

An Object Location Strategy using Shape and Grey-level Models

David H. Cooper, Noel Bryson, Christopher J. Taylor.

Wolfson Image Analysis Unit
Department of Medical Biophysics
University of Manchester,
MANCHESTER M13 9PT

Any model-based image interpretation system must be capable of describing objects, whose appearance in real images can vary widely, in sufficient detail to ensure that robust location of objects is possible. The system must cope with circumstances where data is incomplete, for example when touching and occlusion occur. We argue that to achieve this it is necessary to describe grey-level properties as well as geometric ones, and their expected variations. This paper proposes an object description which combines an explicit shape model with models of expected grey-level boundary appearance together with a mechanism for evaluating image data for correspondences to the model. The results of applying the method to locating the boundaries of overlapping and touching objects in microscope images of metaphase chromosomes and man-made mineral fibres are presented.

This work forms part of a project which is directed towards producing a computer vision system which may be used by a person who is not familiar with image processing techniques to define tasks and produce solutions to 2-D image interpretation problems. The proposed system models image data and allows a user to describe an image processing task in an hierarchical manner in what is collectively termed a world model¹. The purpose of the hierarchy is to split the task and the data modelling into manageable units, to assist in reducing the combinatorial problems encountered in model-based image interpretation⁴. Execution of the task is synonymous with instantiation of the model from image data.

An essential part of such a system is a mechanism to match features in an image to elements of the model starting from primitives or cues extracted directly from the image data. The mechanism must cope with indistinct or missing data caused by touching, occlusion and noise. Encouraging results have been reported by Bolles, Faugeras, Horaud and others^{5 6}. In this work edge detection operators provide boundary evidence which is matched directly to rigid (wire-frame or CAD)

models. The applicability of these methods to domains other than simple industrial part recognition is, however, limited because the models employed only cope with rigid shapes and the edge detection operators are often unable to cope with the enormous variation in appearance which may be exhibited by objects in real images.

We argue that there is a requirement to re-evaluate image data whenever a particular hypothesis is being examined, to search for supporting evidence unique to that hypothesis. This paper describes a mechanism to provide such supporting evidence, applicable across a variety of domains. It can be used independently or within a knowledge-based image interpretation system and overcomes some of the limitations discussed. It is intended to cope with objects whose shape may vary and whose boundary characteristics may vary from object to object or within an object. The user requires minimal knowledge of the internal representation of shape or boundary models, but will create the models during a training process in which examples are shown to the system.

MODEL-DRIVEN SEARCH

The assumption made throughout this paper is that an initial choice of object model (hypothesis) and of location has been suggested by one or more cues or by some higher-level process which is seeking to establish or refine an interpretation. Possible mechanisms for generating such cues are discussed elsewhere in these proceedings².

We use an object description, which contains both boundary and symmetry (axis) information, to direct a search for further evidence to support the hypothesis that such an object is present in the image at the suggested place.

The search is effected by extracting grey-level profiles along straight lines cast in positions determined by the typical shape predicted by the shape description. Candidate boundary points are extracted from the profiles using a technique which matches each profile to one or more parameterised models of expected grey-level appearance across the boundary³. A candidate boundary is then computed from the set of points using a dynamic programming technique which takes into account data from the shape model and also from the boundary

This work has been funded by SERC and has been carried out as part of Alvey project MMI 093: "Techniques for User Programmable Image Processing"

grey-level models. We expect this process to be one of refinement and re-evaluation. The candidate boundary could well be used to generate a new hypothesis to which the same or some other object model could be applied.

Planting Search Lines

In order to plant the search lines the model must predict where the expected boundary lies. The shape model must therefore capture the distributions of positional variation along the boundary. We can choose to place the expected boundary at the mean of the distribution, or at a percentile if it is skewed and we want to bias towards the majority. Lines are placed normal to the expected boundary with their mid-points on the expected boundary. The sampling strategy used at the moment is a simple one; a minimum frequency sample suitable for location of straight boundary sections with increased frequency proportional to local curvature.

