An Adaptive Eigenshape Model

Adam Baumberg and David Hogg
School of Computer Studies
University of Leeds, Leeds LS2 9JT, U.K.
amb@scs.leeds.ac.uk

Abstract

There has been a great deal of recent interest in statistical models of 2D landmark data for generating compact deformable models of a given object. This paper extends this work to a class of parametrised shapes where there are no landmarks available. A rigorous statistical framework for the eigenshape model is introduced, which is an extension to the conventional Linear Point Distribution Model.

One of the problems associated with landmark free methods is that a large degree of variability in any shape descriptor may be due to the choice of parametrisation. An automated training method is described which utilises an iterative feedback method to overcome this problem. The result is an automatically generated compact linear shape model. The model has been successfully applied to a problem of tracking the outline of a walking pedestrian in real time.

1 Introduction

Statistical Analysis of 2D landmark data has become a well established tool in computer vision (e.g. morphological methods [1]). A significant advance in this area is the Point Distribution Model (PDM) introduced by Cootes et al. ([2, 3, 4, 5]). The PDM is a statistical model based on a set of example shapes of a given object. Each shape is described by a set of landmark points which correspond to particular (often biological) features around the object. The advantage of this approach is that a whole class of objects or a single deforming non-rigid object can be described by a relatively small set of shape parameters. These models have proven useful in image analysis (e.g. in medical images [4]) and image sequence analysis (see [6], [7]).

This paper aims to tackle some of the problems associated with extending the landmark based PDM to the problem of modeling continuous deformable contours. The aim is to build a compact, contour model that describes the shapes in a training set. The more compact the model, the fewer shape parameters are required for accurate representation which leads to faster and more efficient image search and object tracking procedures. A more compact model also increases robustness by producing a more restricted solution space of feasible shapes.

Previous work describes one approach to this problem where the control points of a cubic B-spline are treated as landmark points (Baumberg and Hogg [8]). This paper extends this work in two significant ways. Firstly a more rigorous eigenshape model based on the PDM is derived which takes into account the measurement noise characteristics for

BMVC 1995 doi:10.5244/C.9.9

the control points. Secondly, the model is made more compact by eliminating some of the variability caused by control points shifting along the contour (which cause little change to the actual observed shape). This work has some similarity to the work of Revow et al. [9] in that a covariance matrix associated with control point positions is learned from training data using an iterative learning process.

Hill and Taylor outline an approach to automatically choosing landmark points from training shapes [10]. This paper describes a simpler alternative approach to automating the model building process. An advantage of our method is that the complete contour shape is modeled as opposed to selected points on the boundary. The eigenshape model derived here retains many of the benefits of the conventional PDM. The model is object-specific containing a priori knowledge of the expected shape of the class of contours of interest.

2 Background: the PDM

The linear "Point distribution model" of Cootes et al. [2] is a statistical model of a training set of shapes described by n landmark points. The shapes are aligned to the mean shape, $\overline{\mathbf{x}}$, and the differences from the mean are analysed using principal component analysis. Hence for each aligned shape the vector $\mathbf{dx}$ is calculated as follows:

$$\mathbf{dx} = \mathbf{x} - \overline{\mathbf{x}}$$

where each training vector $\mathbf{x} = (x_1, y_1, ..., x_n, y_n)$ describes a set of landmark positions (x_i, y_i).

The $2n$ x $2n$ covariance matrix $C = E(\mathbf{dx}\mathbf{dx}^T)$ is then calculated where $E(...)$ is the expectation or mean value over the training set. The eigenvectors of the covariance matrix correspond to modes of variation of the training data. Moreover, the eigenvector corresponding to the largest eigenvalue describes the most significant mode of variation. The resulting model consists of the mean shape, $\overline{\mathbf{x}}$, and a subset of t eigenvectors $\mathbf{p_1}, \mathbf{p_2}, \ldots, \mathbf{p_t}$ corresponding to the t most significant modes of variation in the training data.

