(Image and Vision Computing, Vol.12, No.6, July 1994. pp.355−366)

The Use of Active Shape Models For Locating Structures in Medical Images

T.F.Cootes, A.Hill, C.J.Taylor and J.Haslam

Department of Medical Biophysics,
University of Manchester,
Manchester M13 9PT
England.

Abstract. This paper describes a technique for building compact models of the shape and appearance of flexible objects (such as organs) seen in 2−D images. The models are derived from the statistics of sets of labelled images of examples of the objects. Each model consists of a flexible shape template, describing how important points of the object can vary, and a statistical model of the expected grey levels in regions around each model point. The shape models are parameterised in such a way as to allow 'legal' configurations. Such models have proved useful in a wide variety of applications. We describe how the models can be used in local image search and give examples of their application to medical images. We also describe how the method can be simply extended to segment 3−D objects in volume images and to track structures in image sequences.

1 Introduction

Almost every object of interest in the human body can vary in size, shape and appearance. Often this makes the task of automatically identifying and segmenting interesting structures, such as organs or bones, very difficult. It can be made easier if suitable models of appearance of the desired structure are available. To be effective such models must be able to allow for the expected variations in size, shape and appearance of the structure in the image. There has been considerable interest in developing such flexible models or deformable templates [1−18]. However most previous approaches have either used 'hand−crafted' models designed specifically for one particular application or have been too general to give suitably constrained flexible models (see below). In this paper we present a statistically based technique for building compact models of the shape and appearance of almost any flexible object and show how they can be used when searching an image for a new example of the object. Though applicable in many domains these models are particularly useful for medical applications where variations in shape and appearance are otherwise difficult to model.

Our shape models rely on representing objects by sets of labelled points; each point is placed on a particular part of the object. By examining the statistics of the positions of the labelled points a 'Point Distribution Model' is derived. The model gives the average positions of the points, and a description of the main modes of variation found in the training set. The grey−level appearance of objects is represented by statistical models of the grey−levels in regions around each of the shape model points.

Locating an example of such a model in an image involves choosing values for each of the parameters so as to best fit the model to the image. An initial guess for the best shape, orientation, scale and position can be refined by comparing the hypothesised model example with image data and using differences between model and image to deform the shape. The method has similarities with the Active Contour Models (or snakes) of Kass *et al* [6], but differs in that global shape constraints are applied; to make this distinction clear we have adopted the term Active *Shape* Models. The key point is that an instance of a model can only deform in ways found in its training set. The methods we present can be thought of as a two dimensional application of Lowe's refinement technique [7]. Because of the linear nature of the Point Distribution Model, the mathematics is considerably simpler and can lead to rapid execution.

As well as being able to capture the variability in shape from different examples, a 2−D PDM can capture the apparent changes in shape due to *slice* position through 3−D objects or the temporal variation of 2−D objects. They can thus be used to segment 3−D volumes or to track objects in image sequences.

In this paper we describe the modelling method, present examples of models and show examples of image segmentation using Active Shape Models. We also demonstrate a model locating ventricles in a 3D MR image of the brain and tracking the left ventricle of the heart in an echocardiogram sequence.

2 Background

A number of authors have used flexible models or deformable templates to aid image interpretation. Such models usually have a number of parameters to control the shape and pose of all or part of the model. Yuille *et al* [8,9] and Lipson *et al* [10] use 'hand−crafted' models of faces and transaxial slices through vertebrae. These have to be individually tailored for each application. Kass *et al* [6] describe flexible contour models which are attracted to image features. They are usually free to take almost any smooth boundary with few constraints on their overall shape. The method of fitting, that of using image evidence to apply forces and then minimising an energy function, is an effective one. Hinton, Williams and Revow [11] describe a type of spline snake governed by a number of control points which have preferred 'home' locations to give the snake a particular default shape. Deformations are caused by moving the control points away from their 'home' locations. Although the average shape of an object is represented, the modes of shape variation are only coarsely defined by the number and position of control points.

Scott [12], Staib and Duncan [13] and Bozma and Duncan [14] all use closed contour shape models, based on expansions of trigonometric functions, for the interpretation of medical images. However, the basis functions used are unlikely to give the most compact representation of shape and shape variability. Although recording the distributions of each parameter over a set of examples leads to a broad description of a class of shapes, without knowledge of how the parameters tend to vary together over the training set the models lack specificity; ie many examples can be generated which are not 'legal' generalisations of the class of shapes.

Pentland and Sclaroff [15], Karaolani *et al* [16] work with finite element models of flexible objects. Like the trigonometric models the basis functions used in finite element models are not necessarily the most effective at describing the variability occurring in a particular class of shapes, and so may not give the most specific or compact model.

Grenander, Chow and Keenan [17], Mardia, Kent and Walder [18] describe statistical models of shape. These are local models and the methods used to fit to images impractical for real applications. Bookstein [19,20] has applied statistical techniques to learn relationships between shape and other variables for morphometric analysis, but does not generate models which could be used to aid image interpretation.

3 Modelling Object Shape : Point Distribution Models

Suppose we have a set of images containing examples of a variable structure which we wish to model, for instance the left ventricle of the heart as seen in an echocardiogram (the chamber to the upper right in Figure 1). The shape of this chamber can vary both with time, as the heart beats, and between individuals. We wish to model this shape variation and use the resulting models to interpret images.

3.1 Obtaining the Training Data

In order to model the shape, we represent it by a set of points. For the heart example we have chosen points around the ventricle boundary, and also around the nearby edge of the right ventricle and the top of the left atrium. Although we are primarily interseted in the left ventricle, the addition of these components gives a more specific model which is better able to locate the left ventricle during image search. The chosen components must be labelled on each shape in a training set. The labelling of the points is important. Each point represents a particular part of the object or its boundary, so must be planted in the same way on every training example.

