Shape Verification using Belief Updating

C. J. Taylor and D. H. Cooper

Department of Medical Biophysics
University of Manchester
Oxford Rd
Manchester M13 9PT

This paper addresses the problem of 2D shape representation and its application to object verification. We show how knowledge of shape can be integrated in a principled manner with low–level evidence such as an estimate of object position and an edge strength map. We begin by considering the role of shape in image interpretation and discuss the criteria which should be applied in assessing representations of shape. We propose new criteria, particularly as regards the ability to model variability, and describe a Chord Length Distribution (CLD) representation of shape which possesses many desirable properties. We show how the CLD representation can be used in an iterative belief-updating scheme for object location and verification. We give experimental results which demonstrate the feasibility of the method and discuss future developments.

Shape information is often provided as part of the task definition for a vision system. Objects and structures which may be present in the visual input are described in such a manner that they can be automatically located and identified. In the simplest case, the objects and structures of interest are of fixed appearance and independent of each other. More commonly they are of variable appearance and subject to mechanical interaction. Any general scheme for representing and applying knowledge of shape must address these problems.

The simplest method of using shape information is to apply image segmentation methods to extract, from the visual input, candidate objects which can be tested for similarity to the expected objects and structures [1]. Although this approach has been used extensively it is rarely practical with real images [2].

A more realistic approach is to use shape information to organise low–level evidence and generate plausible hypotheses for objects which may be present. There are, for example, numerous variations on the idea of extracting simple primitives, typically edge segments, and applying knowledge of shape to conduct an efficient search through the space of all possible matches between target objects and combinations of these primitives [3,4,5]. The limitations of this approach are the computational cost of exploring plausible combinations of primitives, the limitation to reasonably rigid shapes and, above all, the fact that the globally best solution may not involve the locally most plausible primitives.

A more general framework for combining low–level evidence and high–level knowledge of shape (and other properties) is provided by the Cue-Hypothesize-Verify paradigm. In this approach, data–driven processing is used to generate cues which constrain, to a greater or lesser extent, the possible instances of objects which may be present in the input. These constraints are used in combination with shape information to generate object hypotheses which are refined and verified by collecting additional low-level evidence [6,7]. Methods of verification often involve search, amongst the constrained set of possible solutions, by methods such as dynamic programming [8,9,10,11] and gradient descent [12]. One of the main disadvantages of these approaches is the difficulty of applying other than very local shape knowledge during the refinement and verification process.

The method described in this paper is intended for use within a Cue-Hypothesize-Verify framework. We assume that some estimate is available for the position of an object and that there is low-level evidence such as an edge-strength map. The objective is to find the most plausible instance of a given shape (or family of shapes) and an estimate of the confidence which can be placed in the solution. We confine our attention to 2D shape though the method we describe could be extended to 3D.

SHAPE REPRESENTATION

A number of authors have proposed criteria against which any shape representation may be assessed [13,14]. These criteria include *computability, scope, information preservation, uniqueness, sensitivity, proportionality and local support*. We endorse the utility of most of these criteria but argue that *local support* is not a requirement in the context of a Cue–Hypothesize–Verify strategy, since there is no necessity to extract a representation of shape from each input image. There is, however, a closely related point which is that *global constraints* should be capable of *local application* if the representation is to be used effectively to control the process of refinement and verification. We also propose that important additional criteria are *variability encoding*, the capacity to capture variability within families of shapes, and *support for probabilistic reasoning*. The first of these criteria is closely related to *proportionality* and is important because, even in images of man-made objects and particularly in the natural world, we rarely deal with shapes which are entirely rigid. Given this, it is important to distinguish those changes in shape which are allowed from those which are not. The second idea, that methods of dealing with shape should provide support for a rigourous scheme of probabalistic reasoning, is axiomatic to the Cue-Hypothesise-Verify paradigm.

