Locating Objects of Varying Shape Using Statistical Feature Detectors

T.F. Cootes and C.J.Taylor

Department of Medical Biophysics, University of Manchester, Oxford Road, Manchester. M13 9PT, UK
email: bim,ctaylor@sv1.smb.man.ac.uk

Absract. Most deformable models use a local optimisation scheme to locate their targets in images, and require a 'good enough' starting point. This paper describes an approach for generating such starting points automatically given no prior knowledge of the pose of the target(s) in the image. It relies upon choosing a suitable set of features, candidates for which can be found in the image. Hypotheses are formed from sets of candidates, and their plausibility tested using the statistics of their relative positions and orientations. The most plausible are used as the initial position of an Active Shape Model, which can then accurately locate the target object. The approach is demonstrated for two different image interpretation problems.

1 Introduction

Image search using deformable models has been shown to be an effective approach for interpreting images of objects whose shape can vary [1,2,3]. Usually the object of interest is located by some form of local optimisation so a 'good enough' starting approximation is required. Such starting points are either supplied by user interaction or obtained in some application-specific manner. We wish to develop a system which can automate the generation of such starting points for a general class of models. This paper proposes a framework for generating hypotheses for all plausible instances of variably shaped objects in a scene. The approach is to determine a set of key features, use statistical feature detectors to locate all examples of these in a scene and then to generate a ranked list of all plausible combinations of features. We systematically consider all possible sets of features, ranking or eliminating each by considering the statistics of the relative positions and orientation of feature points using statistical shape models [4]. Missing features are dealt with, allowing robustness to occlusion. The best feature sets are then used to instantiate a deformable model known as an Active Shape Model [3,4] which can be run to locate the full structure. This leads to a generally applicable method, which can locate multiple instances of a target in a given scene.

In the following we will describe the approach in more detail. We will consider related approaches and describe the statistical feature detectors we use. We will then cover the statistical shape models and show how they can be used to determine how plausible a set of features is. Finally we will show the method working on real data and give results of systematic experiments.

2 Background

Many people have studied the problem of locating rigid objects in images. A review is given in Grimson [5]. In most approaches an image is preprocessed to locate features such as edges or corners, and the best matches of these to a model are located by a suitably pruned tree search. In general the simpler the features, the more possible matches there are with the model and the more expensive the resulting combinatorial explosion. The Local Feature Focus method of Bolles and Cain [6] attempts to deal with this problem both by finding sets of more distinctive features

and by choosing sub-sets of model features which are sufficient to identify and position the model. Work by Ashbrook *et al* [7] on Geometric Histograms is an example of an attempt to reduce the combinatorial explosion by making the feature models more detailed, and thus to generate fewer responses for each detector (usually at the expense of greater cost of locating the features).
Where the objects of interest can vary in shape, deformable models such as the 'snakes' of Kass *et al* [1], the finite element models of Pentland and Sclaroff [2] or Active Shape Models [3] have proved useful. However such models are usually used in a local optimisation schemes, requiring a suitable initial position to be provided. Most of the methods proposed for locating rigid objects rely on tight constraints between the positions and orientations of features, which are violated when the objects can deform. Loosening the constraints in simple ways often leads to a combinatorial explosion [5]. More complex models of the constraints are required.
Yow and Cipolla [8] describe a system for locating faces which uses a gaussian derivative filter to locate candidates for the lines of the eyes, nostrils and mouth, then uses a belief network to model the face shape and select a set of features most likely to be those of the face.
Burl *et al* [9] combine feature detectors with statistical shape models to generate a set of plausible configurations for facial features. Their (orientation independent) feature detectors match template responses with the output of multi-scale gaussian derivative filters. They use distributions of shape statistics given by Dryden and Mardia [10] to test configurations of outputs from detectors which locate possible locations of the eyes and nostrils. They build up hypothesis sets by first considering all pairs of features. From each pair they determine the regions in which they would expect candidates for other features to lie, and consider all sets of points which lie in these regions. They allow for missing features in their derivations, to give robustness to feature detector failure.
Our general approach is similar to that of Burl *et al*. However, we use more complex statistical feature detectors in order to minimise the number of false positive responses. Our detectors are orientation dependent, making them more discriminating. Although they must be run at multiple angles, they return both the position and orientation of the found features. We use a simpler method of calculating the shape statistics to test sets of points, and include the feature orientation in our tests. We systematically calculate all plausible sets of features using a depth first tree search, pruning sub-trees as soon as they become implausible. Our aim is to rapidly determine sets good enough to initialise an Active Shape Model to locate the object of interest accurately.