The accuracy of the placement of the initial hypothesis is unknown, and so the length of the search lines is related to the size of the object. On subsequent refinements however the length of each search line can be related to the local variation in the boundary position; shorter lines may be used where the variance is low. This reduces the amount of computation in general and the possibility of producing false positive evidence in particular.

Generation of Candidate Points.

Candidate boundary point generation and evaluation is based on the parameterised grey-level model scheme described by Woods et al³. Any number of models may be used, depending on the cases that are likely to be encountered in the images. Four models are used in the chromosome example, as shown in fig. 1. Each model consists of a fixed number of segments and has the boundary edge clearly identified. Grey-level model matching employs dynamic programming and uses parameters in the cost function which are computed from statistics derived during training.

Candidate boundary points are generated from each profile using an edge detection operator. In the examples shown we have used maxima in the smoothed first derivative; we intend to use an improved method described in these proceedings by Graham et al¹⁰. Each candidate point is labelled with a measure of how well the profile matches the expected boundary characteristics when the boundary is considered to be at the point under consideration. The process is as follows: each candidate point is considered in turn and associated with the appropriate edge in all of the grey-level boundary models; the two partitions generated are matched separately to their corresponding partitions in each model to produce a fit measure corresponding to that model. The smallest of these values, which corre-

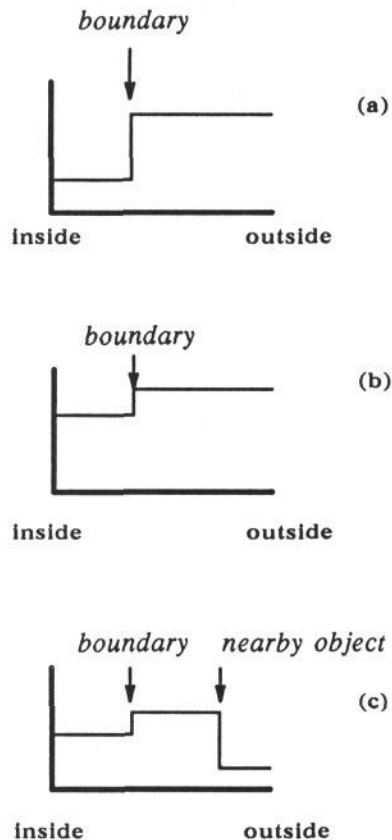


Figure 1: Boundary grey-level models used in the chromosome example: (a) Distinct boundary (b) faint boundary and (c) weak edge and a nearby object. The vertical axis represents grey-level. A strong edge/nearby object model is also used but not shown here. Strong and weak edges are generated by the presence or absence of banding, the term given to the density changes within the chromosome body. (See figure 4a.)

sponds to the best fit, is chosen as a label for the point. The labels for all the candidate points on a profile are normalised with respect to the smallest one.

A further set of *phantom* points are generated by propagating lines parallel to the predicted boundary from a subset of the previous search lines' candidate point set (points corresponding to the best three paths so far) and computing the intersection of these lines with the current search line. These points are labelled in the same way as genuine edge points. In the event that portions of the boundary are obscured for whatever reason the *phantom* points enable the missing sections to be spanned as dictated by the shape information. Experience has shown that, in such cases, edge points are usually generated by noise and do not, of course, correspond to object boundary points. For this reason *phantom* points are always generated.

Generation of the Candidate Boundary.

Associated with each search line we now have a set of labelled points. A dynamic programming technique is used to select which combination of points best represents the object boundary. The notion of best fit is a combination of grey-level profile correspondence to the

expected edge data and overall correspondence to the shape model. In the latter case candidate points nearest to the predicted boundary might be favoured; however this criterion alone can give rise to erratic boundaries [fig. 2a] because the choice of point on each search line is independent. If we favour paths which run parallel to the predicted boundary the correspondence to the model is greatly improved [fig. 2b]. Because the data and model can differ in size and location this criterion is applied most strongly on boundary sections of low curvature.

