An Efficient Method for Constructing Optimal Statistical Shape Models

Rhodri H. Davies[1], Tim F. Cootes[1], John C. Waterton[2], and Chris J. Taylor[1]

[1] Division of Imaging Science, Stopford Building, Oxford Road, University of Manchester, Manchester, M13 9PT, UK.
rhodri.h.davies@stud.man.ac.uk
[2] Enabling Science & Technology, AstraZeneca, Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK.

Abstract. Statistical shape models show considerable promise as a basis for segmenting and interpreting images. A major drawback of the approach is the need to establish a dense correspondence across a training set of segmented shapes. By posing the problem as one of minimising the *description length* of the model, we develop an efficient method that *automatically* defines a correspondence across a set of shapes. As the correspondence does not use an explicit ordering constraint, it generalises to 3D shapes. Results are given for several different training sets of 2D boundaries, showing the automatic method constructs better models than ones built by hand.

1 Introduction

Statistical models of shape show considerable promise as a basis for segmenting and interpreting images [4]. The basic idea is to establish, from a training set, the pattern of 'legal' variation in the shapes and spatial relationships of structures in a given class of images. Statistical analysis is used to give an efficient parameterisation of this variability, providing a compact representation of shape and allowing shape constraints to be applied effectively during image interpretation [5]. One of the main drawbacks of the approach is, however, the need - during training - to establish dense correspondence between shape boundaries over a reasonably large set of example images. It is important to establish the 'correct' correspondence, otherwise an inefficient parameterisation of shape can result, leading to difficulty in defining shape constraints. In practice, correspondence has often been established using manually defined 'landmarks'; this is both time-consuming and subjective. The problems are exacerbated when the approach is applied to 3D images.

Several previous attempts have been made to automate model building [1, 2, 3, 9, 10, 11, 12, 14, 16, 18] . The problem of establishing dense correspondence over a set of training boundaries can be posed as that of defining a parameterisation for each of the training shapes, leading to implicit correspondence between equivalently parameterised points. Different arbitrary parameterisations of the training boundaries have been proposed [1, 12] , but do not address the issue of

optimality. Shape 'features' (e.g. regions of high curvature) have been used to establish point correspondences, [2, 11, 18] but, although this approach corresponds with human intuition, it is still not clear that it is in any sense optimal. A third approach, and that followed in this paper, is to treat finding the correct parameterisation of the training shape boundaries as an explicit optimisation problem.

The optimisation approach has been described by several authors [3, 6, 9, 14, 16]. The basic idea is to find the parameterisation of the training set that yields, in some sense, the 'best' model. We have previously described a minimum description length criterion that describes a set of shapes as efficiently as possible [6]. We showed that, by optimising the parameterisation of each shape, we could produce models that were superior to those built by hand. The optimisation scheme was, however, inefficient and took many hours to converge. In this paper, we describe a more efficient method and also consider the pose transformation for each shape. As the method applies only an implicit ordering constraint, it can be used on 3D shapes.

2   Statistical Shape Models

A statistical shape model is built from a training set of aligned example shapes. Each shape, Ψ_i ($i = 1 \ldots n_s$), can (without loss of generality) be represented by a set of n points sampled along the boundary. These points are sampled according to the parameterisation, Φ_i of the shape. By concatenating the coordinates of its sample points into a $2n$-dimensional vector, $\mathbf{x}$ (where n is the number of points sampled on each shape), and using principal component analysis, each shape vector can be explained by a linear model of the form

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

where $\bar{\mathbf{x}}$ is the mean shape vector, the columns of $\mathbf{P}$ describe a set of orthogonal modes of shape variation and $\mathbf{b}$ is a vector of shape parameters. New examples of the class of shapes can be generated by choosing values of $\mathbf{b}$ within the range found in the training set. This approach can be extended easily to deal with continuous boundary functions [14], but for clarity we limit our discussion here to the discrete case.