3 Extending the PDM to parametrised curves

3.1 Theoretical considerations

An implicit assumption of the PDM is that each landmark point position or displacement is measured independently with 2D isotropic Gaussian noise of fixed variance r (i.e. the measurement covariance matrix is rI_{2n}). Point measurements are mapped to the "modes of variation" shape parameters using:

$$\mathbf{b} = P^T(\mathbf{x} - \overline{\mathbf{x}})$$

where $\mathbf{b}$ is a vector of t shape parameters, $\mathbf{x}$ is a $2n$ vector describing the (2D) point positions and P is a matrix whose columns are eigenvectors of the covariance matrix. As P^T is an orthonormal transformation, the above assumption leads to the result that the shape parameters are also measured independently (with a measurement covariance matrix equal to rI_t).

This result coupled with the fact that the modes are *a priori* derived to be linearly independent, provides a theoretical basis for treating the shape parameters independently in the Active Shape Model search mechanism described by Cootes and Taylor ([3]).

Consider a set of parametrised training shapes represented by a set of nodal positions (e.g. an n control point cubic B-spline). The shapes are continuous closed curves and the measurement model is assumed Gaussian, isotropic unbiased and homogeneous. Intuitively this corresponds to a dense set of point measurements along the continuous curve at regular parametric intervals. In general any particular point measurement will effect the position of several nodal positions. (e.g. in the case of a cubic spline each point measurement effects 4 control points) and hence it is incorrect to assume independence of nodal measurements. In fact Blake et al. show the appropriate measurement covariance matrix is given by:

$$R = r\mathcal{H}^{-1}$$

with

$$\mathcal{H}_{i,j} = \int H_i(u)H_j(u)du$$

, where H_i is the interpolation function for the i'th nodal parameter (see [11]).

3.2 The linear eigenshape model

Given a known *a priori* covariance matrix S obtained from a training set the following eigenproblem can be solved to obtain a set of $2n$ eigenshapes

$$S\mathcal{H}\mathbf{e_i} = \lambda_i \mathbf{e_i}$$

Defining the inner product $\langle \mathbf{u}, \mathbf{u} \rangle = \mathbf{u}^T\mathcal{H}\mathbf{u}$ the following results are obtained:

1. The eigenvectors $\mathbf{e_i}$ are orthogonal with respect to this inner product. By scaling the vectors appropriately we can enforce orthonormality.

2. Given a new shape vector $\mathbf{u} = \sum_{i=1}^{2n} \mu_i \mathbf{e_i} + \overline{\mathbf{u}}$

$$\mu_i = \langle \mathbf{u} - \overline{\mathbf{u}}, \mathbf{e_i} \rangle$$

3. Assuming the previous measurement model, the covariance matrix for the coefficients μ_i is given by

$$R_\mu = rI$$

   i.e. measurements are independent

4. Over the training set the coefficients μ_i are linearly independent i.e.

$$E(\mu_i \mu_j) = \begin{cases} 0 & : \quad i \neq j \\ \lambda_i & : \quad i = j \end{cases}$$

The above properties allow the eigenshape basis to be treated in an analogous way to the eigenvectors in the standard PDM. Hence the coefficients associated with the eigenshape modes of variation are updated independently and a truncated basis of eigenshapes can be used as in the PDM. This model can be regarded as a Finite Element physical system with mass matrix $\mathcal{H}$ and stiffness matrix S^{-1}.

4 Generating an initial model from a training set

In order to generate the eigenshapes, a covariance matrix for the training set is required. Each training shape consists of an arbitrary long set of boundary points. The training information is usually the output of some image processing segmentation such as background subtraction or colour-based segmentation. Alternatively shapes may be hand segmented using a suitable interactive curve drawing tool.

In order to proceed, each shape is represented by a fixed length shape vector consisting of nodal parameters that represent an approximation to the curve. The method used is detailed in previous work [8] but a brief summary is given here. Each point on the boundary is associated with a parameter value. A suitable fixed point is chosen to have parameter value zero using the principal axis of the boundary set. (The fixed point is chosen to be the point at which the principal axis crosses the object boundary). The remaining parameter values are defined by using the arc-length around the boundary from the fixed point to each boundary point. The boundary can then be approximated using a cubic B-spline with a fixed number of control points.