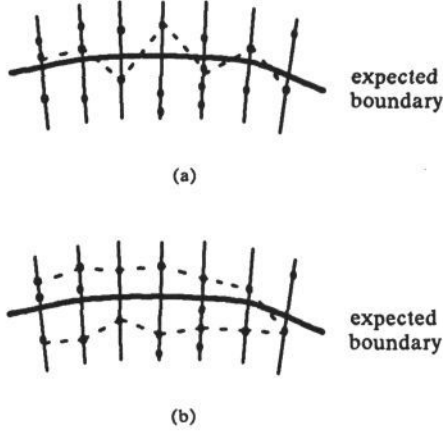


Figure 2: boundary matching: (a) boundary fit when points nearest to the predicted boundary are favoured and (b) when paths parallel to the predicted boundary are favoured.

We implement these ideas by forming a feature vector for each point. The component features are:

- 1) The label associated with each point (grey-level profile fit). This is the minimum cost associated with fitting to each of the grey-level models; it is normalised with respect to all of the costs of the points on the current search line, so that the best point has a cost of 1, all others having costs greater than 1.
- 2) Divergence of the path from parallel to the model prediction, weighted inversely by local absolute curvature. The curvature value is provided by the shape model and included in the label associated with the search line.
- 3) The distance between the candidate point and the predicted point. Using this feature favours the upper path depicted in figure 2b which is, overall, closer to the expected boundary.

Our objective is to find the path through the set of candidate points which is most likely to be the correct path. We have various forms of evidence and need to combine them in a principled way.

If we assume that the evidence from each point is independent, then $P(m = m_0)$, the probability of a path m

being the correct path m_0 , is just the probability that all of the points i are on the path m_0 and is given by

$$P(m = m_0) = \prod_{i=1}^N P(i \in m_0 | x_i)$$

where x_i is the evidence associated with the point i and N is the number of points on the path. Using Bayes' theorem we can write

$$P(i \in m_0 | x_i) = \frac{P(x_i | i \in m_0) P(i \in m_0)}{P(x_i)}$$

If we assume that $P(x_i | i \in m_0)$ and $P(x_i)$ are normally distributed and that $P(i \in m_0)$ is a constant i.e. that all points are equally likely to lie on m_0 we can write

$$P(i \in m_0 | x_i) = \prod_{j=1}^F K \frac{e^{-\frac{(x_i(j) - \mu_0(j))^2}{\sigma_0(j)^2}}}{e^{-\frac{(x_i(j) - \mu(j))^2}{\sigma(j)^2}}}$$

for some K and where

- j ranges over the elements of the feature vector
- μ_0 is the mean of x_i , $i \in m_0$
- σ_0 is the standard deviation of x_i , $i \in m_0$
- μ is the mean of x_i , for all i
- σ is the standard deviation of x_i , for all i

and therefore,

$$P(m = m_0) = \prod_{i=1}^N \prod_{j=1}^F K \frac{e^{-\frac{(x_i(j) - \mu_0(j))^2}{\sigma_0(j)^2}}}{e^{-\frac{(x_i(j) - \mu(j))^2}{\sigma(j)^2}}}$$

Taking logarithms and eliminating the constant K we seek the path for which

$$\sum_{i=1}^N \sum_{j=1}^F \left\{ \left(\frac{x_i(j) - \mu_0(j)}{\sigma_0(j)} \right)^2 - \left(\frac{x_i(j) - \mu(j)}{\sigma(j)} \right)^2 \right\}$$

is a minimum. This form of cost function can be minimised across all possible paths using dynamic programming.

The matching process starts at an arbitrary choice of search line/candidate point-set with initial costs of zero. For each candidate point on a search line each path from the previous line's candidate set into the current candidate is considered and the minimum cost path is recorded. After one circuit the start line/point set is reached; the matching process continues, however, until the pattern of cost paths stabilizes, thus overcoming the error introduced by not having correct cost values initially. This typically happens within two or three places after the start and improves performance compared to terminating at the start point.