Figure 2 shows some of a set of example shapes, each represented by 96 points. Theses were obtained by drawing boundaries on each of 66 images. 11 key positions were marked on each boundary, and the 96 points generated by equalling spacing points along the boundaries between the key positions. In Figure 2 the points have been connected by lines in a manner defined by a separate model descriptor file.

3.2 Aligning The Set of Training Shapes

The modelling method works by examining the statistics of the co−ordinates of the labelled points over the training set. In order to be able to compare equivalent points from different shapes, they

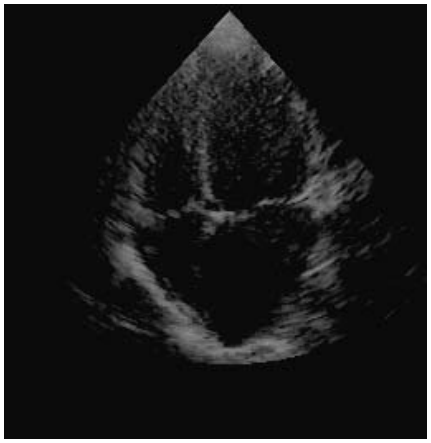


Figure 1. Example of an echocardiogram. The left ventricle is at the top right.

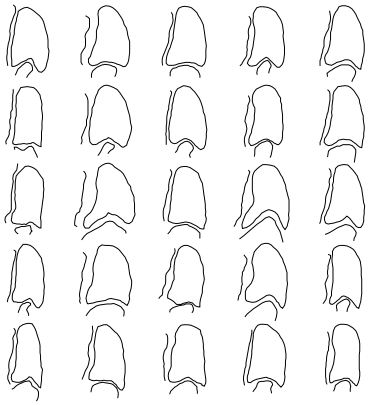


Figure 2. Examples of heart left ventricle shapes, each containing 96 points.

must be aligned in the same way with respect to a set of axes. We achieve the required alignment by scaling, rotating and translating the training shapes so that they correspond as closely as possible. We have previously [1] described an iterative method which minimises a weighted sum of squares of distances between equivalent points on different shapes.

3.3 Capturing the Statistics of a Set of Aligned Shapes

Once a set of N aligned shapes is available the mean shape and variability can be found as follows:

Let $\mathbf{x}_i$ be a vector describing the n points of the i^{th} shape in the set;

$$\boldsymbol{x}_i = (x_{i0},\ y_{i0},\ x_{i1},\ y_{i1},\ \ldots,\ x_{ik},\ y_{ik},\ \ldots,\ x_{in-1},\ y_{in-1})^T$$

where (x_{ij}, y_{ij}) is the j^{th} point of the i^{th} shape.

The mean shape, $\bar{\boldsymbol{x}}$, is calculated using

$$\bar{\boldsymbol{x}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_i \tag{1}$$

The modes of variation, the ways in which the points of the shape tend to move together, can be found by applying principal component analysis to the deviations from the mean as follows [22]. For each shape in the training set we calculate its deviation from the mean, $d\mathbf{x}_i$, where

$$d\boldsymbol{x}_i = \boldsymbol{x}_i - \bar{\boldsymbol{x}} \tag{2}$$

We can then calculate the $2n$ x $2n$ covariance matrix, $\mathbf{S}$, using

$$\boldsymbol{S} = \frac{1}{N}\sum_{i=1}^{N} d\boldsymbol{x}_i d\boldsymbol{x}_i^T \tag{3}$$

The modes of variation of the points of the shape are described by $\boldsymbol{p}_k$ $(k = 1..2n)$, the unit eigenvectors of $\mathbf{S}$ such that

$$\boldsymbol{S}\boldsymbol{p}_k = \lambda_k \boldsymbol{p}_k \tag{4}$$

(where λ_k is the k'th eigenvalue of $\mathbf{S}$, $\lambda_k \geq \lambda_{k+1}$)

$$\boldsymbol{p}_k^T \boldsymbol{p}_k = 1 \tag{5}$$

It can be shown that the eigenvectors of the covariance matrix corresponding to the largest eigenvalues describe the most significant modes of variation in the variables used to derive the covariance matrix, and that the proportion of the total variance explained by each eigenvector is equal to the corresponding eigenvalue [22]. Most of the variation can usually be explained by a small number of modes, t $(<2n)$. One method for calculating t is to choose the smallest number of modes such that the sum of their variances explain a sufficiently large proportion of λ_T, the total variance of all the variables, where

$$\lambda_T = \sum_{k=1}^{2n}\lambda_k \tag{6}$$

The k'th eigenvector affects point l in the model by moving it along a vector parallel to (dx_{kl}, dy_{kl}), which is obtained from the l'th pair of elements in $\mathbf{p}_k$;

$$(dx_{k0},\ dy_{k0},\ \ldots\ ,dx_{kl},\ dy_{kl},\ \ldots\ ,\ dx_{kn-1},\ dy_{kn-1}\ ) \tag{7}$$

Combinations of vectors, one for each mode, move the modelled landmark points around in the regions of the 'clouds' of scattered points from the aligned training set. Any shape in the training set can be approximated using the mean shape and a weighted sum of these deviations obtained from the first t modes

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \boldsymbol{P}\boldsymbol{b} \tag{8}$$

where $\boldsymbol{P} = (\boldsymbol{p}_1\ \boldsymbol{p}_2\ \ldots\ \boldsymbol{p}_t)$ is the matrix of the first t eigenvectors,

$\boldsymbol{b} = (b_1\ b_2\ \ldots\ b_t)^T$ is a vector of weights, one for each eigenvector

the eigenvectors are orthogonal, $\boldsymbol{P}^T\boldsymbol{P} = \boldsymbol{I}$ so

$$\boldsymbol{b} = \boldsymbol{P}^T(\boldsymbol{x} - \bar{\boldsymbol{x}}) \tag{9}$$