It is important to note the following points:

- As this is only the initial step in an iterative process the exact method of representing the shapes is not critical. For instance, if the shapes are reasonably well registered the fixed point may be the upper most boundary point.
- This arc-length parametrisation does not ensure that physically corresponding points will always have the same parameter values.
- Apparently similar shapes may have quite different nodal representations due to variation in the placement of control points. (i.e. due to variation in the material parameter values of corresponding boundary points).
- The resulting eigenshape model may not be as compact as desired due to this additional variability.

The shapes are aligned in exactly the same manner as in the PDM treating the nodal control points as landmark points. The covariance matrix for the nodal points, S, is calculated in the usual way. i.e.

$$S = E(\mathbf{u}\mathbf{u}^T) - E(\mathbf{u})E(\mathbf{u}^T)$$

where $\mathbf{u}$ is an aligned shape vector representing the nodal positions of a training shape.

5 Adaptively improving the model

5.1 Active search using the model

We will require a method for fitting the linear shape model to an image containing an example of the object. In previous work, we describe a (static) Kalman filter mechanism for locally optimising the shape parameters of an eigenshape model along with the position, scale and orientation parameters [7]. The method is an extension of the Active Shape Model of Cootes et al. [3] adapted for continuous curve measurements and allows the shape parameters to vary slowly over time to track a deforming object through an image sequence. The following points are of interest:

- For the first frame of the sequence, the estimated variance of each shape parameter is initialised to some multiple of the associated eigenvalue (allowing the more significant modes to vary more freely). This method allows all the $2n$ shape parameters to be used when necessary. The initial estimate of each shape parameter is set to zero.
- For subsequent image frames a noise term is added to the estimated variance allowing the shape parameter to vary slowly.
- The Kalman filter mechanism can be regarded as a physical system where there are internal forces pulling the shape parameters towards the current shape estimate. The filter is suitable for robust and fast tracking but may lead to compromise solutions when the internal forces balance the image forces.
- The filter mechanism can be adapted for image search on a single frame by regarding the image frame as an image sequence of N identical images. By tracking over this new sequence the final contour position can be very accurately recovered (at the expense of computational speed).

5.2 Improving the model: initial approach

An obvious approach to "bootstraping" the eigenshape model is to utilise the active search mechanism on training images. The resulting shape parameters, μ_i, can be mapped into corresponding shape vectors and this new training set used to calculate a new mean shape and covariance matrix.

It is assumed that high quality (possibly pre-segmented) training images are available in which the approximate location, size and orientation of the object are known. A new set of training shape vectors can be obtained by running the active search method on these images. The new training shape vectors are aligned and a new covariance matrix generated. Note the parametrisation of the shapes is no longer explicitly calculated but implicitly derived from the old eigenshape model. The process may be iterated utilising the previous model to generate a new set of training shape vectors and a new model.

5.3 Iterative Method

Even if the full set of $2n$ eigenmodes are utilised in the active search method, variations which do not occur within the initial training set will never become apparent in subsequent models. In recognition of the fact that the initial model is only an estimate of the optimal model an additional step is taken. The current eigenshape model is perturbed by a simulated noise process. The eigenvalues λ_i are updated as follows

$$\lambda_i' = \lambda_i + \sigma$$

This is equivalent to adding Gaussian isotropic noise with variance σ to the training shapes. i.e. generating a new covariance matrix S' given by

$$S' = S + \sigma\mathcal{H}^{-1}$$

This step allows (arbitrary) small perturbations in the nodal positions. This hybrid model allows fine detail that is not well represented in the original model, to be more accurately recovered. This method is similar to the method employed by Cootes and Taylor to combine the PDM with a Finite Element physical model [12]. It is important to note that *all* the eigenmodes are used since the noise process ensures that no mode of variation can be regarded as insignificant. The Kalman filter active search mechanism allows the more significant modes to vary more easily so that all of the $2n$ modes can be employed without the method becoming unstable.