Training

The parameters used in the grey-level and boundary matching cost functions are estimated by training on examples. For each object model the grey-level model parameters are trained first. Typically four or five search lines are drawn (interactively) across boundary sections which correspond to each of the grey-level models deemed necessary by the user. This number of samples is enough to give sensible initial values for these models to be used for edge matching. In each case a match is made and the position considered to be the boundary is displayed. Errors are corrected interactively and feature statistics are accumulated.

The boundary model parameters μ , μ_0 , σ , σ_0 are trained by initiating a boundary search on example objects. An axis cue is drawn interactively to instantiate an object model. A trial instantiation is generated using default initial values for the parameters and any errors in the generated boundary position are corrected by the user. Feature values corresponding to the correct boundary are used to update the statistics. At the time of writing data generated by the boundary search is not used to update the grey-level model statistics, and this is reflected in the examples shown below. We intend ultimately to use this data.

Generic models – Size and Shape.

In our chromosome example the world model defines a generic chromosome to be a specialisation of a system-defined prototype *blob* from which it inherits an appropriate shape model. Each specific chromosome is a further specialisation. (See figure 3.) The shape model for all chromosomes therefore uses the same representation; the generic version combines the distributions of all of the sub-types, and is less constrained.

Several authors have stressed the need to distinguish between size and shape ^{7 8 9}, but for the purposes of reconstructing boundaries this is not necessary. Indeed it can be a hindrance in generic models where the individuals may vary in size and shape, even though there are recognisable constraints in their spatial occupancy. We therefore record the shape data in un-normalised form for reconstruction purposes, but transform to a size-plus-normalised form whenever necessary, for example in providing goodness-of-fit measures independent of size for matching individuals or providing confidence labels for candidate points.

Choice of shape representation

The cue generators that we have used in the work reported here happen to generate symmetry and edge information. It is desirable to have similar entities expressed explicitly in our model representation in order to simplify the process of generating an initial choice of object model. The shape model that we use is a single axis symmetric ribbon model which is appropriate for

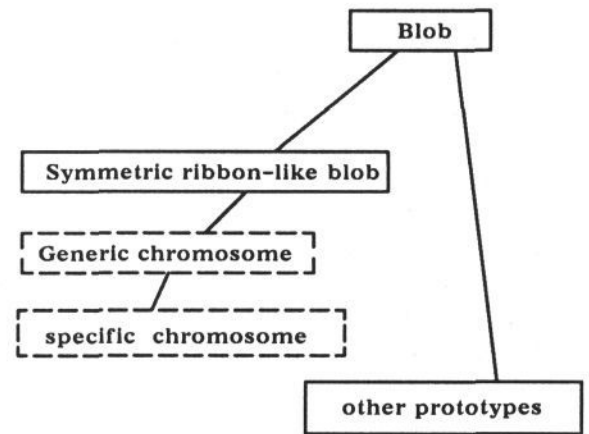


Figure 3: The world model inheritance hierarchy depicting the refinement of blob definitions. User-defined entities (dashed boxes) inherit a shape representation from the system prototypes (solid boxes). See Woods et al ¹ for a fuller description.

our exemplars and makes the mapping from the cues to the model a simple one. This does not detract from the principle of combining explicit shape models with grey-level descriptions or compromise the generality of the method.

Extending the capacity of the shape representation to more complex shapes is a topic currently under consideration. We require to infer shape and shape variability from a set of examples and unfortunately many shape representations described in the literature (Symmetric Axis Transform ^{11 12 15 16 17}, Smoothed Local Symmetries ^{13 14}, Local Rotational Symmetries ⁸) are either degenerate on certain classes of shape, are unstable, or currently require unacceptably large amounts of computation to achieve stability.

APPLICATION

We have applied the method described above to microscope images of chromosomes (Figures 4 and 5) and man-made mineral fibres (Figure 6). All the images used were 512 * 512 pixels with 6-bit grey-level resolution. The process of object instantiation was initiated using grey-level symmetry axis cues, generated automatically by the method described elsewhere in these proceedings by Thornham et al. ² with the exception that interactive editing was used to link across short (and topologically meaningful) gaps. We intend ultimately to apply an improved version of the *good continuation* strategy described by Dixon et al ²⁰ for asbestos fibre counting, to create these links.