THE CLD REPRESENTATION

We propose a representation of shape which we call the CLD (Chord Length Distribution). It is extremely simple but, together with the method of verification described below, scores well against all the criteria listed in the previous section. A shape is first defined by a set of n points $x_1 .. x_n$. These may be equally spaced around the boundary but this is not necessary and it may be the case that, for a given value of n, an unequally spaced set of points may provide a more stable description, particularly for man-made objects. The only requirement is that there is a consistent method of selecting the points when the shape or family of shapes is defined. A reference point $x_\bullet$ is also defined for the object. The shape representation consists of the set of probability distributions $P(r_{ij}) : i, j = 0 .. n, i \neq j$ for the distances r_{ij} between all pairs of points x_i, x_j. The arrangement is illustrated in Figure. 1.
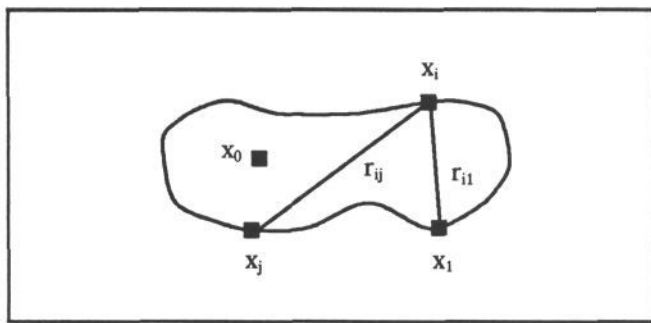

Figure 1: Geometry of the CLD representation.

The probability distributions can be estimated from a set of example images in which the correct locations of the shape-defining points have been established independently, usually via an interactive training procedure. When the objects of interest are rigid, all the $P(r_{ij})$ will have low variance and the shape will be highly constrained. When the objects of interest are variable, some, though generally not all, of the $P(r_{ij})$ will have high variance and some aspects of the shape will be less constrained. For example, suppose a description of cooked spaghetti [15] were created by training using equally spaced shape-defining points along the two sides; the distances between diametrically opposite points would be well-constrained because of the constant width of the strands, whereas the distances between other points would be less well constrained because of the variable length and curvature. It is also interesting to note how the CLD representation deals with scale; the distances between points which are near-neighbours on the boundary capture fine-scale characteristics whilst those associated with well-separated points capture coarse-scale information.

The representation is easily *computable* from examples of the shape though not directly from an image (but see below). It has the *scope* to deal with both rigid man-made shapes and variable natural objects. It is *information preserving* and *unique* except with respect to mirror symmetry. The *sensitivity* to small changes is limited only by the choice of n and the representation exhibits *proportional* behaviour as shapes are distorted. The statistical nature of the model allows natural *variability encoding*. We demonstrate below how the scheme can

support probabalistic reasoning and *local application of global constraints*.

SHAPE VERIFICATION

We now consider the problem of using a given CLD shape representation, s, to refine and verify an object hypothesis. Assume we are given a constraint, c, on the position of x_0 in the form of a probability distribution $P(x_0)$. Assume also that we are given edge evidence, e, in the form of a probability distribution $P_e(x)$, derived from an edge-strength map.

The key to our method is to store a probability map $P(x_i)$ for each of the n points which define the shape. Initial estimates $P(x_i|c)$ for these maps, given c, can be obtained from $P(x_0)$ using $P(r_{0i})$. If we consider a particular point x in the x_0 map the corresponding probability map for x_i is given by:

$$P(x_0 = x)P(x_i|x_0 = x)$$

where $P(x_i|x_0 = x)$ is the annular probability function obtained by rotating $P(r_{0i})$ about x. That is, the probability of point i being located at x_i, based on the possible location of x_0 at x, is the product of the probability that x is actually the location of x_0 with the probability that points x_0 and x_i will be found a distance (x_i-x) apart. Summing over all (mutually exclusive) possible values for x we obtain simply:

$$P(x_i|c) = P(x_0) * P(x_i|x_0 = 0)$$

where * denotes convolution.

If we make the reasonable assumption that the positional constraint, or cue, is independent of the edge evidence we can combine the two sources of information to give an initial estimate for each of the probability maps, given c and e:

$$P(x_i|c, e) = P(x_i|c)P_e(x) = P^0(x_i|c, e, s)$$

In general these maps will not tightly constrain the positions of the x_i but neither will all the possible solutions they represent be consistent with the given CLD. We can update the maps to improve consistency with the CLD by choosing each map in turn as the target and using each of the other maps, in combination with the appropriate $P(r_{ij})$, to generate an independent estimate of the target map, in a manner analogous to that described above. The updated target map is simply the product of these independent estimates with the current map:

$$P^{k+1}(x_i|c, e, s) = P^k(x_i|c, e, s)$$
$$\Pi_{j \neq i}[P^k(x_j|c, e, s) * P(x_i|x_j = 0)]$$

This belief updating process is repeated until a stable, maximally consistent interpretation is reached.

