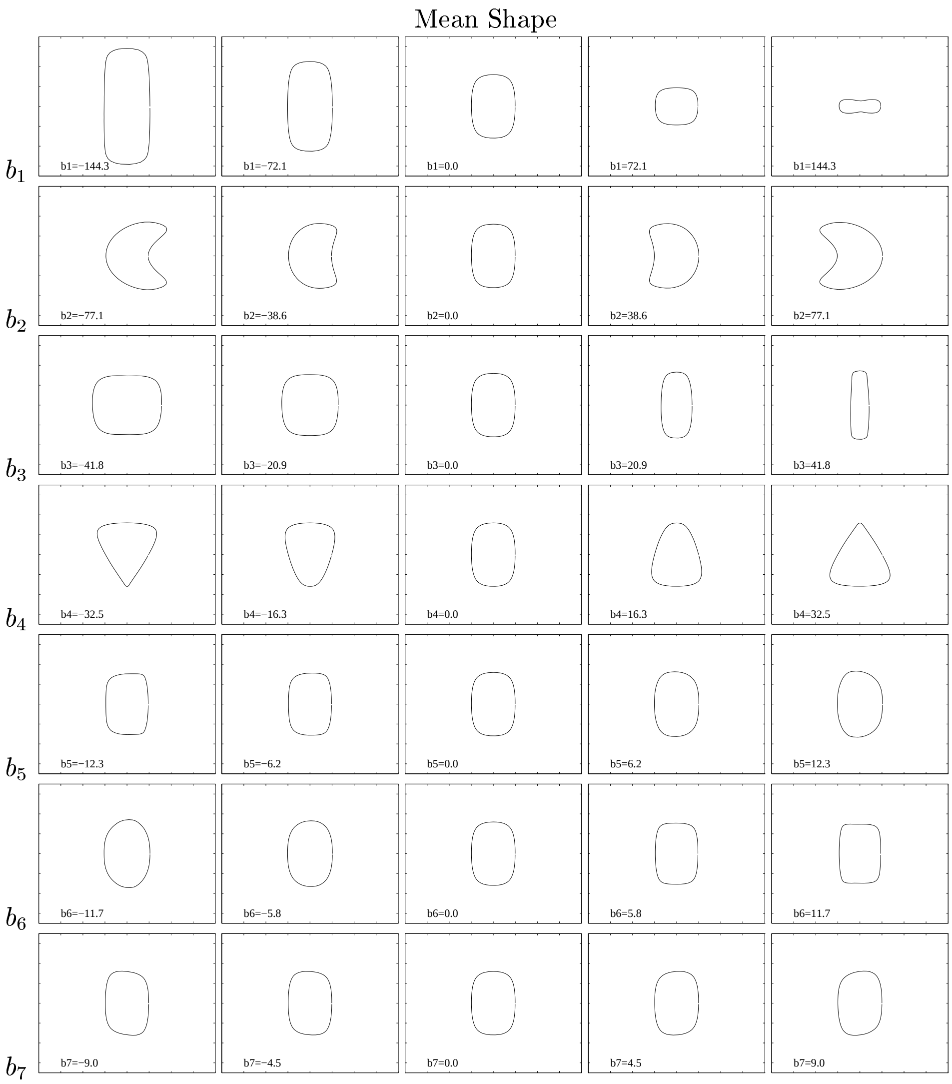


Figure 5: Parametrisation of the PDM model. The modes of variation controls the actual PDM shape in a rather neat way: the first four modes straightly control vertical height, bending, width and tapering, respectively, whereas the last three produce, in combination and unexpectedly, slight horizontal tapering, squaring and shearing.

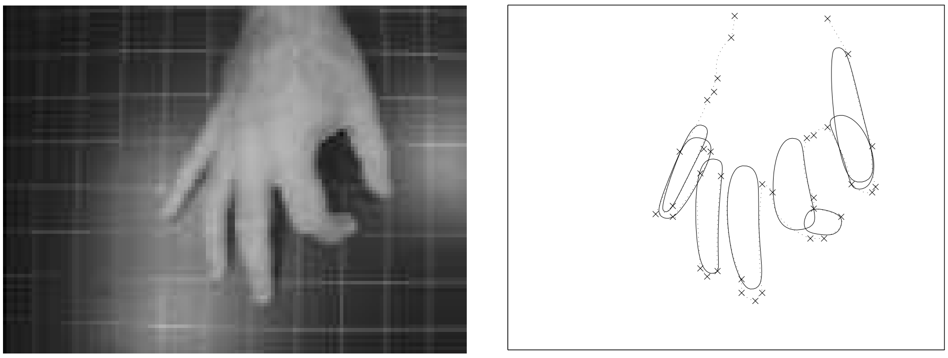


Figure 6: Example of use of the linear superellipse model fitted to some arbitrary object parts. More examples appear in these proceedings in [12]

This correlation between modes and shape parameters comes to the fore in Figure 5, in which the PDMs are shown for five different values of the first seven modes, one per row. The first four modes neatly control single shape features of height, bending, width and tapering, whereas the last three produce, in combination and unexpectedly, slight horizontal tapering, squaring and shearing, which also nicely enhances the model representational power.