The utility of the linear model of shape shown in (1) depends on the appropriateness of the set of boundary parameterisations $\{\Phi_i\}$ that are chosen. An inappropriate choice can result in the need for a large set of modes (and corresponding shape parameters) to approximate the training shapes to a given accuracy and may lead to 'legal' values of $\mathbf{b}$ generating 'illegal' shape instances. For example, consider two models generated from a set of 17 hand outlines. Model A uses a set of parameterisations of the outlines that cause 'natural' landmarks such as the tips of the fingers to correspond. Model B uses one such correspondence but then uses a simple path length parameterisation to position the other sample points. The variance of the three most significant modes of models A and B are (2.13, 1.16, 0.61) and (4.39, 1.56, 1.08) respectively. This

suggests that model A is more compact than model B. All the example shapes generated by model A using values of **b** within the range found in the training set are 'legal' examples of hands, whilst model B generates implausible examples - this is illustrated in Fig. 1 .
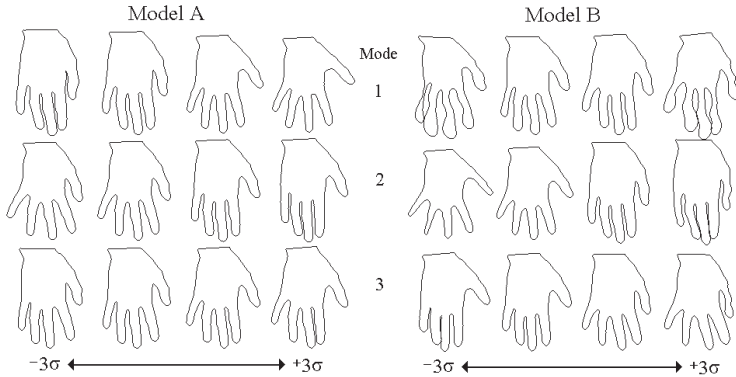

Fig. 1. The first three modes of variation ($\pm 3\sigma$) of models A and B

The set of parameterisations used for model A were obtained by marking the 'natural' landmarks manually on each training example, then using simple path length parameterisation to sample a fixed number of equally spaced points between them. This manual mark-up is a time-consuming and subjective process. In principle, the modelling approach extends to 3D, but in practice, manual landmarking becomes impractical. We aim to overcome these problems by building shape models *automatically*.

3   Automatic Model Building

We wish to optimise the parameterisations $\{\Phi_i\}$ of each shape in our training set $\{\Psi_i\}$. Since we wish to obtain a compact model with good generalisation ability we define the 'best' model as that which can account for the observations (the training shapes) in as simple a way as possible.

The configuration space of $\{\Phi_i\}$ is highly non-linear and has many local minima. Although stochastic optimisation techniques such as simulated annealing [13] and genetic algorithms [7] search for a truly *global* minima, they take many hours to converge. We overcome this problem by optimising $\{\Phi_i\}$ using a *multiresolution* approach. This allows a local optimisation method to be used at each resolution. We have used the Nelder-Mead simplex algorithm to produce the results in section 4.

3.1   An Information Theoretic Objective Function

To select a suitable objective function we must state the desirable properties of a statistical shape model. Ideally, we would like a model that is general (it can

represent any instance of the object - not just those seen in the training set), specific (it can only represent valid instances of the object) and compact (it can represent the variation with as few parameters as possible). We therefore choose to follow the principle of Occam's razor : the simplest explanation generalises best. In our case, we need to find the simplest explanation of the training set.

We choose to code the training set (to some given accuracy δ) with a linear shape model of the form:

$$\mathbf{x}_i = \bar{\mathbf{x}} + \mathbf{P}\mathbf{b}_i + \mathbf{r}_i \tag{2}$$

where $\bar{\mathbf{x}}$ is the mean of $\{\mathbf{x}_i\}$, $\mathbf{P}$ has m columns which are the m eigenvectors of the covariance matrix of $\{\mathbf{x}_i\}$ corresponding to the m largest eigenvalues λ_j, $\mathbf{b}_i$ is a vector of shape parameters, and $\mathbf{r}_i$ is a vector of residuals.