3 Overview Of Approach

The approach we use is as follows. We assume that we have sets of training images, in which points are labelled on the the objects of interest. In advance we determine a sub-set of points which can be used as features to detect the objects of interest, by considering the performance of feature detectors trained at every point. We train feature detectors for the chosen points, and build statistical models of both the shape of the whole model and that of the sub-set of feature points.
Given a new image, we proceed as follows:

- We find all responses for the feature detectors
- We systematically search through all these responses to determine all plausible sets of features. How plausible a set is is determined by the relative positions of the candidate features and their relative orientations
- We fit the full shape model to the best sub-set, and use this as the starting point for an Active Shape Model, which can then accurately locate the object of interest.

4 Statistical Feature Detectors

In order to locate features in new images we use statistical models of the grey levels in regions around the features, and use a coarse-to-fine search strategy to find all plausible instances in a new image [14].

4.1 Statistical Feature Models

We wish to locate accurately examples of a given feature in a new image. To deal with variations in appearance of features we use statistical models derived from a set of training examples. These are simply rectangular regions of image containing instances of the feature of interest. For each n_x x n_y patch we sample the image at pixel intervals to obtain an $n = n_x n_y$ element vector, $\mathbf{g}$.

A statistical representation of the grey-levels is built from a set of s example patches, $\mathbf{g}_i$ $(i = 1..s)$. A Principle Component Analysis is applied to obtain the mean, $\bar{\mathbf{g}}$, and t principle modes of variation represented by the n x t matrix of eigenvectors, $\mathbf{Q}$. The value of t is chosen so that the model represents a suitable proportion of the variation in the training set (eg 95%) [4].

Our statistical model of the data is

$$\mathbf{g} = \bar{\mathbf{g}} + \mathbf{Qc} + \mathbf{r}_g \tag{1}$$

where the elements of $\mathbf{c}$ are zero mean gaussian with variance λ_i, the elements of $\mathbf{r}$ are zero mean gaussian with variance v_j and the columns of $\mathbf{Q}$ are mutually orthogonal. This is the form of a Factor Model [12], with $\mathbf{c}$ as the common factors and $\mathbf{r}_g$ as the errors.

Given a new example, $\mathbf{g}$, we wish to test how well it fits to the model. We define two quality of fit measures as follows,

$$f_1 = M_t + \frac{R^2}{V_r} \quad (2) \qquad\qquad f_2 = M_t + \sum_{j=1}^{j=n} \frac{r_j^2}{v_j} \quad (3)$$

where

$$M_t = \sum_{i=1}^{i=t} \frac{c_i^2}{\lambda_i} \qquad\qquad \mathbf{c} = \mathbf{Q}^T(\mathbf{g} - \bar{\mathbf{g}}) \tag{4}$$

$$R^2 = \mathbf{r}_g{}^T\mathbf{r}_g = (\mathbf{g} - \bar{\mathbf{g}})^T(\mathbf{g} - \bar{\mathbf{g}}) - \mathbf{c}^T\mathbf{c} \qquad\qquad \mathbf{r}_g = \mathbf{g} - (\bar{\mathbf{g}} + \mathbf{Qc}) \tag{5}$$

In [14] we show that the distribution of these fit values is a scaled chi-squared distribution of degree k, $p(f) = (n/k)\mathrm{X}^2(kf/n, k)$ where k is the number of degrees of freedom of the pixel intensities. k can be estimated from a verification set of examples; $k = 2(n/\sigma)^2$, where σ is the standard deviation of the distribution of fit values across the set.

The time to calculate f_1 is about half that for f_2, but gives slightly less predictable distributions and poorer discrimination between true positive and false positive responses. Knowledge of $p(f)$ allows us to set thresholds which will produce predictable numbers of false negatives (missed true features). More details of the statistical feature models are given in [14].

