Locating Facial Features Using Genetic Algorithms

A. Lanitis, A. Hill, T.F.Cootes and C.J.Taylor
Wolfson Image Analysis Unit, Dept. of Medical Biophysics,
University of Manchester, Oxford Road,
Manchester M13 9PT, UK
e-mail: lan@wiau.mb.man.ac.uk

Abstract
Detection of facial features is difficult because face images are highly variable. Sources of variability include individual appearance, 3D pose, facial expression and lighting. We describe a system for locating facial features, which uses flexible models of appearance as a means of providing prior knowledge about the expected appearance of features and the spatial relationships between them. A global search procedure based on Genetic Algorithms is used to fit the model to new face images, so that the exact position of a face and facial features can be recovered.

Introduction
Face images have received considerable attention from both the computer vision and signal processing communities. This interest is motivated by the broad range of potential applications for systems able to code and interpret face images. Examples include personal identification and access control, low-bandwidth communication for videophone and teleconferencing, automated surveillance and human-computer interaction. In order to tackle these applications successfully it is essential to be able to detect and locate the position of whole faces and possibly a number of facial features in face images. However, locating facial features is a difficult problem because the images involved are complex and are also highly variable, even for a particular individual. Sources of variability include pose, facial expression, individual appearance, lighting, and occluding structure (facial hair, spectacles etc).

We describe a top down approach for locating facial features. We use a statistical shape model (or a Point Distribution Model) of facial appearance which is derived from a training set of face images. The model provides a compact parametrised description of the shape for any instance of a face. This compact parameterisation enables the efficient use of the model for local and global image search. We have described elsewhere [6] image search using similar types of flexible models and genetic algorithms. We have built on that approach in order to produce a more reliable and more practical system suitable for a wide range of object recognition problems. In this paper we describe the improved system and present experimental results for locating facial features.

For our experiments we have used a face database [8] containing 600 images (10 training, 10 test and 3 difficult test images of each of 30 individuals). A description of the database and its contents are publicly available at the World Wide Web at the address: http://pipa.essex.ac.uk/ftp/ipa/pix/faces/manchester.

Point Distribution Models (PDMs)

Training
Point Distribution Models [2] are generated from sets of training shapes. Each training shape (X_i) is represented by a number of co-ordinate points.

\[ X_i = (x_{i1}, x_{i2}, \ldots, x_{in}) \] (1)

where \( x_{ik} \) is the kth landmark in the ith training example.

Equations 2 and 3 are used to calculate the average example (X) among the n training examples and the deviation of each example from the mean (d_i).

\[ X = \frac{1}{n} \sum_{i=1}^{n} X_i \] (2)

\[ d_i = X_i - X \] (3)

A Principal Component Analysis (PCA) is applied by calculating the eigenvectors and eigenvalues of the covariance matrix of the deviations. Usually there are correlations in the way that points move, so that most of the variation in the training set is explained by a small number of eigenvectors or by a small number of main modes of variation. Any shape in the training set can be approximated by the mean plus a weighted sum of the most important modes of variation.
\[ X_i = \overline{X} + Pb \]  

where \( P \) is the matrix of unit eigenvectors of the covariance of deviations.

\( b \) is a vector of weights (These are referred as shape parameters) By modifying \( b \) new instances of the model can be generated; if the elements of \( b \) are kept within some limits (typically within 3 standard deviations from the mean) the corresponding model instances are plausible examples of the modelled object. Since \( P \) is orthogonal \( P^{-1} = P^T \) thus equation 4 can be solved with respect to \( b \):

\[ b = P^T(X_i - \overline{X}) \]  

Equation 5 can be used to represent training examples by their shape parameters. To define a model instance in an image a set of shape parameters and 4 pose parameters (the \( x \) and \( y \) origin, a scaling factor and a rotation angle) are required.

PDMs can be used in Active Shape Models [4] to fit to new image objects in an iterative local optimization scheme. A PDM is placed on the image and is allowed to interact dynamically until it fits to the shape of the object presented. During each iteration two main operations take place: the definition of a new suggested position for each model point and the deformation of the model in order to move as close as possible to the new preferred positions. We have described this procedure in detail elsewhere [4].

**Building a Face PDM**

We have built a face PDM using 144 points which were planted on 160 training examples (8 examples from 20 individuals from the data set). Typical training examples and the locations of the model points are shown in figures 1 and 2 respectively. Sixteen shape parameters are needed to describe 95% of the variation within the training set, thus 20 parameters all together are required to define a model instance in an image (16 shape and 4 pose parameters). The effects of the 4 most important modes of variation on the model are shown in figure 3.