The parameter σ is initially set to around 8 pixels and subsequently decreased gradually. A diagram illustrating the scheme is shown in fig (1).

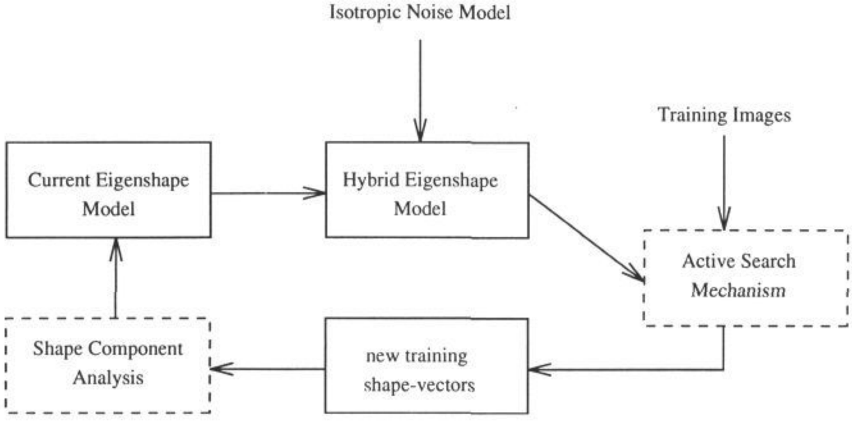


Figure 1: Diagram illustrating the method

6 Results

6.1 Single walk data set

The data set contained 59 shapes (silhouettes) segmented from an image sequence of a pedestrian walking from left to right across the image. Background subtraction was used to segment the silhouette of the walker (The method is described in previous work [8]). Four of these training shapes are show in fig.2. The feedback scheme described above

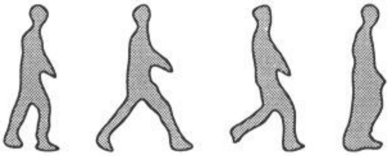


Figure 2: training shapes

was implemented with and without the additional noise process. Each iterative step generated a new eigenshape model which was then used for subsequent active image search.

A "compactness" measure was calculated for each model as follows:

$$\text{compactness} = \frac{\lambda_1 + \lambda_2}{\sum_{i=1}^{2n} \lambda_i} \times 100\%$$

where λ_1, λ_2 are the two largest eigenvalues in the model. The compactness measures the percentage the principal two "modes of variation" contribute to the total variance. A large compactness measure indicates that most of the variance is encapsulated by these two modes. The compactness of each model is shown in fig.4. The graph indicates that in both cases the compactness increases from under 65% for the initial model to almost 90% for the final, adapted model. It is apparent that the additive noise process has little effect on this increase in compactness.

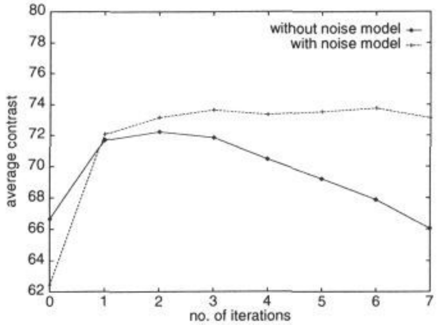


Figure 3: Average 'fitness'