4.2 Properties of the Statistical Feature Detectors

The quality of fit of the feature models to an image is sensitive to position, orientation and scale. By systematically displacing the models from their true positions on the training set we can quantify this sensitivity. This information allows us to calculate the number of different scales and orientations at which the detector should be run in order to cover a given range of object orientation and scales.

4.3 Searching for Features

Given a new image we wish to locate all plausible instances of a given feature. We train feature models at several levels of a gaussian pyramid [13] and determine their sensitivity to angle and scale. We then use a coarse–to–fine search strategy as follows;

- Test every pixel of a coarse resolution image with the matching feature model at a set of angles and scales. Determine the peaks in response which pass the statistical threshold.
- Refine the position and angle estimates of each response on finer resolution images. The accuracy in angle required is determined by the (pre–computed) sensitivity of the model.

Those which pass the threshold at the finest resolution are candidate features. We can choose suitable resolutions at which to perform the search by performing tests on the training or verification set. The total work required to run the detector starting and ending at different resolution levels can be calculated. We choose those ranges of levels which can accurately relocate the true features, require least work and generate the fewest false positive responses.

4.4 Choice of Features

We wish to be able to select automatically a set of features suitable for locating the object of interest. The suitability of a given feature depends on a number of factors including the success rate of the feature detector, the number of false positives it tends to generate and the time it takes to run the detector over a given region.
For the experiments described below we have used training sets labelled with many points on and around structures of interest, which we use for building statistical shape models [3]. In order to obtain a set of suitable features we build models for every shape model point and test each one. By running each over the training set we can estimate the success rate, the number of false positives the detector generates and the computational cost for that detector. Those detectors which have the highest success rate (hopefully 100% on the training set) are ranked by the number of false positives.
For instance, Figure 1 shows the 40 best features for detectors trained on 122 landmark points on 40 face examples. Figure 2 shows the 8 best features for detectors trained on points on the outline of the brain stem in a 2D slice of a 3D MR image of the brain.

Fig. 1. The 40 best features for locating a face. (Tested 122 points)

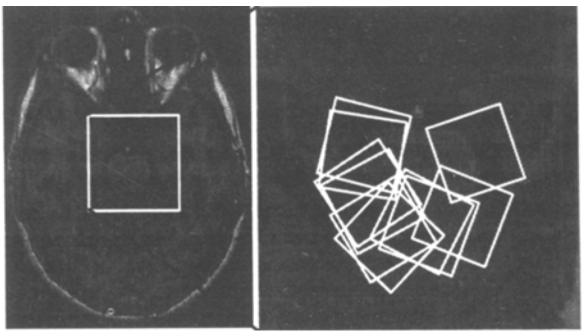

Fig. 2. The 10 best features for locating the brain stem in an MR image. (Tested 60 points).

We are experimenting with algorithms for automatically deciding on sets of features, but at present the sets are chosen by the user, guided by these rankings and the desire to have the features reasonably spread out over the target object.

5 Use of Statistical Shape Models to Test Hypotheses

When presented with a new image we apply the selected feature detectors over target regions to generate a number of candidates for each feature. By choosing one candidate for each feature we can generate a hypothesis for sets of features belonging to the same object. We will use statistical shape models to determine the plausible sets by considering the relative positions of the features and the orientations at which each was detected.

5.1 Statistical Shape Models

We have previously described methods for building statistical shape models. Given a training set of shapes, each representing n labelled points, we can find the mean configuration (shape) and the way the points tend to vary from the mean [3,4]. The approach is to align each example set into a common reference frame, represent the points as a vector of ordinates in this frame and apply a PCA to the data. We can use the same formulation as for the grey-level models above,