Once a set of stable probability maps have been generated the point locations which constitute a solution can be extracted as follows:

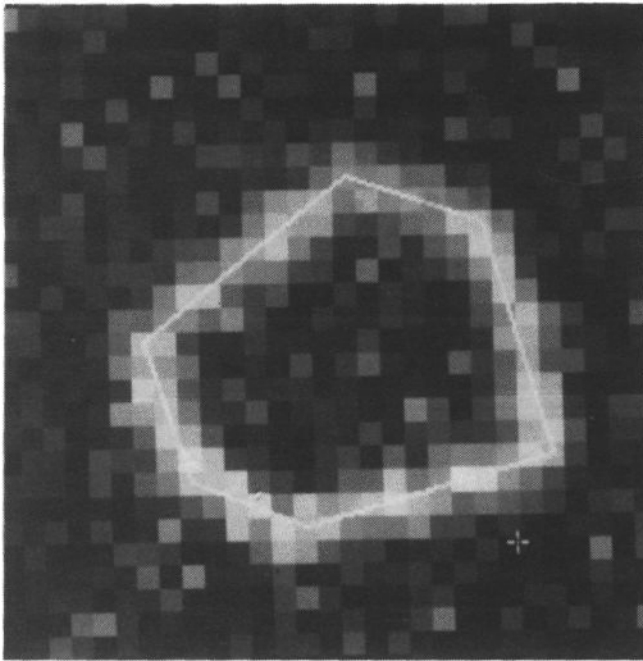(a) choose the highest peak in any of the probability maps,

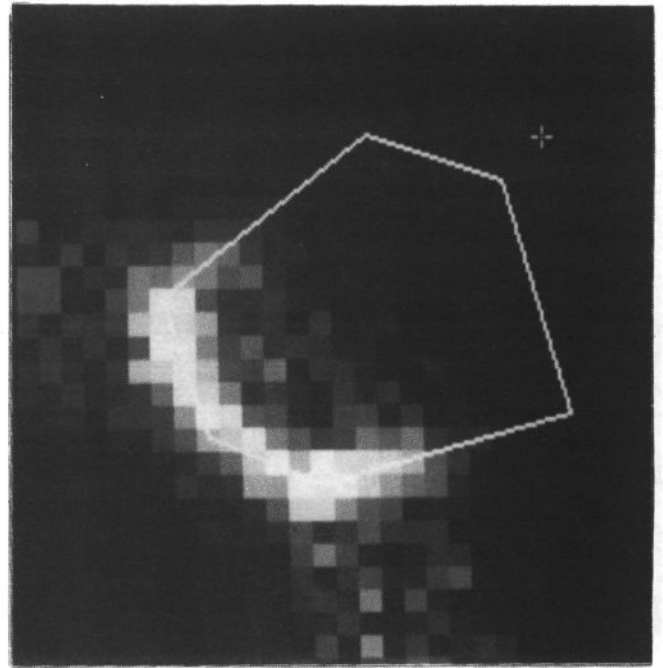*Figure 2(a): $P_e(x)$, a poorly localised, noisy edge map.*

Figure 2(b): $P^0(x_i)$, the initial estimate for one of the point probability maps.
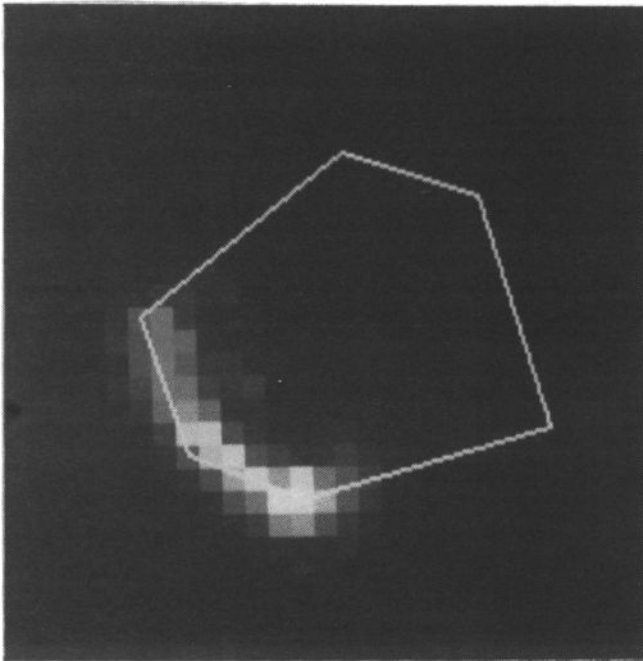