The above equations allow us to generate new examples of the shapes by varying the parameters (**b**) within suitable limits, so the new shapes will be similar to those in the training set. The parameters are linearly independent, though there may be non−linear dependencies still present. The limits for each element of **b**, b_k, are derived by examining the distributions of the parameter values required to generate the training set. If gaussian distributions are assumed one can choose sets of parameters $\{ b_1 \ldots b_t \}$ such that the Mahalanobis distance (D_m) from the mean is less than a suitable value, D_{max};

$$D_m^2 = \sum_{k=1}^{t}\left(\frac{b_k^2}{\lambda_k}\right) \le D_{\max}^2 \tag{10}$$

3.4 An Example of a Shape Model

A heart model was trained on a set of 66 examples (Figure 2) each comprising 96 points. This illustrates how one model can represent both shapes and relationships between several objects. The points represent the boundary of the left ventricle, part of the boundary of the right ventricle and part of the boundary of the left atrium (below the ventricle in the examples). Figure 3 shows shapes reconstructed by varying each of the first four model parameters in turn, keeping the others zero. The first mode varies the width of the model, the second the position of the lower part of the ventricle. The third and fourth modes give other deformations of the chamber boundary.

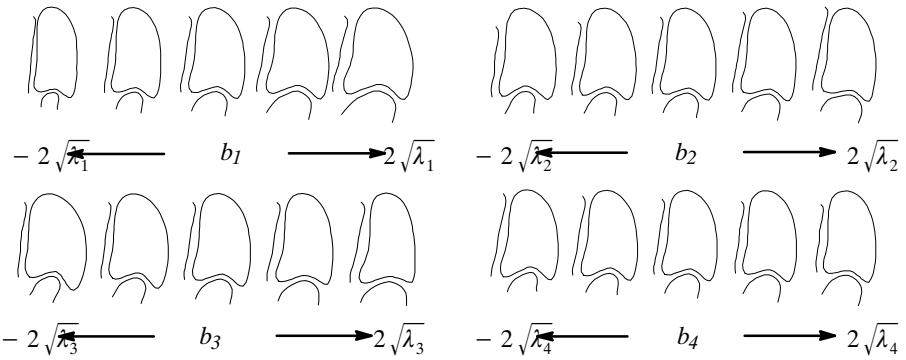


Figure 3. Effects of varying each of the first four parameters of the heart ventricle model individually.

The modelling technique has been applied successfully to a wide variety of examples including the ventricles of the brain in MR images [5], spinal vertebrae in X−rays [21], the outline of the abdomen and the prostate in MR images (see below), faces [3], hands [1] chromosomes and industrial components.

4 Modelling Grey Level Appearance

We wish to use our models for locating examples of objects in new images. For this purpose, not only shape, but also grey−level appearance is important. We account for this by examining the statistics of the grey levels in regions around each of the labelled model points. Since a given point corresponds

to a particular part of the object, the grey−level patterns about that point in images of different examples will often be similar. The work of Bailes and Taylor [23] suggests that the location of model points in images can be improved by incorporating each point's grey−level environment into the model.

Although in general we could consider a region of any shape around each point , we will concentrate on one−dimensional profiles normal to curves passing through the points. This requires that we define the connectivity of the model points; In many cases this is straightforward, particularly when the points lie around a boundary.

In some cases it is efficient to assume that the points lie on strong edges and to seach for such in an image. However this is not always satisfactory, and it is necessary to have a more general model of the grey−level appearance. For every model point i in each image, j, we can extract a profile, $\boldsymbol{g}_{ij}$, of length n_p pixels, centred at the point. Following Bailes and Taylor [23], we choose to sample the derivative of the grey levels along the profile in the image and normalise. This gives invariance to uniform scaling of the grey levels and the addition of a constant.

If the profile runs from $\boldsymbol{p}_{istart}$ to $\boldsymbol{p}_{iend}$ and is of length n_p pixels, the k^{th} element of the derivative profile is

$$g_{ijk} = I_j(\mathbf{y}_{i(k+1)}) - I_j(\mathbf{y}_{i(k-1)}) \tag{11}$$

where $\mathbf{y}_{ik}$ is the k^{th} point along the i^{th} profile :

$$\boldsymbol{y}_{ik} = \boldsymbol{p}_{istart} + \frac{k-1}{n_p - 1}(\boldsymbol{p}_{iend} - \boldsymbol{p}_{istart}) \tag{12}$$

and $I_j(\boldsymbol{y}_{ik})$ is the grey level in image j at that point.

We then normalise this profile,

$$\boldsymbol{g}_{ij}{}' = \frac{\boldsymbol{g}_{ij}}{\sum_{k=1}^{n_p} |g_{ijk}{}'|} \tag{13}$$

For each point, i, we can calculate a mean normalised derivative profile,

$$\bar{\boldsymbol{g}}_i = \frac{1}{N_s}\sum_{j=1}^{N_s} \boldsymbol{g}_{ij}{}' \tag{14}$$

We can then calculate an n_p x n_p covariance matrix, $\mathbf{S}_{\mathbf{gi}}$, giving us a statistical description of the expected profiles about the point.

5 Using PDMs in Image Search − Active Shape Models

Having generated a flexible shape model and a description of the grey levels about each model point we would like to find examples of the modelled structures when they are present in images. In general, a procedure for achieving this has two stages:

- A number of hypotheses are made, giving approximate locations of the model points
- Each of the hypotheses is refined and the best chosen.