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{Pb} + \mathbf{r} \tag{6}$$

where $\mathbf{x} = (x_1 \; \ldots \; x_n \; y_1 \; \ldots \; y_n)^T$, $\mathbf{P}$ is a $2n$ x t matrix of eigenvectors and $\mathbf{r}$ is a set of residuals whose variance is determined by miss-one-out experiments. In this case t is the number of *shape* parameters required to explain say 95% of the shape variation in the training set.
Again, the quality of fit measure for a new shape is given by

$$f_{shape} = \sum_{i=1}^{t} \frac{b_i^2}{\lambda_i} + \sum_{j=1}^{j=2n} \frac{r_j^2}{v_j} \qquad \mathbf{b} = \mathbf{P}^T(\mathbf{x} - \bar{\mathbf{x}}) \qquad \mathbf{r} = \mathbf{x} - (\bar{\mathbf{x}} + \mathbf{Pb}) \tag{7}$$

Which should be distributed approximately as chi-squared of degree $2n-4$.
In the case of missing points, we can reformulate this test using weights (1.0 for point present, 0.0 for point missing);

$$f_{shape} = \sum_{i=1}^{t} \frac{b_i^2}{\lambda_i} + \sum_{j=1}^{j=2n} w_i \frac{r_j^2}{v_j} \tag{8}$$

where in this case $\mathbf{b}$ is obtained as the solution to the linear equation

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

($\mathbf{W}$ is a diagonal weight matrix).
This measure will be distributed as chi-squared of degree $2n_v - 4$ where n_v is the number of points present.

5.2 Models of the Feature Sets

Our features represent a sub-set of the points making up the full shape model for the object of interest. For each such sub-set we can generate statistical models of the configurations of the feature positions as described above. For instance for the face model we choose features at four of the 122 points of the full model and build statistical models both of the whole set and of the four points. Each shape model has its own co-ordinate frame (usually centred on the centre of gravity of the points and with some suitably normalised scale and orientation [3]).
To test the validity of a set of image points forming a shape, $\mathbf{X}$, we must calculate the shape parameters $\mathbf{b}$ and the pose Q (mapping from model frame to image) which minimise the distance of the transformed model points, $\mathbf{X'}$, to the target points

$$\mathbf{X} \approx \mathbf{X}' = Q(\bar{\mathbf{x}} + \mathbf{Pb}) \tag{10}$$

(Q is a 2D Euclidean transformation with four parameters, t_x, t_y, s and Θ.)

This is a straightforward minimisation problem [4,3]. Having solved for Q and **b** we can project the points into the model frame using Q^{-1} and calculate the residual terms and hence the quality of fit, f_{shape}. We can test the plausibility of the shape probabalistic limits both to the overall quality of fit f_{shape} and, if desired, to the individual shape parameters b_i. The latter have zero mean and a variance of λ_i, the eigenvalues obtained from the PCA.

By considering the training set we can calculate the average mapping between the co-ordinate frame for the full model and that for a sub-set of points. This allows us to propagate any known constraints on the pose of the whole object model to test the pose of the sub-set of points representing the current feature set. In addition, we can learn the expected orientation and scale of each feature relative to the scale and orientation of the set as a whole, allowing further discrimination tests.

If we assume that the configuration of the sets is independent of the errors in feature orientation and scale we can estimate the probability density for a configuration as follows;

$$p = p(shape)\prod_{i=1}^{i=n_f}[p(\theta_i)p(s_i)] \tag{11}$$

where the probabilities for shape, angle and scale terms are determined from the estimated distributions. If we assume normal distributions for the measured orientations and scales, then

$$ln(p) = const + \frac{f_{shape}}{2} + \sum_{i=1}^{i=n_f}\left[\frac{(a_i - \bar{a}_i)}{2\sigma_{ai}^2} + \frac{(s_i - \bar{s}_i)^2}{2\sigma_{si}^2}\right] \tag{12}$$

This allows us to sort any plausible hypotheses by their estimated probability.

5.3 Systematic Hypothesis Generation and Testing

If we have n_f feature detectors, and detector i produces m_i candidates, then there are $\prod m_i$ possible sets. If we allow for the detectors missing true features, then there are $\prod(m_i + 1)$ possible sets (allowing a wildcard feature match). Selecting plausible sets of features given this potential combinatorial explosion has received much attention [5,9].