Figure 2(c): $P^1(x_i)$, the map for the same point as Fig, 2(b) after 1 iteration.
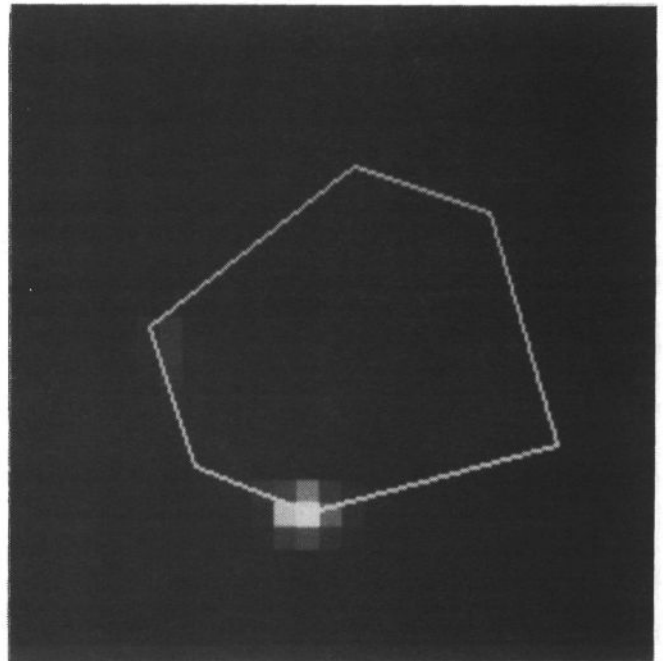

Figure 2(d): $P^2(x_i)$, the map for the same point as Fig 2(b) after 2 iterations.

(b) select that point as part of the solution by setting the probability at all other locations in that map to zero,

(c) iteratively update all maps until a new stable solution is reached,

(d) repeat (a)–(c) for maps in which the solution point has not yet been selected, until a complete solution is obtained.

The belief we attach to the solution is given by the product of the probabilities of the selected points. In practice, a good approximation to the correct solution may be obtained without applying step (c) if the initial convergence of the probability maps results in well–localised peaks. In the special case that the shape to be verified is symmetric, there will be more than one stable

solution and the probability map for each point will contain more than one peak. Under these circumstances it is always necessary to apply step (c) at least once.

EXPERIMENTAL RESULTS

We have carried out a series of experiments to demonstrate the feasibility of the method described above, using both synthetic and real image data.

The experiments with synthetic images use irregular polygons with the shape definition points placed at the vertices. We have simulated uncertainty in the positional constraint, limited flexibility of the shape and noisy,
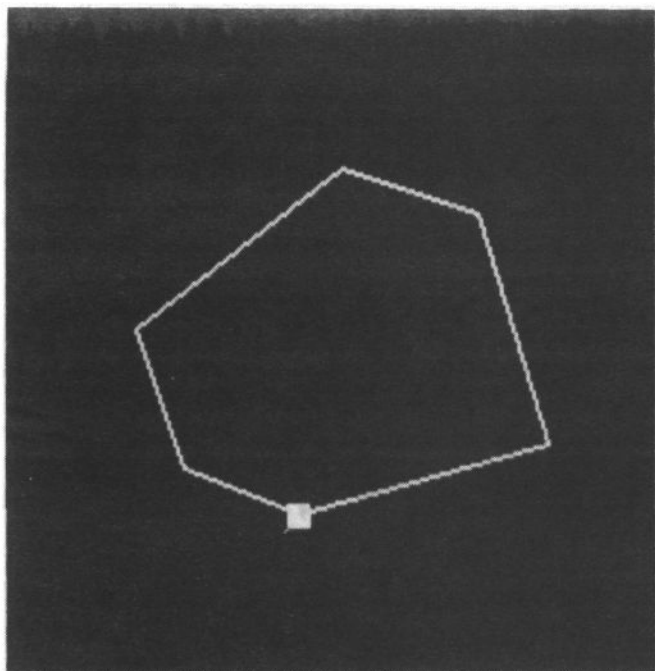
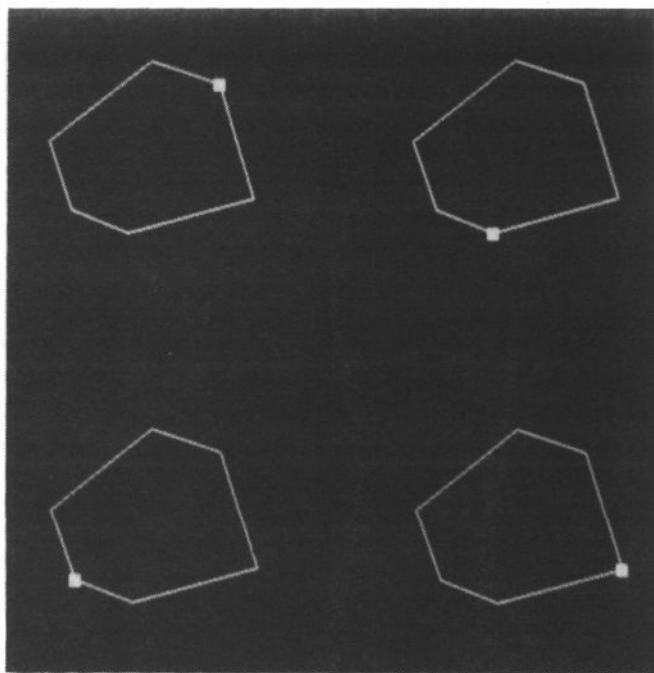*Figure 2(e): $P^4(x_i)$, the result after 4 iterations.*