![Typical training shapes](image1)

![Model points located](image2)

![The main modes of shape variation](image3)

The face PDM can be used successfully for locating facial features in a local optimization search [4]. Example of the model fitting to a face image is shown in figure 4. However, when we have no indication of the approximate location and orientation of a face in an image, local image search is likely to fail; in these cases global rather than local search must be used.

**Genetic Algorithms (GAs)**

Genetic Algorithms [5, 7] are global search algorithms based on the adaptive processes of natural systems which underlie the survival and evolution of species. They are based on the idea that among a population of

![Initial placement](image4)

![5th Iteration](image5)

![10th Iteration](image6)

![Final Fitting](image7)
examples only members which are the fittest can survive and breed so that the examples in following generations are improved. GA search is illustrated diagrammatically in figure 5. The procedure starts by choosing randomly a number of possible solutions within the search space in order to generate an initial population. Based on an objective function calculated for each solution, the numbers of each member of the population allowed to take part in reproduction is determined. This procedure is called selection and results in discarding the members that show no promise in favour of the promising ones. All members are coded in a parametric form so that each solution is represented by a chromosome, which can be represented by a string (see figure 6). The most common way to code variables is by representing them in a binary (often grey-coded) form. Genetic operations are performed on chromosomes with the aim that evolved solutions should combine promising structures from their ancestors to produce an improved population. Usually two types of genetic operations are utilised; crossover and mutation [5]. The GA search procedure terminates when most of the members of the population converge to a single solution, which represents the outcome of the search procedure.

Image Search Using GAs/PDMs

We have already seen that PDMs can provide compact parametrised shape models of variable objects. Even for complex objects, the number of important shape parameters is small. For example, for the model representing human faces only 16 shape parameters are needed to capture most of the variability in the training set. Thus the total number of variables needed to be adjusted during the search procedure (shape parameters + pose parameters) is small and well suited to GA search.

In some cases, searching images for a particular shape structure is not easy because of the possible existence of multiple candidates, which provide local solutions to the problem. It is important that a search algorithm should proceed by investigating all possible plausible solutions in order to ensure that the outcome is a global optimum. Exploring more than one solution at a time is a feature of GAs. The combination of PDM representation and GA search can thus provide an effective basis for image interpretation.

We have described elsewhere [6] an image interpretation system using GAs and PDMs. The 4 pose parameters and the most important shape parameters of a PDM are encoded using binary coding, so that each parameter is represented by a string of bits. These strings are combined to generate a chromosome which corresponds to a possible instance of the model in an image. To reconstruct this instance the strings are decoded back to their numerical values and the resulting pose and shape parameters and equation 4 are used for calculating the corresponding model points in the image. In [6] the selection procedure uses an objective function based on the edge strength in the vicinity of model points (the scheme was used with PDMs for which model points were located on strong edges). The optimum solution of the GA/PDM search is a model instance, for which all model points are located on pixels of high edge strength. When this method is used in conjunction with the face PDM, each chromosome represents a possible instance of the model in the image as shown in figure 7. The aim is to use GA search to define the combination of shape and pose parameters required to approximate the shape of the face in a given image as well as possible.

Improving the performance of GA/PDM Search

Although the basic GA/PDM search method described in [6] have been used successfully for many image interpretation tasks, there are a number of improvements which might be considered:

- It would be desirable to use an objective function based on grey-level rather than edge information.
It would be desirable to implement a multi-resolution GA/PDM in order to reduce the computational search time and possibly improve the performance.

Using Grey-Level Information

We have investigated the use of local grey-level profile models [3], for calculating the objective function. To train these models a number of training profiles are needed to model the grey-level appearance at each model point. To obtain the training profiles, we overlay the training shapes on the corresponding training images and extract grey-level profiles in a direction locally perpendicular to the boundary (see figure 8). If there are n training shapes for the shape model there should also be n training profiles for each model point. For a specific model point, each training profile \( (G_i) \) is given by

\[
G_i = \begin{bmatrix} (g_{i1}), (g_{i2}), \ldots, (g_{iN_x}) \end{bmatrix}
\]

where \( N_x \) is the number of pixels in each training grey-level profile.

By performing a statistical analysis of the training profiles for each model point (similar to the one used for training PDMs), a flexible model of the grey-level landscape around each point can be obtained. Given a shape model instance on a new image, profiles perpendicular to the boundary at each model point can be extracted. The grey-level models associated with each model point can be scanned along an extracted profile (which is longer) and at each point a similarity measure \( (d) \) can be calculated. This procedure is illustrated in figure 9. By summing the best similarity measure \( (d_{\max}) \) recorded for each model point, an overall estimate of the goodness of a model instance can be obtained. This procedure is described in detail elsewhere [3, 11].