We have used a relatively simple scheme amounting to a depth first tree search. The feature candidates are sorted by quality of fit and, if missing features are to be allowed, a wildcard is added. We then recursively construct sets, starting with a candidate from the first detector and adding each candidate from the second detector in turn. The pose and shape of each pair is tested. If a pair passes the tests, each candidate from the third detector is added and the three points tested. Those sets which pass are extended with candidates from the fourth detector and so on. In this manner all possible plausible sets can be generated fairly efficiently. (This approach has the advantage that it can be implemented in a recursive algorithm in a small number of lines of code). We record all the sets which have at least three valid features and pass the statistical tests.

Burl *et al* [9] calculate a probability for each set of candidates which takes into account missing features. We feel that it is difficult to correctly assign probabilities for missing features and instead simply sort our hypotheses first by the number of features present, and secondly by their probability. This avoids comparing sets with different numbers of features directly. In practice it is those which have the fewest missing which tend to be the correct responses.

5.4 Verification of Plausible Feature Sets

Given a plausible set of features, we find the least-squares fit of the full object shape model to these points. This can be achieved by solving a weighted version of (10), with zero weights for all but the points corresponding to the found features. This gives the starting point for an Active Shape Model. We can run the ASM to convergence, using a multi-resolution search scheme to locate all the points [11]. The ASM has grey-level models of what it expects in the region around every one of its points. By considering the quality of fit ot these models to the image after convergence we can determine whether a good example of the object of interest has been found. Where there are several equally plausible sets of features, the ASM can be run for each and the one with the best final fit accepted. To detect multiple instances of the model the best examples which do not overlap in pose space should be accepted.

6 Results of Experiments

6.1 Performance of Hypothesis Tester

We have performed experiments to study how well the hypothesis testing scheme works. We generated synthetic feature detector responses by taking one true set of 5 feature positions and adding varying numbers of random responses to each candidate list. We then ran the hypothesis tester, recording the number of tests required to find all plausible feature sets. We tried this both allowing and not allowing missing features, and either using or not using information on the orientation of the features. Each experiment was performed 10 times. Figure 3 summarises the results for the number of tests required. Allowing missing features increases the number of tests significantly, but including angle information reduces the number by almost as much. Our current implementation tests about 500 sets each second. Of course, ordering the features so that those with the fewest responses are tested first can significantly reduce the work required. The number of plausible sets resulting varied between 1 (the correct one) and 3 (with >20 false responses) for each case except for the case of missing features with no angle information, which produced many spurious plausible sets when the number of false responses was large.

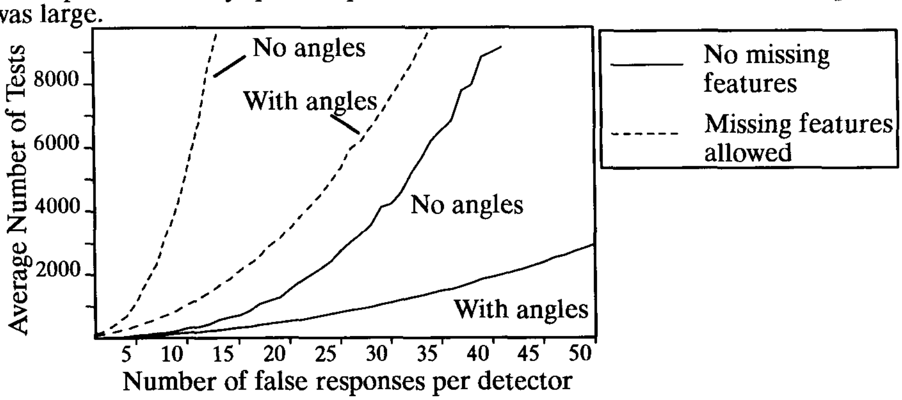

Fig. 3. Average number of tests required to find all plausible sets of candidates, given varying numbers of false responses for each feature detector.