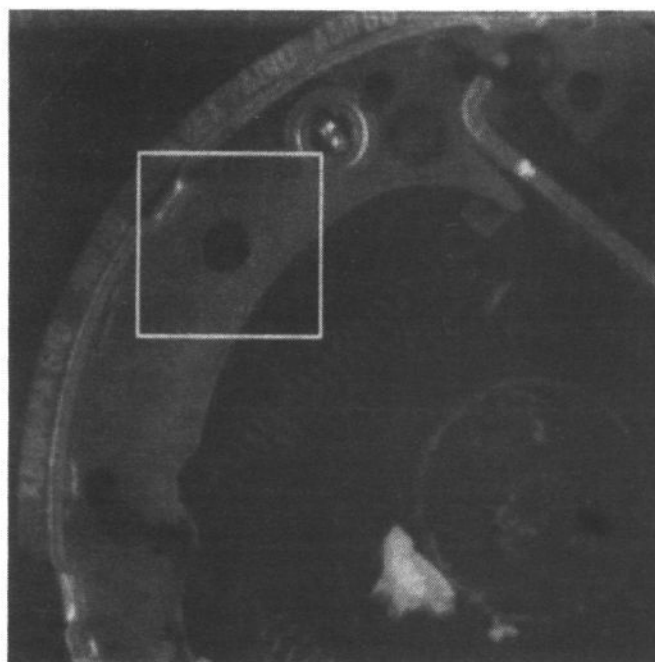*Figure 2(f): Results after 4 iterations on four of the points x_i.*

Figure 3(a): A brake assembly image and region of interest.
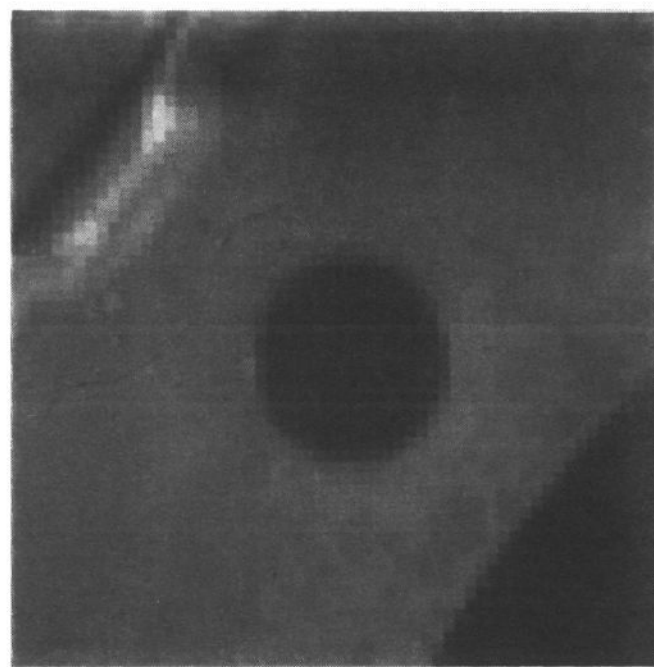

Figure 3(b): 64^2 region of interest.

non–localised edge data. Figure 2 illustrates various stages in the process of shape verification and refinement. Figure 2(a) shows $P_e(x)$, a poorly localised, noisy edge map. The object location cue is at the image origin and only loosely constrains the location of the object. Figure 2(b) shows $P^0(x_i|c, e, s)$, the initial estimate for one of the point probability maps, with the outline of the object from which the edge-strength image was generated superimposed. Figures 2(c), 2(d) and 2(e) show the same map after 1, 2 and 4 iterations respectively. The process can be seen to converge rapidly to a good estimate of the position of one of the vertices of the polygon. Figure 2(f) shows similar results for several other maps.