The initial hypotheses take the form of estimates for the position at which the model should be placed, its orientation, scale and the shape parameters required to fit the model to the image. We assume in what follows that we know roughly the position in which the model should be placed and that our initial hypothesis assumes the mean shape and scale. Hypotheses can be obtained from cue generators or by using suitable search techniques. Hill *et al* [4,5] describe methods for finding flexible objects in images which use Genetic Algorithms to generate a set of hypotheses quite rapidly.

Genetic Algorithms (GAs) employ mechanisms analagous to those involved in natural selection to conduct a search through a given parameter space for the optimum of some *objective function*, in this case the fit of the model to the image. The main features of the GA approach are that a population of different solutions is maintained, and that new solutions are generated by probabalistically recombining parts of existing solutions. Optimal solutions are thus *evolved* by iteratively producing new *generations* of solutions and applying a *selective breeding* strategy which favours more successful solutions (good fits of model to data). Each solution in the population consists of the pose and shape para-

meters of a model instance. Because a GA maintains a population of solutions it can compare many possible alternative image interpretations and select a number of plausible hypotheses.

By choosing a set of shape parameters **b** for a Point Distribution Model, we define the shape of a model instance **x** in a model centred co−ordinate frame. We can then create an instance, **X,** of the model in the image frame by defining the position, orientation and scale:

$$\boldsymbol{X} = M(s,\theta)[\boldsymbol{x}] + \boldsymbol{X}_c \qquad (15)$$

where $\boldsymbol{X}_c = (X_c,\ Y_c,\ X_c,\ Y_c,\ ...,\ X_c,\ Y_c)^T$

$M(s,\theta)[\ ]$ performs a rotation byθ and a scaling by s

(X_c, Y_c) is the position of the centre of the model in the image frame.

In this section we describe an iterative method for refining the shape and pose parameters so as to give a better match between a model instance and structures in the image. The approach is as follows :

i) Examine a region of the image around each point to calculate the displacement of the point required to move it to a better location.
ii) From these displacements calculate adjustments to the pose and scale and to the shape parameters of the Point Distribution Model.
iii) Update the model parameters; by enforcing limits on the shape parameters, global shape constraints can be applied ensuring the shape of the model instance remains similar to those of the training set.

The procedure is repeated until no significant changes result. Because the models deform to better fit the data, but only in ways which are consistent with the shapes found in the training set we call them 'Active Shape Models' (ASMs).

5.1 Calculating A Suggested Movement For Each Model Point

Given an initial estimate of the positions of a set of model points which we are attempting to fit to an image we need to estimate a set of adjustments which will move each point toward a better position. In the case in which the model points represent the boundary of an object (Figure 4) the required adjustments will move them toward the edges of the image object. If we have a profile model for each point, the search involves finding a nearby region which better matches the profile model (Figure 5). At a particular model point we extract a derivative profile, **g**, from the current image of some length, l ($>n_p$), centred at the point and aligned parallel to the orientation we have defined at that point (for instance, normal to the boundary). We then run the profile model along this sampled profile and find the point at which the model best matches. If we assume that the points lie on edges, it is simply a matter of finding the strongest edge along the profile. If we have a more general model of appearance we must calculate how well it fits to sub−sections of the sampled profile.

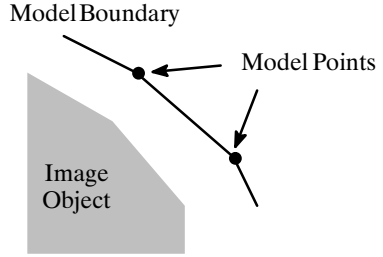


Figure 4. Part of a model boundary approximating to the edge of an image object.

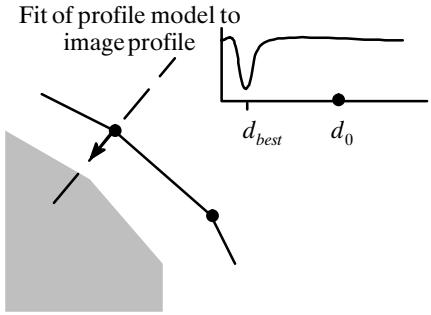


Figure 5. Suggested movement of point is along normal to boundary, in a direction towards the point at which the profile model best fits the profile sampled from the image.

Given a sampled derivative profile the fit of the model at a point d pixels along it is calculated as follows;

$$f_{prof}(d) = (\boldsymbol{h}(d) - \bar{\boldsymbol{g}})^T\ \boldsymbol{S}_g^{-1}(\boldsymbol{h}(d) - \bar{\boldsymbol{g}}) \qquad (16)$$

where $\mathbf{h}(d)$ is a sub−interval of **g** of length n_p pixels centred at d, normalised using (13). This is the square of the Mahalanobis distance of the sample from the mean grey model and for normally distrib-

uted data is proportional to the log of the probability of obtaining $\mathbf{h}(d)$ from the measured distribution of grey−levels.

The value of f_{prof} decreases as the fit improves. The point of best fit is thus the point at which $f_{prof}(d)$ is a minimum.

Suppose d_{best} is the distance along the sampled profile from the model point to the point of best fit. We choose a displacement for the model point of $d\mathbf{X}$ which is parallel to the profile, in the direction of the point of best fit, with magnitude

$$\begin{array}{lll} |d\boldsymbol{X}| = 0 & \text{if } |d_{best}| \le \delta \\ |d\boldsymbol{X}| = 0.5d_{best} & \text{if } \delta < |d_{best}| < d_{\max} \\ |d\boldsymbol{X}| = 0.5d_{\max} & \text{if } d_{\max} \le |d_{best}| \end{array} \qquad (17)$$

We have used $\delta = 0.5$ pixels, $d_{max} = 8$ pixels in our experiments.