We have shown elsewhere [6] that a lower bound for the total information required can be obtained that is independent of the choice of representation of $(\bar{\mathbf{x}}, \{\mathbf{x}_i\}, \mathbf{P}, \{\mathbf{b}_i\}, \{\mathbf{r}_i\})$. The only free parameter is m, the number of shape modes in $\mathbf{P}$. The variable part of this lower bound is given by

$$F(m) = (n_p + n_s) \sum_{j=1}^{m} \log(\lambda_j) + [n_p(n_s - 1) - (n_p + n_s)]m \log(\lambda_{\mathbf{r}}) \tag{3}$$

where n_s is the number of shapes, $n_p = 2n$, twice the number of sample points on each shape and $\lambda_{\mathbf{r}} = \frac{1}{n_p} \sum_{j=m+1}^{n_p} \lambda_j$.

The first term is analogous to the determinant of the covariance of $\{\mathbf{x}_i\}$ used by Kotcheff and Taylor [14] and favours a model with much variation described by a small number of modes. The second term penalises a large number of modes and/or large residuals. The optimal trade-off between describing systematic variability using the model versus describing each shape individually can be determined by varying m exhaustively between zero and an upper bound on m given by $\lambda_m > 12\alpha\lambda_{\mathbf{r}}/(2\pi e)$, where $\alpha = (\frac{n_p n_s}{n_p(n_s-1) - m(n_p - n_s)})$.

F_{min}, the minimum of F with respect to m can be used to asses the quality of a given model.

3.2   Multiresolution Parameterisation

Our training data are a set of shapes $\{\Psi_i\}$ represented parametrically as curves: $\Psi_i(t), (0 \le t \le 1)$. We can manipulate the correspondences of the shapes by reparameterising these curves:

$$\Psi_i(t) \rightarrow \Psi_i(t'), \quad t' = \Phi_i(t), \quad \text{where} \quad \{\Phi_i : [0, 1] \rightarrow [0, 1]\} \tag{4}$$

We select n corresponding uniformly sampled points from the reparameterised shapes. The method described in this section is applicable to both open and closed curves; for clarity, we will limit our discussion to the closed case.

Each $\Phi_i(t)$ must be a homeomorphic mapping of the interval [0,1]. We use a piecewise-linear approximation of each parameterisation $\Phi_i(t)$ by specifying a set

of n_c *control points*, on each shape and linearly interpolating between them. The configuration space is therefore (sn_c)-dimensional. This search space is generally too large for a direct optimisation scheme to converge rapidly and reliably. We overcome this by using the following multiresolution approach:

- We begin with one control point, p_{1i}, on each shape. This point is allowed to assume any value in the range $[0, 1]$. We use linear interpolation to sample $n/2$ points between $[0, p_{1i}]$ and $[p_{1i}, 1]$, see figure (2a). We find the values of $\{p_{1i}\}$ that minimise F_{min}. Once these values are found, they are fixed and recorded.
- We place two additional control points p_{2i} and p_{3i} between 0 and p_{1i} and between p_{1i} and 1 respectively. We equally space $n/4$ points in the intervals $[0, p_{2i}]$, $[p_{2i}, p_{1i}]$, $[p_{1i}, p_{3i}]$ and $[p_{3i}, 1]$ (see figure (2b)). We fix and record the optimal positions of $\{p_{2i}\}$ and $\{p_{3i}\}$.
- We continue adding additional control points in a similar fashion between the fixed control points $\{p_i\}$ until the parameterisation is suitably defined. See figures (2c) and (2d).
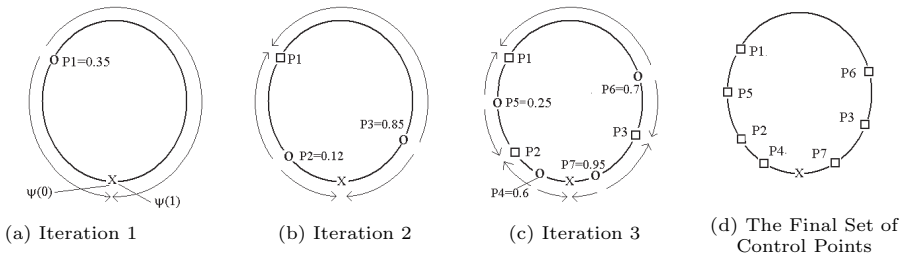