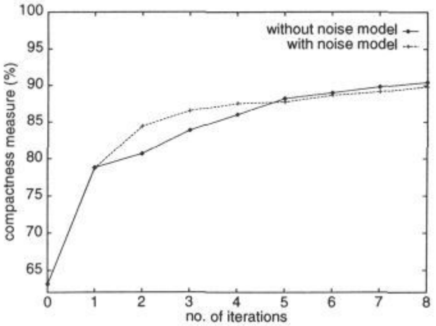


Figure 4: Compactness of models

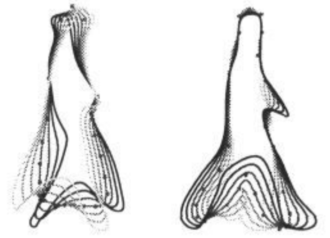
A "fitness" measure was also calculated at each iteration. This was a crude measure of how close the final contour lay to the true object shape after each image search. In fact the average image contrast at sampled points on the contour was used and this "fitness" was averaged over the training set. A high average fit indicates that most of the contour points lie close to an edge and hence the segmentation is accurate. The results for both methods are show in fig.3. The plot shows that without the noise process the benefits of increasing compactness are offset by the decrease in average fit. However the inclusion of the noise process generally gives a better fit and the average fit reaches a stable maximum.

Note that these plots show that the iterative process converges quickly (due to the fact that the initial model is fairly good), with the significant improvements occurring within the first few iterations.

Figure 5(a) shows a graphical representation of the effect of varying the principal shape parameter in the initial model. Figure 5(b) shows the principal mode of variation for the final adapted model. It appears that there is more information encapsulated in the principal mode of the adapted model.

6.2 Large data set

A second data set was generated containing 462 shapes of the silhouette of a pedestrian walking in a variety of directions. A sample of the training shapes is show in fig.6. In this experiment the results of the two methods were very similar. This was probably due to the



(a) initial model (b) final model

Figure 5: Principal (eigenshape) modes of variation

fact that the initial data set is very large and already quite noisy. Hence there is no need to add simulated noise. Results are shown for the simpler scheme outlined in sec.5.2.

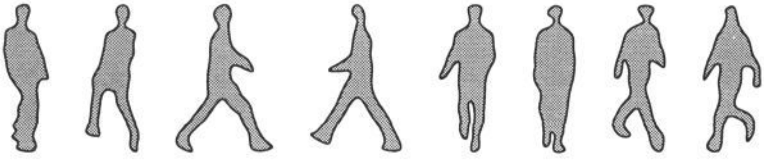


Figure 6: training shapes from large training set

As before the fitness and compactness measures were calculated. The results are shown in figs 7, 8. Figure 9 shows the first 10 eigenvalues for each successive model. The principal variation modes of the initial and adapted eigenshape models are shown in figs 10(a) and 10(b). This adapted model has been successfully utilised to track a moving pedestrian in real time with an increase in performance (over the initial model) due to a smaller number of shape components being required for shape representation.