An alternative approach is to generate potential images such as to those described by Kass *et al* [6], possibly one for each model point, describing how likely each point in the image is to be the model point [24]. Adjustments to the position of each point can then be derived from the gradient of the potential image at the current estimate of the point's position.

5.2 Calculating the Adjustments to the Pose and Shape Parameters

Using the procedure described above a set of adjustments can be calculated, one for each point of the shape. We denote such a set as a vector $d\mathbf{X}$, where

$$d\boldsymbol{X} = (dX_0,\ dY_0,\ \ldots,\ dX_{n-1},\ dY_{n-1})^T$$

We aim to adjust the pose and shape parameters to move the points from their current locations in the image frame, $\mathbf{X}$, to be as close to the suggested new locations $(\mathbf{X} + d\mathbf{X})$ as can be arranged whilst still satisfying the shape constraints of the model. If the current estimate of the model is centred at (X_c, Y_c) with orientation θ and scale s we would like first to calculate how to update these parameters to better fit the image. This is achieved by finding the translation (dX_c, dY_c), rotation $d\theta$ and scaling factor $(1+ds)$ which best maps the current set of points, $\mathbf{X}$, onto the set of points given by $(\mathbf{X} + d\mathbf{X})$. This can be done by a weighted least squares fit [1]. The choice of possible weights is discussed below.

Having adjusted the pose variables there remain residual adjustments which can only be achieved by deforming the shape of the model. We wish to calculate the adjustments, $d\mathbf{x}$, to the original model points in the local co−ordinate frame required to cause the scaled, rotated and translated points $\mathbf{X}$ to move by $d\mathbf{X}$ when combined with the new scale, rotation and translation parameters.

In [2] we show that

$$d\boldsymbol{x} = M((s(1+ds))^{-1}, -(\theta + d\theta))[M(s,\theta)[\boldsymbol{x}] + d\boldsymbol{X} - d\boldsymbol{X}_c] \qquad (18)$$

This gives a way of calculating $d\boldsymbol{x}$, the suggested movements to the points in the local model co−ordinate frame. Since there are only $t\ (<2n)$ modes of variation available and $d\mathbf{x}$ can move the points in $2n$ different degrees of freedom, we can only achieve an approximation to the deformation required. The movements are not in general consistent with our shape model. We wish to calculate the adjustments to the shape parameters, $d\mathbf{b}$, which will best match the model to the suggested new positions. This can be thought of as minimising a (possibly weighted) sum of squares of differences between model points and desired points. We wish to minimise

$$(d\boldsymbol{x}')^T \boldsymbol{W}_s (d\boldsymbol{x}') \qquad (19)$$

where $\boldsymbol{W}_s$ is a diagonal matrix of weights, one for each co−ordinate of each point.

$$\begin{aligned} d\boldsymbol{x}' &= ((\boldsymbol{x} + d\boldsymbol{x}) - (\bar{\boldsymbol{x}} + \boldsymbol{P}(\boldsymbol{b} + d\boldsymbol{b})) \\ &= (d\boldsymbol{x} - \boldsymbol{P}d\boldsymbol{b}) \end{aligned}$$

It can be shown that (19) is minimised when

$$(\boldsymbol{P}^T\boldsymbol{W}_s)d\boldsymbol{x} = (\boldsymbol{P}^T\boldsymbol{W}_s\boldsymbol{P})d\boldsymbol{b} \qquad (20)$$

This is a set of t linear equations in the t variables of $d\mathbf{b}$, and can be solved using standard matrix algebra. In the special case in which all weights are set to unity, $\mathbf{W}_s = \mathbf{I}$, Eq.(20) simplifies to

$$d\boldsymbol{b} = \boldsymbol{P}^T d\boldsymbol{x} \qquad (21)$$

5.3 Choice of Weights For Pose and Shape Parameter Adjustment Calculation

In the calculation of the adjustments to both the pose and shape parameters we can give a weight to each point to indicate the confidence we have in the suggested new position for the point. If all weights are set to unity simplifications can be made which reduce the complexity of the calculations required (Equation 21).

We have found that decreasing the weight for points which are found to be further away from the current model points than average, and may be outliers, works well:

$$w_i = \frac{1}{2 + |d\mathbf{X}_i|^2} \tag{22}$$

5.4 Updating the Pose and Shape Parameters

The equations above allow us to calculate changes to the pose variables, dX_c , dY_c, $d\theta$ and ds, and adjustments to the shape parameters $d\mathbf{b}$ required to improve the match between an object model and image evidence. When these changes are applied we can ensure that the model only deforms into shapes consistent with the training set by placing limits on the values of b_k. As mentioned above, a shape can be considered acceptable if the Mahalanobis distance D_m is less than a suitable constant, D_{max}, for instance 3.0 (See Eq. 10). That is to say, the vector $\mathbf{b}$ should lie within a hyper−ellipsoid about the origin. If updating $\mathbf{b}$ leads to an implausible shape, ie $D_m > D_{\max}$ and the point lies outside the ellipsoid, $\mathbf{b}$ can be re−scaled to lie on the closest point of the allowed volume using

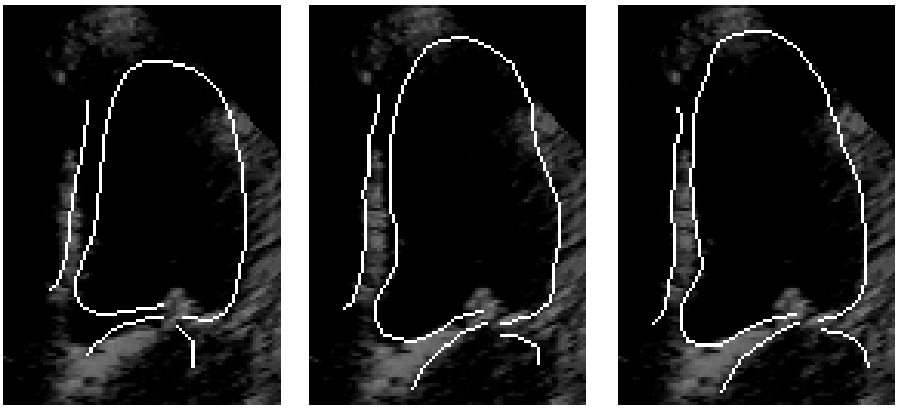
$$b_k \rightarrow b_k \cdot \frac{D_{\max}}{D_m} \quad (k = 1..t) \tag{23}$$