(a) Iteration 1   (b) Iteration 2   (c) Iteration 3   (d) The Final Set of Control Points

Fig. 2. A Demonstration of Parameterising a Circle. The 'X' represents the origin, the circles represent the current (flexible) control points and the squares represent the fixed control points. At each iteration, the current control points are allowed to move between the endpoints of the arrow.

At each iteration, the position of each control point is initialised as halfway along its allowed range - the equivalent of an arc-length parameterisation. As we have not used an explicit ordering constraint, the method may be used on shapes in 3D (see [6] for details).

3.3   Dealing with Pose Transformations

The pose of each shape affects the value of F_{min}. We therefore need to optimise the four parameters that allow a rigid transformation of each shape: translations d_x, d_y, scaling s and rotation θ. We have found that adding an additional $4n_s$ dimensions to each iteration significantly slows the optimisation and introduces many additional false minima. Better results can be achieved by performing a procrustes analysis [8] of the reparameterised shapes inside the objective function, before calculating the value of F_{min}.

4   Results

We tested our method on four different sets of object outlines. The algorithm was run for four iterations, giving 16 control points per shape. We compare the results to models built by equally-spacing points along the boundary and hand-built models, produced by identifying a set of 'natural' landmarks on each shape.

Our training sets consisted of manual-segmentations taken from:

Hand Outlines 17 images of a hand.

Rat Kidney 15 transverse, multislice T2-weighted magnetic resonance images (MRI) of rat kidneys. The repetition time was (TR) 2 sec; echo time (TE) 20 msec and slice thickness 1 mm. 41 contiguous transverse slices were acquired with a 64×64mm field of view and a $256 \times 256 \times 41$ image matrix. A single slice - containing most evidence of the collecting apparatus - was chosen from each image and the right kidney segmented.

Stroke Model 23 images collected from a previous study where permanent focal cerebral ischaemia was induced in rats [19]. For this study, only data from saline-treated animals were used. The experiment used in vivo multislice T2-weighted MRI as described in [19]. A single transverse slice was segmented, chosen with reference to an atlas of anatomy [15] and corresponded to an anatomic location 6.3 mm posterior to the bregma.

Knee Cartilage 15 T1-weighted MR images of the femoral articular cartilage [17]. A single sagital slice was chosen from the centre of the lateral femoral condyles. As the width of the femur varies from subject to subject, we identified comparable slices by selecting the slices halfway between (1) the first evidence of the lateral aspect of the meniscal horn and (2) the full extent of the posterior cruciate ligament.

In figure 3, we show qualitative results by displaying the variation captured by the first three modes of each model (the first three elements of **b** varied by $\pm 2\sigma$). We also give quantitative results in table 1, tabulating the value of F_{Min}, the total variance, and variance explained by each mode for each of the models.

The qualitative results in figure 3 show that the shapes generated within the allowed range of **b** are all plausible. The quantitative results in table 1 show that our method produces models that are significantly more compact than either the models built by hand or those obtained using equally-spaced points.

To test the generalisation ability of the models, we performed leave-one-out tests on each model described in table 1. In figure 4, we report the results on the hand outlines although the same trends appear in all datasets. As can be seen from the figure, the optimised model performs significantly better than both the manual and arc-length parameterised models for the entire range of included modes, suggesting better generalisation ability.