Running Multi-Resolution GA/PDM Search

A possible problem with the approach described in the preceding section is the dependency of the search procedure on the length of extracted profiles \( (l) \) in figure 9. If the lengths are short the image evidence used for assessing a solution may not be reliable enough to achieve robust performance. If \( l \) is too long then the computational complexity increases and the accuracy of locating points may decrease because a good fit measure may
be detected at a distance as far as \( l/2 \) from the model point, without any penalty on the objective function. It is reasonable to use a search strategy for which \( l \) is long for the starting generations, in order to ensure the approximate location of the object, and gradually decreases \( l \) in order to get a more accurate solution. A multi-resolution search scheme has the right properties [4]. During training a multi-resolution pyramid of images, is generated for each training image, by using gaussian smoothing and sub-sampling (see figure 10). The method described in the previous section is used for training a local grey-level model for each model point, at each resolution level. GA/PDM search starts at the coarsest resolution level, by using the grey-level models trained at that level for calculating the objective function and moves up the resolution levels as the correct solution is approached; the search is terminated at the finest resolution level. During the whole procedure, the lengths in pixels of the search profiles are kept constant. However, because the same profile length is used at different resolution levels, the actual image area sampled during the search procedure is quite large at the coarser resolution levels; at the higher resolution levels it decreases.

To implement the multi-resolution search scheme it is important to define a criterion to detect convergence at a particular resolution level, so that the progression to a higher resolution level can take place automatically. A criterion based on the distance between model points, and the points at which the best profile match is recorded \( (p_A \) in figure 9), is used [4, 11]. Convergence at a particular resolution level is declared when the proportion of model points for all members of the population for which \( p_A \) is within the central 50% of the extracted profile, is greater than a threshold. Usually this threshold is set to around 80%. Experiments have shown that the search results are not critically dependent on the numerical value of the threshold [11].

**Experimental Evaluation**

**Design of The Experiments**

In this section 3 different ways of performing GA/PDM search are investigated experimentally:

**Experiment 1** Single resolution search using an objective function based on edge strength.

**Experiment 2** Single resolution search using an objective function based on grey-level information.

**Experiment 3** Multi-resolution search using an objective function based on grey-level information.

A standard experimental procedure was used to assess each of the methods listed above. All experiments were run on 30 face images chosen randomly from the training images used for training the shape model (for this experiment we used images from the training set for testing because for these images we already had manually located landmarks, which we could use to assess the accuracy of automatic location). The performance of each method was assessed against the accuracy of locating facial features and the computational complexity. To assess the accuracy of locating facial features the average euclidean distance \( (d_i) \) between the model points of the best member of the population and the correct landmark positions was calculated.

**Results--Conclusions**

The results of our experiments are summarised in figure 11. From these graphs it is clear that GA/PDM search based on grey-level rather than edge information, performs better. Multi-resolution GA/PDM search achieves the best performance in terms of the accuracy of locating points. Model points were located with an average accuracy of about 3 pixels when multi-resolution GA/PDM search was used. The computational complexity is also reduced significantly when multi-resolution search is employed (when compared with single resolution search based on grey-level information), but is still not fast enough for real time applications. We have also used multi-resolution GA/PDM search for locating facial features on unseen and occluded face images from the test sets in the face database, as part of a face identification procedure. The recognition rates obtained [11] were very good implying that the accuracy of locating points was good enough to enable the task of automatic person identification.

**Discussion**

We have presented an image interpretation method, which can be used successfully for accurately locating facial features. Our approach to the problem is generic and can be used for other object recognition problems. In addition to locating facial features, the authors' implementation has been used successfully for detection of the endocardium and epicardium of the left ventricle of hearts in ultrasound images [12] and for locating hands in hand images [1] as part of a gesture recognition system.
The main disadvantage of the system presented is the long computational time. In the future we would like to use parallel processing techniques for the execution of GA search by assessing members of the population in parallel during selection.

The work described in this paper represents the initial steps toward a complete face image interpretation system. We have developed techniques which can perform successfully tasks such as face identification [9], low-bit rate coding [8], expression recognition, pose recovery and gender recognition [10], based on the automatic location of facial features on unseen images. The results obtained for these applications prove the potential of image search methods using GAs/PDMs.

Acknowledgements

We would like to thank all the members of the Dept. of Medical Biophysics University of Manchester, and especially Dr. Adrian Ratter for their help and advice. Also all those who provided the face images used in our experiments.

References