Note that we have already applied implicit limits of zero to the weights of the eigenvectors truncated from our representation (ie $b_i = 0 \ \forall \ i > t$). Once the parameters have been updated, and limits applied where necessary, a new example can be calculated, and new suggested movements derived for each point. The procedure is repeated until no significant change results.

6 Examples of ASMs in Action

6.1 Echocardiograms

Figure 6a shows the initial placement of an instance of the 96 point heart model described above on part of an echocardiogram not used in the training set (Figure 1 shows a complete image). The initial placement was obtained by choosing a set of pose parameters and setting all the shape parameters to zero (corresponding to the mean model shape). In practice such an initial estimate can be obtained from cues or a global search (Hill *et al* [5]). Figure 6b shows the model instance after 20 iterations. After 50 iterations (Figure 6c) the model appears to give a good fit to the data. The shape model used has 12 degrees of freedom. Each iteration takes about 0.05s on a Sun Sparc workstation.



(a) Initial State (b) After 20 iterations (c) After 50 iterations

Figure 6. Echocardiogram image with heart chamber boundary model superimposed, showing its initial position and its location after 20 and 50 iterations.

The grey−levels around each point are modelled as step edges in the correct direction. During image search the strongest edge along the profile through each point is taken to be the desired new position for the point. In this example the model is able to infer the position of the parts of the boundary where

the data is poor (for example, near the top of the ventricle) by using the knowledge of the expected shape combined with information from the areas of the image where there is good evidence for the ventricle wall. Without the prior knowledge of the shape given by the model it would not be possible to delineate the ventricle boundary.

6.2 Locating the prostate.

Figure 7 shows an MR slice through the lower abdomen overlayed with the mean shape of a model built to represent parts of such an image. The model uses 121 points to represent the boundary of the abdomen, a section through the prostate (in the centre), the femur heads (on the left and right), two femoral arteries (upper left and right) and the symphasis pubis (a T−shaped structure useful as a cue to the position of the prostate). The model was trained on 32 examples, taking two slices and their reflections (through the line of symmetry) from each of 8 patients. The reflections were used to increase the effective size of the training set, making use of the assumption that the model would be symmetrical. The slices were taken at approximately the same position in each patient. As a simple edge model would be inappropriate (some model points lie on weak edges, and some not on edges at all), the grey level environment about the model points was represented using grey−levels profile models 7 pixels long, centred on each point. Figure 7 shows an initial estimate of the pose required to fit the mean model to an image not used in the training set. Figure 8 shows how the ASM has fitted to the image after 100 iterations. Each iteration took 0.15 seconds on a Sun Sparc workstation. The prostate is difficult to locate on its own. By modelling a number of nearby features in the slice it is much easier to find its boundary.

Figure 7. Example of an MR slice through the abdomen showing the initial placing of the mean model.

Figure 8. Model after 100 iterations

6.3 Other Examples

We have applied the techniques described above to a number of different problems in both medical and industrial fields. For instance segmenting vertebrae in lateral X−rays of the spine [21] and locating facial features [3]. In each case the same procedures were used to build the models and search images.

6.4 Active Shape Models Applied to 3−D and Time Sequence Images

We have found that a simple 2−D PDM of the first and second ventricles of the brain can capture not only biological variation between individuals but also changes in shape associated with slice position and imaging angle (tilt). It is reasonable to assume, therefore, that this 2−D model might be used to segment brain ventricles from 3−D volume images. This might be achieved by applying a global search (such as a Genetic Algorithm −GA) to each slice independently. However, because the changes in ventricular shape between slices are small, we have adopted the following approach :

1 The GA is applied to a single slice and the ventricles located approximately.
2 The approximate solution is refined using the ASM local optimiser.
3 The segmentation generated by the ASM is projected into the next slice of the 3−D volume to provide an initial approximation of ventricular position in the next slice. The projection is achieved by measuring the differences in the pose and shape parameters between previous slices and using a first order prediction of the parameters for the next slice.
4 Steps 2 and 3 are repeated for each slice.

This proves to be a rapid and accurate mechanism for segmenting the ventricles in 3−D MR datasets. Figure 9 shows the result of applying the approach to an unseen 3−D MR image of the brain.

We have also applied the technique to temporal image sequences of echocardiograms in order to track the left ventricle of a beating heart (Figure 10). The model (see Section 3.4) captured the change in

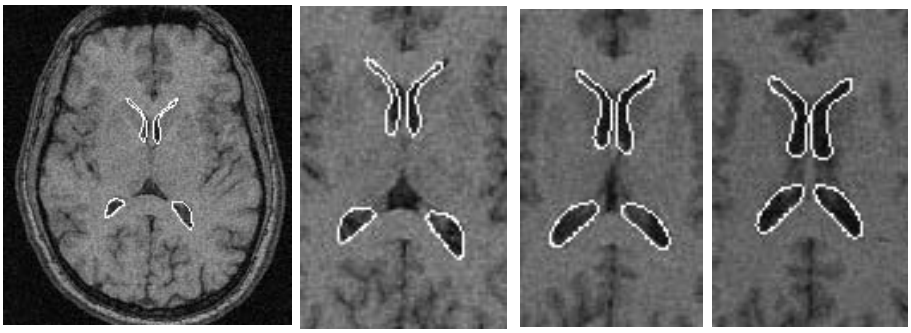


Figure 9. Tracking the First and Second Ventricles of the Brain through slices of a 3D MR Dataset (selected from 15 slices).