Figure 4a shows a subset of the chromosomes of a human cell in metaphase. The grey-level variation within each chromosome (the banding pattern) is clearly visible. The lines overlaid in white mark positions of grey-level symmetry at a scale commensurate with the typical width of the chromosomes. External symmetries are shown as well as internal ones.

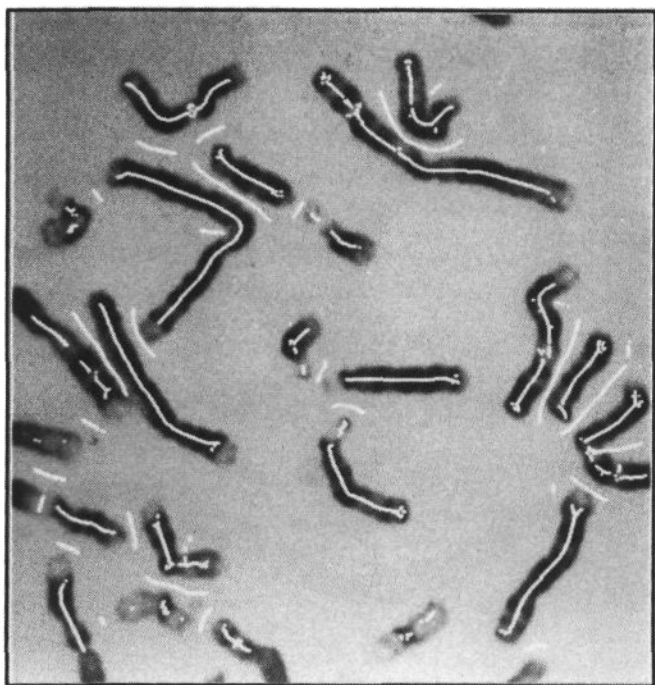


Figure 4a: Chromosomes showing characteristic banding patterns and with positions of grey-level symmetries overlaid.

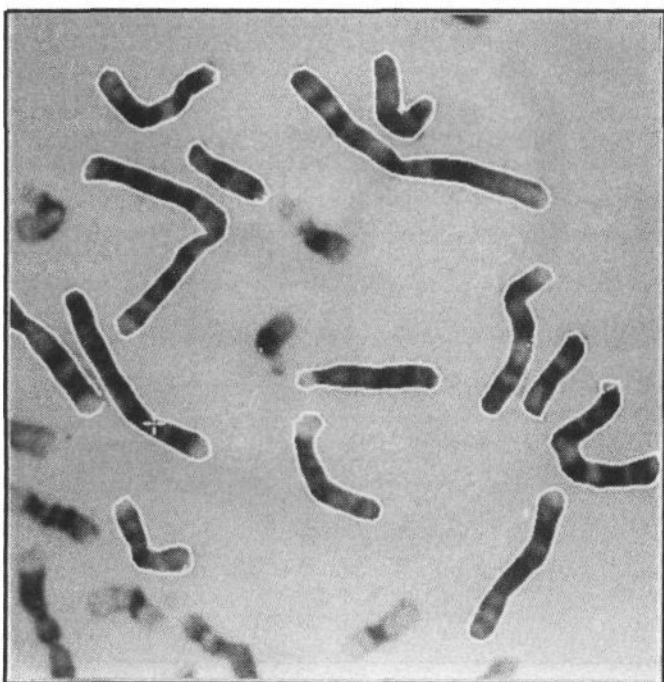


Figure 4b: Best boundaries after the training session.

Figure 4b shows the instantiated boundaries at the end of a training session. These boundaries represent the best that the system can achieve given the limitations of the edge detector and choice of models currently employed. Figure 4c shows the result of automatic boundary instantiation. The *nearby object* grey-level model is particularly useful in ensuring that the correct edge is chosen when objects are adjacent.