6.2 Locating Facial Features

We have used the system to locate features as a way of initialising a face shape model which can be refined with ASM search. The full model is trained on 122 points marked in each of 40 images (a subset of those used by Lanitis *et al* [15]). Four features based around the eyes and nose were chosen from the set of most distinctive points shown above (Figure 1). Figure 4 shows an new face image, the positions of the detected features, the best set of such features and the position of the full 122 point shape model determined from this set. Figure 5 shows the points after running an ASM to convergence from this starting position. We assumed that we knew the approximate scale, but that the orientation and position were unknown. Each feature detector was run over the whole image and allowed any orientation. It took about 10 seconds on a Sun Sparc20 to run each feature detector over the 512^2 images, then one to two seconds to consider all plausible sets of features. In a real face detection system the orientation is likely to be better constrained, but the features would have to be allowed to vary in scale. Figures 6 and 7 show results for a different image, in which one of the feature detectors has failed to locate a satisfactory candidate. The quality of the full model fit to the found features is worse, but still quite adequate for the ASM to converge to a good solution. This demonstrates the robustness to missing features (and thus to occlusions).

Fig. 4. Candidate features (crosses), best feature set (boxes) and full model fit to the best feature set.

Fig. 5. Full model after running ASM to convergence.

Fig. 6. Candidate features (crosses), best feature set (boxes) (only 3 of 4 found) and full model fit to the best feature set.

Fig. 7. Full model after running ASM to convergence.

To test the performance more systematically we ran the system on a test set of 40 different images, which had been marked up with the target point positions by hand. On 5 of the images it failed to find any plausible sets, due to multiple feature detector failure. On one image false positives conspired to give a plausible (but wrong) result. On the other 35 the best set gave a good fit. The mean distance between the (known) target points and the full set of points estimated from the best features was 8.7 pixels. The mean error after running the ASM was 6.5 pixels. On average 3.5 of the 4 features were used, and each feature detector found 12 candidates in the image. By more careful choice of size of feature detectors, and by using more detectors (giving more robustness) and a larger training set we expect to improve the results significantly.

6.3 Locating the Brain Stem in MR Slices

We used the same approach to locate the outline of the brain stem in 2D MR slices. The full outline is represented by 60 points, and models trained from those points marked in 15 images. Five features were chosen from the set of the most distinctive points shown in Figure 2. Figure 8 shows candidates for these 5 in a new image, the best set of features and the full model fitted to these features. Figure 9 shows the points after running an ASM to convergence from this starting position. Again we assumed the scale was fixed but that the position and orientation of the brain stem were unknown. It took about 5 seconds on a Sun Sparc20 to run each feature detector over the 256^2 images, then a fraction of a second to consider all plausible sets of features. Of course in practice the position and orientation of the structure is fairly well constrained, so the system would run far quicker. In cases in which some of the feature detectors fail to locate satisfactory candidates the quality of the full model fit to the found features is worse, but again still quite adequate for the ASM to converge to a good solution.

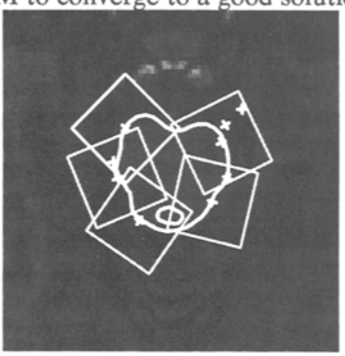

Fig. 8. Candidate features (crosses), best feature set (boxes) and full model fit to the best feature set.

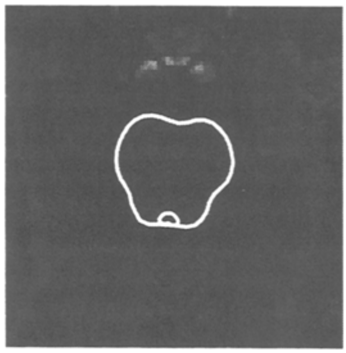

Fig. 9. Full model after running the ASM to convergence.

Again we ran systematic tests over a set of 15 new images. The system failed to find any plausible feature sets on 2 of them, but gave good results on the remaining 13. The mean difference between target points and the full set of points estimated from the best features was 1.5 pixels, falling to 1.1 pixels after the ASM was run to convergence. On average 3.8 of the 5 features were used, and each feature detector found 3 candidates in the image. We expect to improve on these results by using a larger training set and more careful choice of feature model size.

7 Discussion

Although we have used one particular form of statistical feature detector, any approach which located the features of interest could be used.