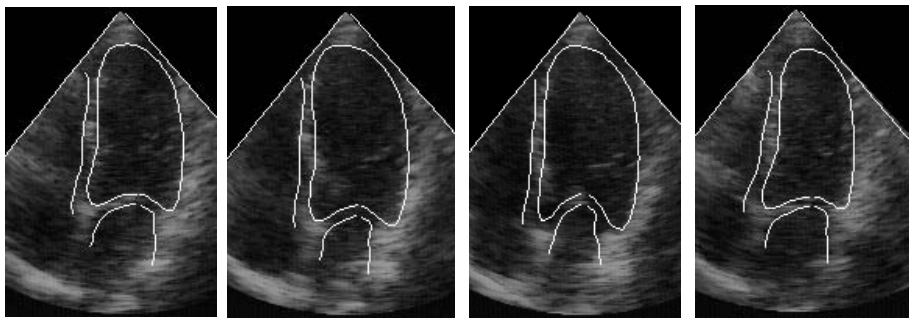


Figure 10. Tracking the left ventricle of a heart in an echocardiogram sequence.

shape resulting from both the biological variation between individuals and changes due to the motion of the heart through time.

7 Discussion

7.1 The Shape Model

When building shape models it is important that the points are placed on the training images as accurately as possible, and the shapes aligned similarly. If not the model will be unable to represent correctly the position of each point − it will include terms describing the noise caused by errors in point location and shape orientation.

The Point Distribution Model can be extended to deal with volume data (eg 3D medical images). For this labelled sets of $3-D$ points are required, each of which must be placed in a particular location on each training example. The problems of devising ways of choosing suitable model points and placing them consistently on sets of examples are the subject of continuing research.

The PDM can be used in a classifier. Given an example of a shape, an estimate can be made of how likely that example is to be a member of the class of shapes described by a model. If labelled points are placed on the example and the point set aligned with the mean shape, Eq. 9 can be used to calculate the model parameters required to generate the example. The distributions of the parameters can be estimated from the training set, allowing probabilities to be assigned. This technique has been successfully used in a simple handwritten character recognition application [24].

7.2 The Active Shape Model

The iterative approach described above, using image evidence to deform a Point Distribution Model, is effective at locating objects, given an initial estimate of their position, scale and orientation. How good an initial estimate is required will depend on how cluttered the image is and how well the model describes the object in the image.

By allowing the model to deform, but only in ways seen in the set of examples used as a training set, we have a powerful technique for refinement. The constraints on the shape of the model are applied by the limits on the shape parameters.

It can be shown that including weights in the calculation of pose and shape parameter adjustments can slow convergence but leads to better overall results, partly because spurious profile model matches away from the majority of the model points can be discriminated against [3]. More detailed models of the grey−level environment about each point give more specific overall models of image

structures and ASMs using such models are less likely to be confused by clutter and noise during image search.

7.3 Comparison with Other Work

Our Active Shape Models have similarities with the snakes of Kass *et al* [6] in that they move around in potential fields, attempting to minimise some function subject to certain constraints on their form. Snakes are usually free to take up a wide variety of shapes; when it is necessary to ensure they retain a certain shape, this is usually done by adding terms to the objective function favouring the desired shape [8]. Active Shape Models explicitly model the shape of an object and the allowed variations giving a more powerful description. Each model can be generated from a set of examples − the only requirement is a labelled set of examples upon which to train the model.

7.4 Tracking Structures in Image Sequences

We have demonstrated how Active Shape Models can track structures through sequences of images, either slices of 3−D volume data or time sequences from videos. In each case we took as our initial estimate of the shape in a frame a prediction based on the shape in the preceding frames. Since the ASM is a local optimiser, if fails to find the best location in one frame, this error will be propogated to the following frames which could lead to a poor result. A more robust scheme could involve more global searching of each frame to ensure the best fit is achieved. Clearly for the 3−D volume data a complete 3−D PDM would be more appropriate. The problems of building and using such a model are currently being addressed and the solutions will be presented in the near future.

7.5 Combining ASMs with Genetic Algorithms

Being a local search technique, the ASM needs a reasonable start position. The global search of a GA can be used to find good but approximate solutions rapidly. It is thus attractive to consider combining the two approaches. There are two ways this might be achieved :

a) Consider the techniques as separate but complementary. As in 6.4 above the GA search is conducted and the ASM applied as a refinement procedure to the suggested GA solutions. If the GA suggests a solution in a non−optimal area of the search space the ASM can do no better than to locate the local optimum in that area.
b) Incorporate the ASM directly into the GA search; in the GA literature it has been suggested that incorporating heuristic information and local optimisation techniques within a GA search can improve performance significantly [25]. The basis of this approach is that the GA can locate *hills* in the search space whilst the ASM embedded within the GA can climb to the top of the hills.

Each of the population of solutions for a GA is assessed by evaluating an *objective function* which measures the evidence for a given hypothesis. Our experiments with the latter show that applying a single iteration of the ASM procedure each time the objective function is evaluated requires minimal extra work and leads to significant improvements in the rate of convergence of the GA and the quality of the final fit.

8 Conclusions

The methods we describe allow flexible models of image objects such as organs to be built easily from sets of example images. The techniques can be applied to a wide variety of objects in different imaging modalities. The local search technique we presented can be used to locate new examples of the modelled objects in new images.