5   Conclusions

We have described an efficient method for automatically defining a set of dense correspondences to build statistical shape models. Results show that the models
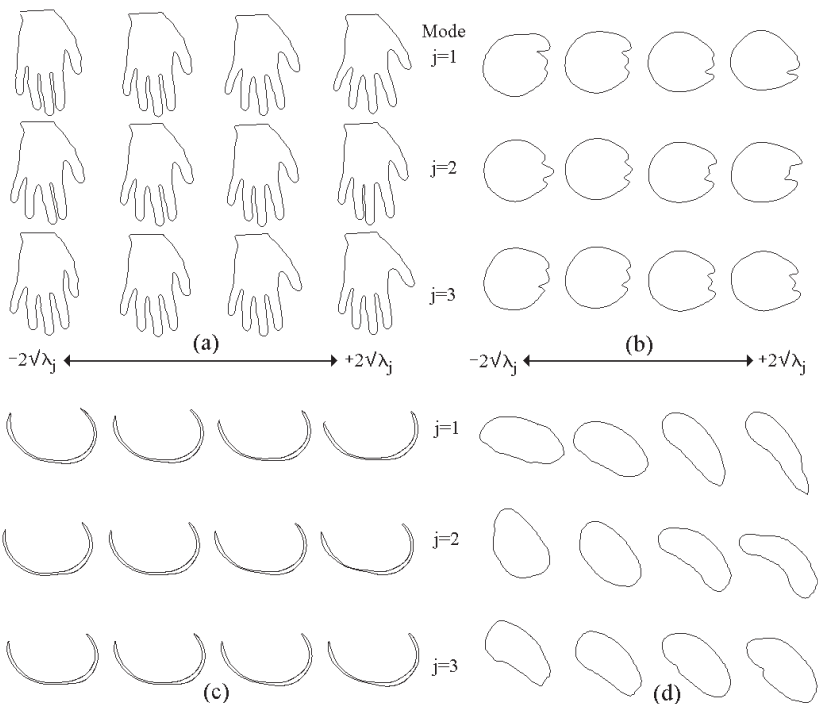
Fig. 3. The first three modes ($j = 1, j = 2, j = 3$) of variation ($\pm 2\sigma$)of the models automatically generated from (a)hand outlines, (b)kidneys, (c) knee cartilage and (d)Stroke Model.

produced by our method are more compact and general than models built by hand - the current gold standard.

Our method finds an isomorphic mapping between each of a set of shape contours and the mean. Since it does not use an explicit ordering constraint, it generalises to higher-dimensional spaces. By starting our optimisation from equally-spaced points, the worst model the method can possibly produce is a procrustes-aligned, arc-length parameterisation.

Regular Principal Component Analysis can not capture non-linear variations (e.g caused by a sub-part rotating in the plane) with a single mode - this affects the generalisation ability, specificity and compactness of such linear models. The method described in this paper overcomes this by allowing points to 'slide' along the parameterisation to compensate for the non-linear movement - this allows the variation to be explained by a single mode.

We are currently experimenting with more continuous representations of the parameterisation using the cumulative distribution defined by a set of basis functions. We intend to generalise the principle so that it can be used to reparameterise the sphere and hence surfaces rendered from 3D images.

Hands

Mode	Automatic	Hand Built	Equally-spaced
1	1.20	2.13	4.39
2	0.68	1.16	1.56
3	0.39	0.61	1.08
4	0.21	0.50	0.67
5	0.08	0.15	0.36
6	0.04	0.14	0.26
V_T	2.69	5.04	8.72
F_{min}	31477	41563	45272

Kidneys

Mode	Automatic	Equally Spaced
1	128.2	306.19
2	53.98	197.18
3	33.65	109.86
4	28.57	70.20
5	14.50	41.98
6	8.55	23.63
V_T	284.48	802.72
F_{min}	4722	5485

Knee Cartilage

Mode	Automatic	Hand Built	Equally-spaced
1	6.90	8.03	8.07
2	0.68	1.29	1.36
3	0.32	0.66	0.69
4	0.17	0.22	0.23
5	0.11	0.17	0.18
6	0.08	0.10	0.11
V_T	8.47	10.73	11.24
F_{min}	25133	35969	37941

Stroke Model

Mode	Automatic	Equally Spaced
1	684.29	1389
2	306.62	581.7
3	59.47	174.5
4	54.65	88.70
5	36.59	40.51
6	27.15	34.9
V_T	12385	23135
F_{min}	3412	4752

Table 1. A quantitative comparison of each model showing the variance explained by each mode. F_{min} is the value of the objective function and V_T is the total variance.