We have investigated the behaviour of the method by varying the parameters of the simulation and have found

that the correct solution is obtained over a broad range of operating conditions. Convergence rarely takes more than a few iterations and robustness is improved, as would be expected, by increasing the number of points used in the shape representation.

Figure 3 shows results obtained using data from a practical industrial inspection task. A training set of 10 brake assembly images was obtained (Figure 3(a)) and a square region of interest was selected at the same absolute position in each image (Figure 3(b)). The feature selected for testing was a punched hole in the brake shoe mounting plate, which we described by 8 shape–defining points. The position of the hole within the region of interest was variable. An edge-strength image was created for each image using a morphological edge detector [16, 17]
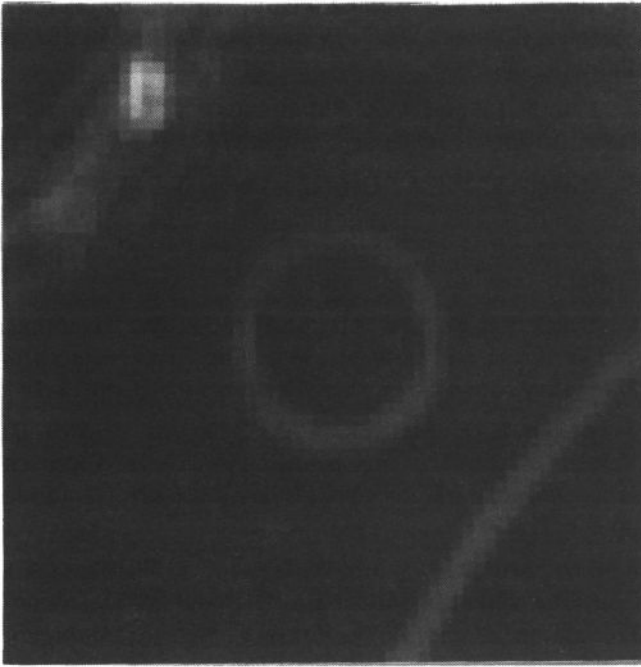
64

Figure 3(c): Morphological "edge strength" image.

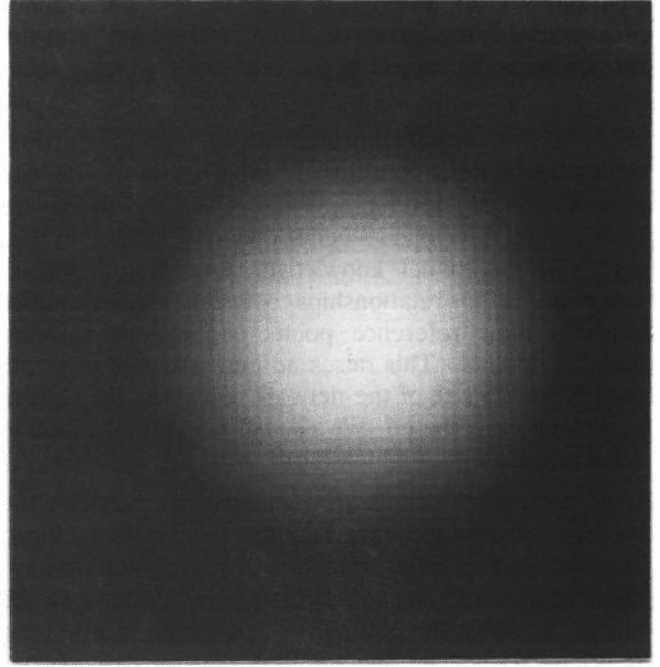

Figure 3(d): $P(x_0)$, the centre-of-hole distribution.