We have shown that the local optimisation method described can be fruitfully used in conjunction with a Genetic Algorithm (GA) search [4,5]. The GA can be run as a cue generator to produce a number of object hypotheses, which can be refined using the Active Shape Model. Alternatively the ASM can be embedded within the GA search directly [5]. Both techniques give good results.

Acknowledgements

Tim Cootes is funded by the Science and Engineering Research Council under the Information Engineering Advanced Technology Programme (VISAGE − Project Number 3/2114).

Andrew Hill is seconded from CN Software Ltd, funded under the Information Engineering Advanced Technology Programme (MIRIAD − Project Number 1660).

Jane Haslam is funded by the SERC.

The authors would like to thank the other members of the Wolfson Image Analysis Unit for their help and advice, particularly David Cooper, Jim Graham, David Bailes and Jeff Hunter.

9 References

1. T.F.Cootes, C.J.Taylor, D.H.Cooper and J.Graham: Training Models of Shape from Sets of Examples. In: Proc. British Machine Vision Conference. Springer−Verlag, 1992, pp.9−18.
2. T.F.Cootes, C.J.Taylor,: Active Shape Models − 'Smart Snakes'. In: Proc. British Machine Vision Conference. Springer−Verlag, 1992, pp.266−275.
3. T.F.Cootes, C.J.Taylor, A.Lanitis, D.H.Cooper and J.Graham: Building and Using Flexible Models Incorporating Grey−Level Information. Proc International Conference on Computer Vision, Berlin 1993.
4. A. Hill,C.J. Taylor and T.Cootes: Object Recognition by Flexible Template Matching using Genetic Algorithms. In: Proc. European Conference on Computer Vision (G.Sandini. Ed.). pp. 852−856, Springer−Verlag, 1992.
5. A. Hill, T.F. Cootes and C.J. Taylor: A Generic System for Image Interpretation Using Flexible Templates. In: Proc. British Machine Vision Conference. Springer−Verlag, 1992, 276−285.
6. M. Kass, A. Witkin and D. Terzopoulos : Snakes: Active Contour Models. In: Proc. First International Conference on Computer Vision, pp 259−268 IEEE Computer Society Press, 1987.
7. D.G. Lowe: Fitting Parameterized Three−Dimensional Models to Images. IEEE Trans. on Pattern Analysis and Machine Intelligence, **13**, 1991, 441−450.
8. A.L. Yuille, D.S. Cohen and P. Hallinan: Feature extraction from faces using deformable templates. Proc. Comp. Vision Patt. Rec., 1989, pp. 104−109.
9. A.L. Yuille, P. Hallinan and D.S. Cohen: Feature extraction from faces using deformable templates, IJCV, **8**, August 1992, pp. 99−112.
10. P. Lipson, A.L. Yuille, D. O'Keeffe, J. Cavanaugh, J. Taaffe and D. Rosenthal: Deformable Templates for Feature Extraction from Medical Images. In: Proceedings of the First European Conference on Computer Vision (Lecture Notes in Computer Science), (O. Faugeras. Ed.) pp. 413−417, Springer−Verlag, 1990.
11. G.E.Hinton, C.K.I. Williams and M.D. Revow: Adaptive Elastic Models for Hand−Printed Character Recognition. in Advances in Neural Information Processing Systems 4, (J.E.Moody, S.J.Hanson, R.P.Lippmann. Ed.s) Morgan Kauffmann, San Mateo, CA, 1992.
12. G.L. Scott: The Alternative Snake − And Other Animals. Proceedings 3rd Alvey Vision Conference (Cambridge), 1987, pp. 341−347.
13. L.H. Staib and J.S. Duncan: Parametrically Deformable Contour Models, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, 1989, pp. 427−430.
14. H.I. Bozma and J.S. Duncan: Model−Based Recognition of Multiple Deformable Objects Using a Game−Theoretic Framework, in Information Processing in Medical Imaging − Proceedings of the 12th International Conference, pp358−372, Springer−Verlag, 1991.
15. A. Pentland and S. Sclaroff: Closed−Form Solutions for Physically Based Modelling and Recognition, IEEE Trans. on Pattern Analysis and Machine Intelligence. **13**, 1991, 715−729.
16. P. Karaolani, G.D. Sullivan, K.D. Baker and M.J. Baines: A Finite Element Method for Deformable Models. Proceedings of the Fifth Alvey Vision Conference, Reading, 1989, pp. 73−78.
17. U. Grenander, Y. Chow and D.M. Keenan: Hands. A Pattern Theoretic Study of Biological Shapes. Springer−Verlag, New York, 1991.
18. K.V. Mardia, J.T. Kent and A.N. Walder: Statistical Shape Models in Image Analysis, Proceedings of the 23rd Symposium on the Interface, Seattle 1991, pp 550−557.
19. F.L. Bookstein, Morphometric Tools for Landmark Data, Cambridge University Press, 1991.
20. F.L. Bookstein: Principal Warps : Thin−Plate Splines and the Decomposition of Deformations. IEEE Trans. on Pattern Analysis and Machine Intelligence. **11**, 1989, 567−585.
21. K. Lindley, Model Based Interpretation of Lumbar Spine Radiographs. MSc Thesis, University of Manchester, 1992.
22. R.A. Johnson & D.W. Wichern: Multivariate Statistics, a Practical Approach, Chapman & Hall, 1988.
23. D.R.Bailes, C.J.Taylor: The Use of Symmetry Chords for Expressing Grey Level Constraints. in Proc. British Machine Vision Conference. Springer−Verlag, 1992, pp.296−305.
24. A.Lanitis: Applications of Point Distribution Models in Handwritten OCR and Face Recognition. Transfer Report, Wolfson Image Analysis Unit, Manchester University, 1992.
25. L.Davis : Genetic Algorithms and Simulated Annealing,Pitman, London, 1987.