Figure 5 shows a series of images (taken at twice the magnification of those in figure 4) of crossing chromosomes. Figure 5a shows the output of the symmetry cue generator, and Figure 5b the best possible boundaries

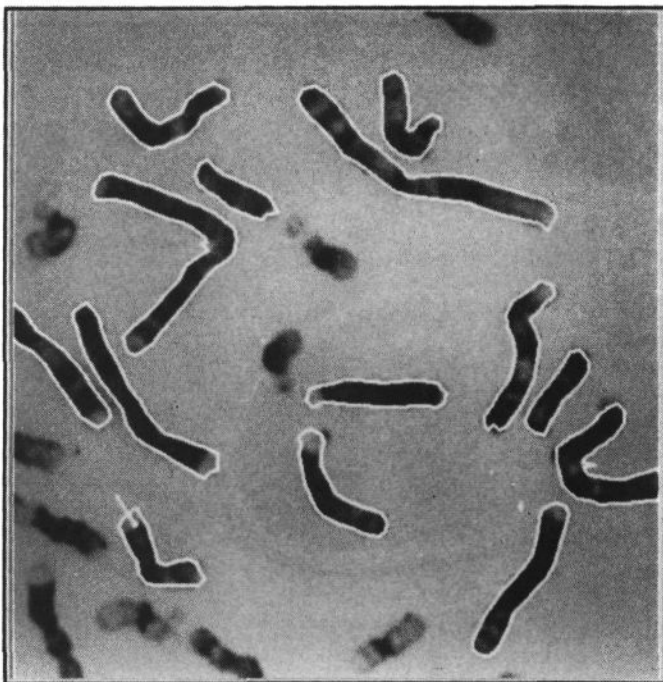


Figure 4c: Result of the search on isolated chromosomes.

after training. Figures 5c and 5d show the result of automatic boundary instantiation and demonstrate the effectiveness of the method in bridging gaps at the intersection.

Errors occur mainly at points of high curvature on the boundary where the parallel-to-model criterion is least effective. Further work is required to establish the precise effect of each feature in the cost function on the resultant boundary. The current implementation does not generate phantom points on the first search line, which happens to correspond to one end of the shape model; lack of candidate points may be the cause of failure on poorly defined ends.

Finally Figure 6 shows the technique applied to man-made mineral fibres. The complex boundary structure where two fibres are touching has been correctly interpreted, despite the crude edge detection methods currently employed. The same object model was used for all the fibres in this image. This sometimes gives rise to width errors which can be seen on the vertical fibre boundary. We expect that this will not occur when the scale information from the cue generator is applied to the production of the initial hypothesis generation.

DISCUSSION

The work described is a generalisation of the method reported by Azzopardi et al ¹⁸ for detecting the boundaries of muscle fibres. In that work an implicit circular shape model was used, candidate points were not assessed with respect to any global boundary description and ad-hoc rather than trained values were used in the cost function. Similar comments apply to work described by Pope et al ¹⁹ in which circular models and dynamic programming are used in the context of ven-



Figure 5a: Crossing chromosomes and corresponding symmetry cues.

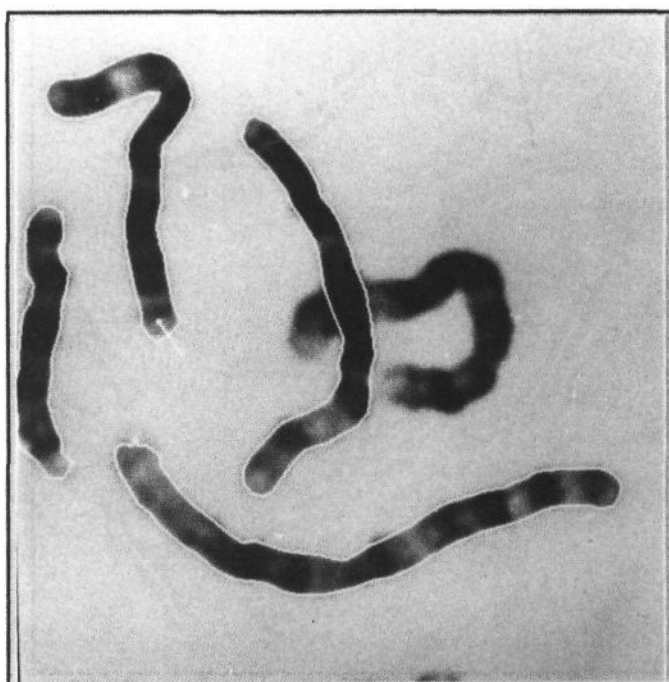


Figure 5c: Result of search, showing one of the crossing chromosome boundaries.