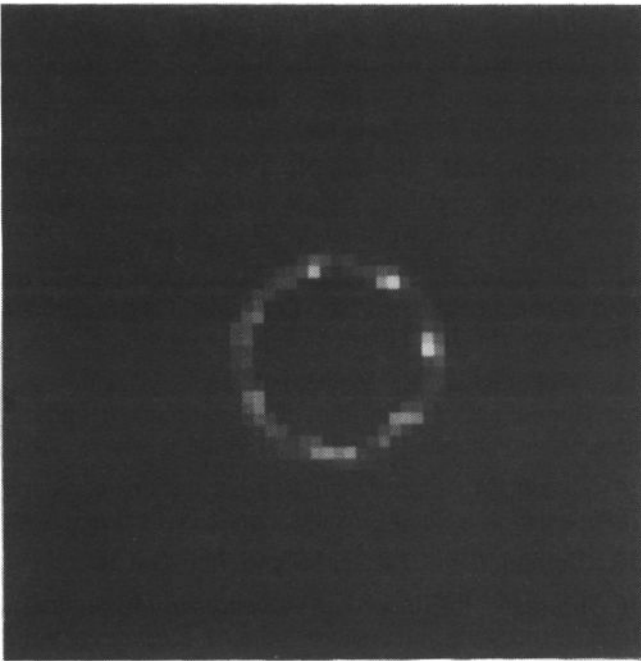
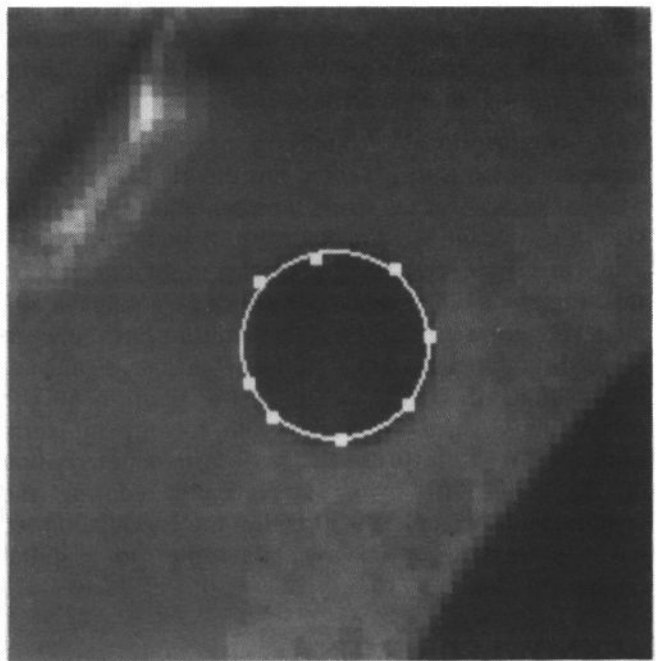Figure 3(e): $P^3(x_i)$, a point probability map after 3 iterations.

Figure 3(f): The most probable points superimposed on the original image.

(Figure 3(c)). The system was trained on a leave-one-out basis to obtain statistics for x_0, r_{ij} and for both on- and off-edge responses of the edge detector. The true location of the hole was obtained, for each example image, using an interactive procedure in which an existing automatic method was used to suggest a solution which could be modified, if necessary, by the user. The edge response statistics were used to transform the edge-strength images into edge probability images. Figure 3(d) shows $P(x_0)$, the distribution of locations for the centre of the hole and demonstrates considerable variability. Figure 3(e) shows the probability map for one of the shape-defining points after 3 iterations. As expected, the symmetry of the hole leads to several plausible solutions. The single largest peak in $P(x_1)$ was selected as the correct solution and belief updating was continued until a new stable solution was reached. $P(x_2)$ was localised similarly. The final result is shown in Figure 3(f) in which the most probable point in each $P(x_i)$ is shown superimposed on the original image. Most of the points have been correctly located though one or two have been shifted slightly, probably because the full solution extraction method was not used.

DISCUSSION

The CLD representation has the potential to form the basis of a unified approach to the integration of high-level and low-level evidence within a Cue-Hypothesize-Verify framework. Our initial experiments have demonstrated the feasibility of the approach over a range of operating conditions. We do not yet understand what conditions must hold for the method to reach a stable, plausible

solution but an attempt to analyse the convergence properties of the method, using a Bayesian network formalism, is the subject of current work.

The scheme reported here can be extended, reasonably easily, to deal with scaling and with cues which constrain orientation as well as position. It would also be desirable to incorporate grey-level properties of the chords but we have not yet developed a belief updating scheme to deal with such additional knowledge. Another important extension is to relationships between shapes, either between their reference points or between selected boundary points. This raises no new issues of principle though, as the size of the network increases, the control problem is exacerbated. We also believe that extension to 3D is reasonably straightforward.