Interestingly, there is a suggestive comparison of these results with Leyton's causal analysis of shape [8] (see also [13]), which proposes a natural order with which shapes are deformed. In fact, the contributions η_i of each mode to the overall point distribution variance indicate what are the most influential shape factors, which might give a solid mathematical basis to Leyton's theory, although at the moment it can be seen as a mere speculation.

5 Discussion

The spirit of the method proposed in this paper is general; it suggests that, whenever convenient, complicated mathematical shape models should be substituted with other with similar representational power to the foremost cause of fitting performance. For doing so, models such as PDMs can be used with a training set built up with examples of the original model itself.

In this paper the approach has been shown for the case of deformable superellipses. We have seen that the linear deformable superellipse model thus created has a similar representational power as DSEs and its shape features are controlled by parameters with a precise geometrical meaning.

Our concern was not to create a precise shape model, since very few objects, if any, can be exactly represented by DSEs and DSQs. As a matter of fact, a high degree of precision of representation is a lesser problem in generic shape analysis, which is the very domain DSEs and DSQs are normally used for.

In [10], the linear deformable superellipse has been used as a generic model for performing part-based grouping and segmentation. There, the fitting is carried out as in [2] but the PDM initialisations is performed by fitting ellipses to the set of pixels belonging to a small set of *seed codons*, as more extensively discussed in [12]. Figure 6 shows some fitting examples to a real image but some more examples can be found in [12] (these proceedings).

The extension to superquadrics is straightforward. In the near future, we also plan to apply the same model-trains-model strategy to other domains, for instance in the training on a shape class of high order polynomials (which can be fitted by closed-form least-square methods) in order to parametrise their dominant shape features.

Acknowledgements: Maurizio Pilu was partially sponsored by SGS-THOMSON Microelectronics.

References

[1] A. Barr. Superquadrics and angle-preserving transformations. *IEEE Computer Graphics and Applications*, 1(1):11–23, Jan. 1981.

[2] T. Cootes and C. Taylor. Active shape models - 'smart snakes'. In *Proceedings of the British Machine Vision Conference*, pages 266–275, 1992.

[3] T. Cootes, C. Taylor, D. Cooper, and J. Graham. Training models of shape from sets of examples. In *Proceedings of the British Machine Vision Conference*, pages 8–18, 1992.

[4] W. Franklin and A. Barr. Faster calculation of superquadric shapes. *IEEE Computer Graphics and Applications*, July 1981.

[5] M. Gardiner. The superellipse: a curve that lies between the ellipse and the rectangle. *Scientific American*, Sept. 1965.

[6] A. Gros and T. Boult. Error of fit measures for recovering parametric solids. In *International Conference on Computer Vision*, pages 690–694, Tampa, FL, Dec. 1988.

[7] I. T. Joliffe. *Principal Components Analysis*. Springer-Verlag, 1986.

[8] M. Leyton. *Simmetry, Casuality, Mind*. MIT Press, 1992.

[9] A. Pentland. Perceptual organization and the representation of natural form. *Artificial Intelligence*, 28:293–331, 1986.

[10] M. Pilu. *Part-based Grouping and Recogntion: A Model-Guided Approach*. PhD Thesis, Department of Artificial Intelligence, University of Edinburgh, Scotland, 1996. Forthcoming.

[11] M. Pilu and R. Fisher. Equal-distance sampling of superellipse models. In *Proceedings of the British Machine Vision Conference*, volume 1, pages 257–266, Birmingham, Sept. 1995.

[12] M. Pilu and R. Fisher. Part segmentation from 2D edge images by the MDL criterion. In *Proceedings of the British Machine Vision Conference*, Edinburgh, Sept. 1996. These proceedings.

[13] F. Solina and R. Bajcsy. Recovery of parametric models from range images: The case of superquadrics with global deformations. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 12(2):131–147, Feb. 1990.