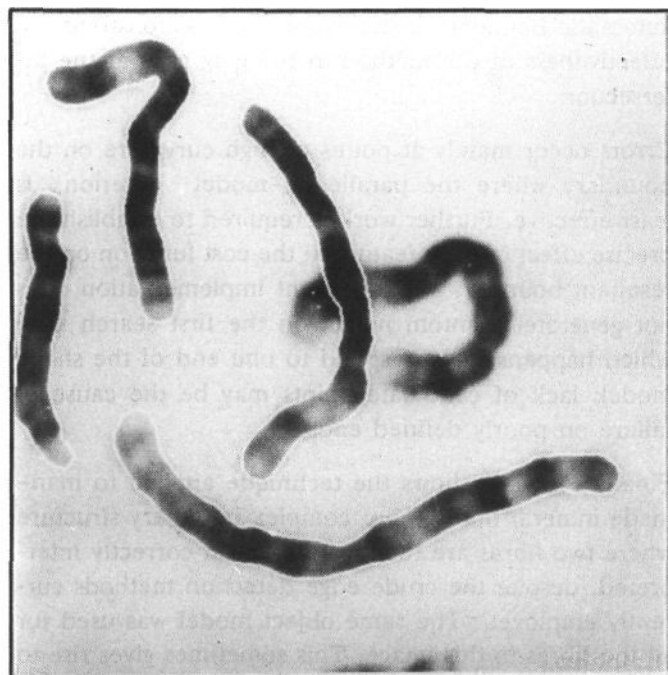


Figure 5b: Best boundaries after training.



Figure 5d: Result of the search, showing the second crossing chromosome boundary.

tricular border recognition. In that paper only *edge likelihood* is used as a matching feature.

The work described here offers significant improvements which are reflected in the success of its application, especially in cases where data is missing. In particular, the technique employed is principled and extendible. We believe that improvements in the generation of candidate points referred to in the text will result in the applicability of the method to images with a wider range of edge characteristics and also give increased robustness.

The key aspect of this work is the combination of explicit shape and grey-level information in an attempt to

model the appearance of objects and apply these models to locating instances in scenes. If we can succeed in developing this approach we will be able to predict the effect of cue generators and thus be able to generate hypotheses in a systematic way.

REFERENCES

1. Woods, P.W., Pycock, D., Taylor, C.J. "A Frame-based System for Modelling and executing Visual Tasks". *Proceedings of the Alvey Vision Conference 1988*.

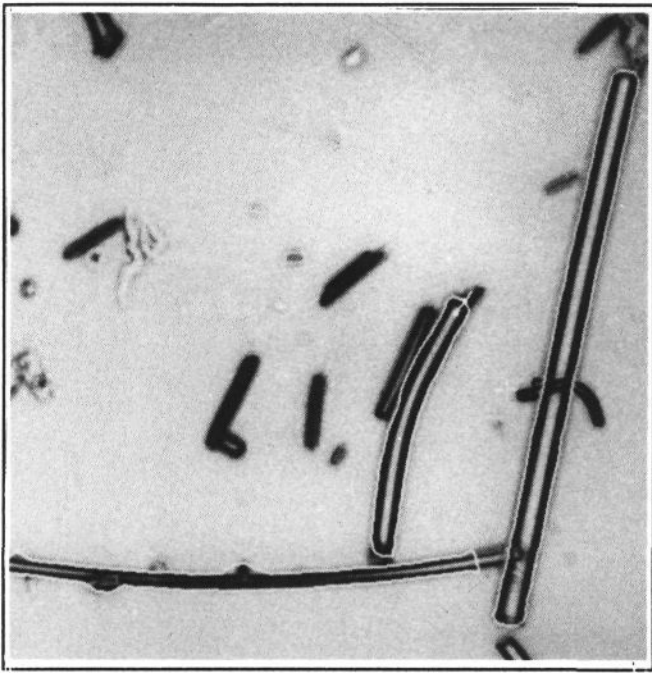


Figure 6: Man-made mineral fibres: Result of search showing instantiated boundaries.