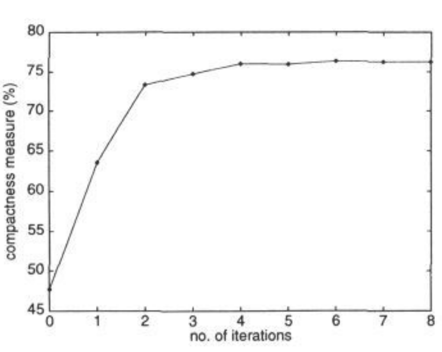


Figure 7: Compactness of models

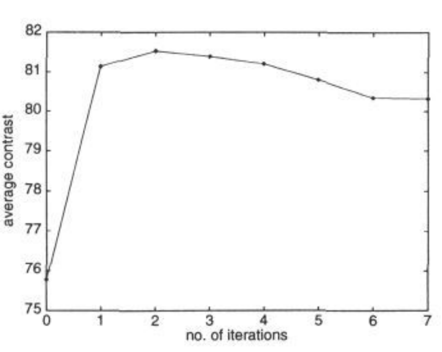


Figure 8: Average fit

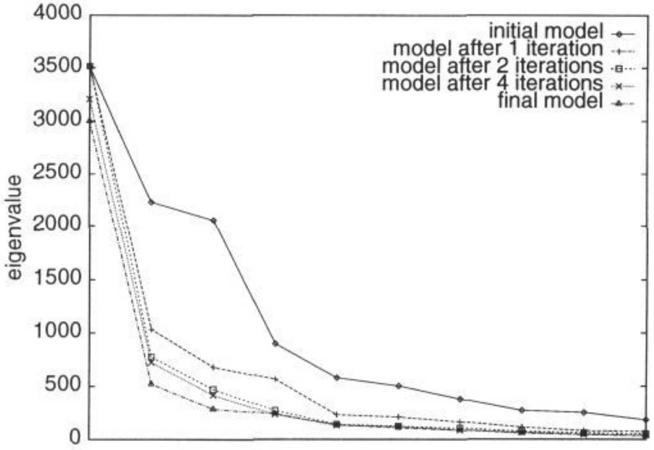
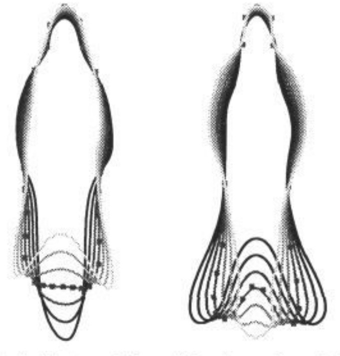


Figure 9: Eigenvalues of shape models



(a) initial model (b) adapted model

Figure 10: Principal modes of variation

7 Conclusions

We have described a method for the treatment of closed parametric contours which is related to the landmark based method of the PDM. One advantage of utilising a spline shape representation is that object shape is modeled between nodes allowing efficient calculation of the position and normal to the curve for any parametric value. The spline representation also allows a more efficient method for calculating a statistical shape model for continuous curves than using a suitably dense set of sampled boundary points and the resulting shape model is also more compact in terms of memory requirements.

A simple initial training method is described which can then be adapted using an iterative feedback mechanism. The iterative scheme reduces the variability in the model due to control points shifting along the contour (causing little change in the observed shape). There is no loss of accuracy in the adapted model (i.e. the training shapes are well represented by the new model). The result is a compact eigenshape model suitable for image search and object tracking. The method is completely automatic and results have been shown for several real noisy data sets.

References

[1] Bookstein F. *Morphometric Tools for Landmark Data*. Cambridge University Press, 1991.

[2] Cootes T.J., Taylor C.J., Cooper D.H., and Graham J. Training models of shape from sets of examples. In *British Machine Vision Conference*, pages 9–18, September 1992.

[3] Cootes T.F. and Taylor C.J. Active shape models - 'smart snakes'. In *British Machine Vision Conference*, pages 276–285, September 1992.

[4] Cootes T.F., Hill A., Taylor C.J., and Haslam J. The use of active shape models for locating structures in medical images. In Barrett H.H. and Gmitro A.F., editors, *Information Processing in Medical Imaging*, pages 33–47, 1993.

[5] Cootes T.F., Taylor C.J., Lanitis A., Cooper D.H., and Graham J. Building and using flexible models incorporating grey-level information. In *International Conference on Computer Vision*, May 1993.

[6] Kervann C. and Heitz F. Robust tracking of stochastic deformable models in long image sequences. In *IEEE International Conference on Image Processing*, volume 3, pages 88–92. IEEE Computer Society Press, November 1994.

[7] Baumberg A. and Hogg D. An efficient method for contour tracking using active shape models. In *IEEE Workshop on Motion of Non-rigid and Articulated Objects*, pages 194–199, November 1994. available as tech report 11, http://agora.leeds.ac.uk/scs/doc/reports.

[8] Baumberg A. and Hogg. D. Learning flexible models from image sequences. In *European Conference on Computer Vision*, volume 1, pages 299–308, May 1994.

[9] Revow M., Williams C.K.I., and Hinton G.E. Using generative models for handwritten digit recognition. Submitted to PAMI, July 1994.

[10] Hill A. and Taylor C.J. Automatic landmark generation for point distribution models. In *British Machine Vision Conference*, volume 2, pages 429–438. BMVA Press, 1994.

[11] Blake A., Curwen R., and Zisserman A. A framework for spatio-temporal control in the tracking of visual contours. *International Journal of computer Vision*, 1993.

[12] Cootes T.F. and Taylor C.J. Combining point distribution models with shape models based on finite element analysis. In *British Machine Vision Conference*, volume 2, pages 419–428, 1994.