The calculation of plausible feature sets can take a long time if a large number of candidates are to be tested. We are currently interested in generating all plausible hypotheses. However, since the candidates are sorted by quality of fit, we usually find the best overall set quite early in the search. We intend to investigate early verification strategies, terminating the combinatorial search when we find a 'good enough' solution to explain the data.

8 Conclusions

We have shown how statistical feature detectors can be used to find good starting positions for deformable models, given no prior information on the position or orientation of the object of interest in the image. We used statistical models of the relative positions and orientations of the detected features to determine the plausible sets of features and to limit the possible combinatorial explosion by pruning bad sets as soon as possible. The plausible feature sets can be used to instantiate a statistical shape model representing the whole of the object of interest, which can then be refined using an Active Shape Model. This approach can locate multiple instances of the objects of interest in an image, and can be applied to a wide variety of image interpretation problems.

Acknowledgements

T.F.Cootes is funded by an EPSRC Advanced Fellowship Award.
The face images were gathered and annotated by A. Lanitis. The brain stem images were gathered by C. Hutchinson and his team. They were annotated by A.Hill.

References

1 M. Kass, A. Witkin and D. Terzopoulos , Snakes: Active Contour Models, in *Proc. First International Conference on Computer Vision*, pp 259-268 IEEE Comp. Society Press, 1987.
2 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.
3. T.F.Cootes, C.J.Taylor, D.H.Cooper and J.Graham, Active Shape Models - Their Training and Application. *CVIU* Vol. 61, No. 1, 1995. pp.38–59.
4. T.F.Cootes, A.Hill, C.J.Taylor, J.Haslam, The Use of Active Shape Models for Locating Structures in Medical Images. *Image and Vision Computing* Vol.12, No.6 1994, 355–366.
5. W.E.L. Grimson. Object Recognition By Computer : The Role of Geometric Constraints. Pub. MIT Press, Cambridge MA, 1990.
6. R.C. Bolles and R.A. Cain, Recognising and locating partially visible objects : The Local Feature Focus Method. *Int. J. Robotics Res.* **1**(3) 1982, pp. 57–82.
7. A.P.Ashbrook, N.A. Thacker, P.I. Rockett and C.I.Brown. Robust Recognition of Scaled Shapes using Pairwise Geometric Histograms. *in Proc. British Machine Vision Conference*, (Ed. D.Pycock) BMVA Press 1995, pp.503–512.
8. K.C.Yow, R.Cipolla, Towards an Automatic Human Face Localisation System. *in Proc. British Machine Vision Conference*, (Ed. D.Pycock) BMVA Press 1995, pp.701–710.
9. M.C.Burl, T.K.Leung, P.Perona. Face Localization via Shape Statistics. Proc. Int. Workshop on Automatic Face– and Gesture–Recognition (ed. M.Bichsel), 1995. pp. 154–159
10. I.L.Dryden and K.V.Mardia, General Shape Distributions in a Plane. *Adv. Appl. Prob.* 23, 1991. pp.259–276.
11. T.F.Cootes , C.J.Taylor, A.Lanitis, Active Shape Models : Evaluation of a Multi–Resolution Method for Improving Image Search, *in Proc. British Machine Vision Conference*, (Ed. E.Hancock) BMVA Press 1994, pp.327–338.
12. R.A.Johnson, D.W. Wichern, Applied Multivariate Statistical Analysis. *Prentice–Hall* 1982.
13. P.J.Burt, The Pyramid as a Structure for Efficient Computation. in *Multi–Resolution Image Processing and Analysis*. Ed. Rosenfield, pub. Springer–Verlag. 1984. pp. 6 - 37.
14. T.F.Cootes , G.J.Page, C.B.Jackson, C.J.Taylor, Statistical Grey–Level Models for Object Location and Identification, *in Proc. British Machine Vision Conference*, (Ed. D.Pycock) BMVA Press 1995, pp.533–542. (Also submitted to IVC)
15. A.Lanitis, C.J.Taylor and T.F.Cootes, A Unified Approach to Coding and Interpretting Face Images. Proc. 5th ICCV, IEEE Comp. Soc. Press, 1995, pp. 368–373.