2. Thornham, A., Taylor, C.J., Cooper, D.H. "Object Cues for Model-Based Image Interpretation." *Proceedings of the Alvey Vision Conference Manchester 1988*.
3. Woods, P.W., Taylor, C.J., Cooper, D.H., Dixon, R.N. "The use of geometric and grey-level models for industrial inspection." *Patt. Rec. Letters*. Vol. 5 No 1 (Jan 1987)
4. Taylor, C.J., Graham, J., Cooper, D.H. "System architectures for interactive knowledge-based image interpretation." *Phil. Trans. Royal Soc. London A* 324 pp. 457-465 (1988)
5. Horaud, R., Skordas, T. "Model-based strategy planning for recognising partially occluded parts." *IEEE Computer* Vol. 20 No 8 (Aug 1987)
6. Chin, R.T., Dyer, C.R. SURVEY: "Model-based recognition in robot vision." *ACM Computing Surveys* Vol. 18 No 1 (Mar 1986)
7. Pavlidis, T. SURVEY: "A review of algorithms for shape analysis." *Computer Graphics and Image processing* Vol.7 pp. 243-258 (1978)
8. Fleck, M.M. "Local rotational symmetries." *MIT AI technical report*. AI-TR-852. (1985)
9. Mokhtarian, F., Mackworth, A. "Scale-based description and recognition of planar curves and two dimensional shapes." *IEEE PAMI* Vol. 8 p 34-43 (1986)
10. Graham, J., Taylor, C.J. "Boundary Cue Operators for Model-based Image Processing". *Proceedings of the Alvey Vision Conference Manchester 1988*.
11. Blum, H. "Biological Shape and Visual Science (part 1)" *J. Theoretical Biology*, Vol. 38, pp. 205-287. (1973)
12. Blum, H., Nagel, R.N. "Shape description using Weighted Symmetric Axis Features" *Pattern Recognition* Vol 10. pp. 167-180. (1978)
13. Brady, M., Asada, H. "Smoothed local symmetries and their Implementation" *MIT AI Memo AIM-757*.(1984)
14. Connel, J.H. "Learning Shape Descriptions: Generating and Generalising Models of Visual Objects" *MIT AI tech. report AI-TR-853*. (1985)
15. Fairfield, J.R.C. "Segmenting blobs into Subregions" *IEEE Systems, Man & Cybernetics*" Vol. 13 pp. 363-384 (1983)
16. Ho, B-S., Dyer, C.R. "Shape Smoothing using Medial Axis Properties" *IEEE PAMI* Vol. 8 pp. 512-520 (1986)
17. Pizer, S.M., Oliver, W.R., Bloomberg, S.H. "Hierarchical Shape Description via the Multiresolution Symmetric Axis Transform" *IEEE PAMI* Vol.9 pp. 505-511 (1987)
18. Azzopardi, P.J., Pycocock, D., Taylor, C.J., Wareham, A.C. "An Experiment in Model-Based Boundary Detection" *Proceedings of the Alvey Vision Conference Manchester 1988*.
19. Pope, D.L., Parker, D.L., Clayton, P.D., Gustafson, D.E. "Left Ventricular Border Recognition Using a Dynamic Search algorithm" *Radiology* Vol. 155 pp. 513-518. (1985)
20. R.N. Dixon. "Aspects of Picture Processing: Hardware and Software for General-Purpose Image Analysis." *Ph.D. Thesis*: University of Manchester (1980)