It is interesting to note some similarities between the method we describe here and the Generalised Hough Transform [18]. The two forms of representation have elements in common and there are parallels between the manner in which the initial evidence is used to vote for a solution. It is tempting to think of our method as a generalisation of the Generalised Hough Transform. The idea of using pairwise relationships between model elements has similarities to the work of Grimson and Lozano-Perez[5,19]. Their approach assumes, however, that suitable primitives can be extracted from the visual input, prior to model-matching,

The computational complexity of the current implementation is n^2m^4 where n is the number of points used as the basis for the shape representation and m is the side of the image window known to contain the object. There are, however, a number of ways of improving upon this. First n^2 could be replaced by $n(\log n)$, noting that the model is highly redundant and that omitting links between some distant neighbours is unlikely to lose significant information. Secondly, we currently perform all the convolutions which are responsible for the m^4 term explicitly. We intend to use a closed form parameterisation of the probability maps, dramatically reducing the computational complexity. Finally, the formulation we have presented is clearly amenable to parallel implementation.

ACKNOWLEDGEMENT

REFERENCES

1. **Pavlidis, T.** "A Review of Algorithms for Shape Analysis" *Computer Graphics and Image Processing*. Vol 7 pp 243–258 (1978).

2. **Marr, D**. *Vision*. W.H Freeman, San Francisco (1982).

3. **Bolles, R. C.** "Robust Feature Matching Through Maximal Cliques" *SPIE, Bellingham, Wash.* Vol 182 pp 140–149. (1979).

4. **Chin, R.T., Dyer, C.R.** "Model-Based Recognition in Robot Vision" *Computing Surveys* Vol 18 No 1 (1986)

5. **Grimson, W.E.L., Lonzano-Perez, T.** "Model-Based Recognition and Localisation from Sparse Range or Tactile Data" *Int. J. Robotics Research* Vol 3 No 3 pp 3–35 (1984).

6. **Sykes, F.P., Pollard, S.B., Mayhew, J.E.W.** "Hypothesis and Verification in 3D Model Matching." *Proceedings of the Fifth Alvey Vision Conference*, Reading (1989), pp 13–18

7. **Baker, K.D., Sullivan, G.D.** "The Alvey MMI 007 Vehicle Exemplar: The Knowledge Based Approach" *Proceedings of the Third Alvey Vision Conference*, Cambridge pp 1–4. (1987)

8. **Azzopardi, P.J., Pycock, D., Taylor, C.J., Wareham, A.C.** "An Experiment in Model-Based Boundary Detection" *Proceedings of the Alvey Vision Conference Manchester (1988)*.

9. **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)

10. **Gorman, J.W., Mitchell, R., Kuhl, F.P.** "Partial Shape Recognition Using Dynamic Programming" *IEEE PAMI* Vol 10 No 2 pp 257–266 (1988)

11. **Cooper, D.H., Bryson, N., Taylor, C.J.** "An Object Location Strategy using Shape and Grey-level Models" *Image and Vision Computing* Vol 7 No 1 pp 50–56 (1989).

12. **Kass, M et al.** "Snakes: Active Contour Models" *Proc. First International Conf. on Computer Vision*, London. (IEEE) pp 259–268 (1987)

13. **Mokhtarian, F., Mackworth, A.** "Scale-based description and recognition of planar curves and two dimensional shapes." *IEEE PAMI* Vol. 8 p 34–43 (1986)

14. **Brady, M.** "Criteria for Representations of Shape" *Human and Machine Vision*. Academic Press (1983).

15. **Buitoni**. *Pasta D'Italia*, Milano, Italy (est 1827).

16. **Maragos, P.** "Tutorial on Advances in Morphological Image Processing and Analysis" *Optical Engineering* Vol 28 No 7 pp 623–632 (1987)

17. **Goetcherian, V.** From Binary to Grey Tone Image Processing using Fuzzy Logic Concepts. *Pattern Recognition* Vol 12 pp 7–15 (1979).

18. **Ballard, D., Brown, C.** *Computer Vision*. Prentice Hall (1982).

19. **Grimson, W.E.L., Lonzano-Perez, T.** "Localising Overlapping Parts by Searching the Interpretation Tree" *IEEE PAMI* Vol 9 No 4 pp 469–482 (1987).