Acknowledgements

References

[1] Baumberg, A. and D. Hogg, Learning Flexible Models from Image Sequences, in *European Conference on Computer Vision, ECCV'94*. p. 299-308.

[2] Benayoun, A., N. Ayache, and I. Cohen. Adaptive meshes and nonrigid motion computation. in *International Conference on Pattern Recognition. ICPR'94*

[3] Bookstein, F.L., Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Medical Image Analysis*, 1997. 1(3): p. 225-243.

[4] Cootes, T., A. Hill, C. Taylor, and J. Haslam, The use of Active shape models for locating structures in medical images. *Image and Vision Computing*, 1994. 12: p. 355-366.

[5] Cootes, T., C. Taylor, D. Cooper and J. Graham, Active shape models - their training and application. *Computer Vision and Image Understanding*, 1995. 61: p. 38-59.

[6] Davies, Rh. H, T. Cootes and C. J. Taylor, A Minimum Description Length Approach to Statistical Shape Modelling, in *Information processing in medical imaging, IPMI'01*, (to appear)
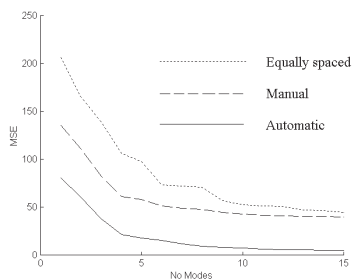
Fig. 4. Leave one out tests. The plot shows the number of modes used against the mean squared approximation error

[7] Goldberg, D.E., *Genetic Algorithms in Search, Optimisation and Machine Learning.* 1989: Addison Wesley.

[8] Goodall, C., Procrustes Methods in the Statistical Analysis of Shape. *Journal of the Royal Statistical Society B*, 1991. 53(2): p. 285-339.

[9] Hill, A. and C. Taylor. Automatic landmark generation for point distribution models. in *British Machine Vision Conference. BMVC'94.*

[10] Hill, A. and C.J. Taylor, A framework for automatic landmark identification using a new method of non-rigid correspondence. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI*, April, 2000.

[11] Kambhamettu, C. and D.B. Goldgof, Point Correspondence Recovery in Non-rigid Motion, in *CVPR'92*. p. 222-227.

[12] Kelemen, A., G. Szekely, and G. Gerig, Elastic model-based segmentation of 3-D neuroradiological data sets. *IEEE trans. medical imaging*, 1999. 18(10): p. 828-839.

[13] Kirkpatrick, S., C. Gelatt, and M. Vecchi, Optimization by Simulated Annealing. *Science*, 1983. 220: p. 671-680.

[14] Kotcheff, A.C.W. and C.J. Taylor, Automatic Construction of Eigenshape Models by Direct Optimisation. *Medical Image Analysis*, 1998. 2: p. 303-314.

[15] G. Paxinos and C. Watson. *The rat brain in streotactic coordinates.* Academic Press, San Diego 1986

[16] Rangarajan, A., H. Chui and F. L. Bookstein,The Softassign Procrustes Matching Algorithm, in *IPMI'97*. p. 29-42.

[17] Solloway, S., Taylor, C.J., Hutchinson, C.E., Waterton, J.C., Quantification of Articular Cartilage from MR images using active shape models, in *Proceedings of the 4th European Conference on Computer Vision, ECCV' 96*, pp. 400-412.

[18] Wang, Y., B. S. Peterson, and L. H. Staib. Shape-based 3D surface correspondence using geodesics and local geometry. *CVPR 2000*, v. 2: p. 644-51.

[19] Waterton J. C. , B. J. Middleton, R. Pickford, C. P. Allott, D. Checkley and R. A. Keith, Reduced animal use in efficacy testing in disease models with use of sequential experimental designs. Developments in Animal and Veterinary Sciences, 31: Progress in the Reduction, Refinement and Replacement of Animal Experimentation, 737-745 (2000)