

Locating Overlapping Flexible Shapes Using Geometrical Constraints

David H. Cooper, Christopher J. Taylor, Jim Graham, Tim F. Cootes.

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

Abstract

In an earlier paper [1] we have proposed a shape representation called the CLD (Chord Length Distribution) which possesses many of the often-quoted desirable properties of a shape representation. It also captures shape variability and complements an object location method using belief updating which integrates low-level evidence and shape constraints. Promising results on synthetic and real rigid objects were given. This paper describes a development to the original definition which makes the location method robust with respect to clutter. We give experimental results which demonstrate the performance of the revised scheme on a class of flexible shapes, both singly and overlapping.

We are currently engaged in a research project [see acknowledgements] concerned with automated 2-D inspection of complex (industrial) assemblies. In common with many machine vision applications we seek to exploit object shape and other geometrical constraints to assist in locating objects in scenes and evaluating interpretations with respect to expected appearance. To this end we need suitable representations for shape (intra-object) and inter-object relationships together with location and verification schemes capable of exploiting such representations. Ideally we seek a scheme capable of addressing both shape and inter-object relationships in a uniform manner.

We have argued [1] that a shape representation not only needs to satisfy often-quoted [2,3] properties of being easily computable, unique, and exhibiting proportional behaviour, but must also describe expected variability and invariance within a class of shapes and be capable of describing a wide range of shape classes. We have proposed such a representation called a Chord Length Distribution (CLD) and an associated object location scheme which exploits and integrates geometrical (shape) constraints with low-level (edge) evidence in a principled way, originally based on ideas derived from probabilistic reasoning using networks [4].

Unlike many reported methods of applying shape models [5,6,7,8] our approach *does not work by matching image primitives to related model elements*. Rather, it seeks to label each point in an ordinate space with a likelihood of correspondence to the

model. This likelihood is maximised with respect to the image evidence (edge data) and the shape constraints in the model. The advantage of this approach is the late commitment to an interpretation – the highest level primitive used is the pixel. This is particularly important in the context of overlapping or occluded objects.

This paper presents further developments and investigations into the properties of the CLD but first we give a brief outline of the CLD and the object location method. The reader is referred to [1] for a detailed description.

1 CLD and OBJECT LOCATION

A shape is first defined by a set of n points $x_1 \dots x_n$. 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_0 is also defined for the object. The shape representation consists of the set of probability distributions $P(r_{ij}) : i, j = 0 \dots n, i \neq j$ for the distances r_{ij} between all pairs of points x_i, x_j . The arrangement is illustrated in Fig. 1.

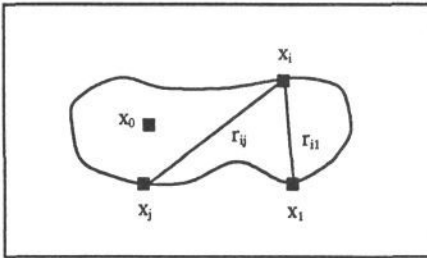


Fig. 1: Geometry of the CLD representation.

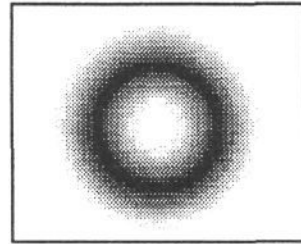


Fig. 2: $P(x_j | x_i)$

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. Various other properties are discussed in [1] but the only one of relevance here is that the representation is unique except with respect to mirror symmetry.

Object location depends on the fact that the radial distributions $P(r_{ij})$ allow us to predict where x_j is given the position of x_i by rotating the radial distribution about the origin $x_i = 0$ as shown in Fig. 2.

The key to our method is to store a probability map $P(x_i)$ for each of the n points which define the shape. Each location in the map is labelled with a likelihood of finding x_i there. We can compute a prediction for x_j at all points by correlating $P(x_i)$

with $P(x_i|x_j)$. For each x_i in turn we compute $(n-1)$ predictions for $P(x_i)$ from each of the other x_j and combine them with the original x_i to produce new estimates for the locations of each x_i . This belief updating process is repeated until a stable, maximally consistent interpretation is reached.

The initial values of the maps are generated by combining predictions made from the expected (prior) position of the reference point x_0 and edge data obtained from the image.

1.1 Behaviour With Clutter

The scheme outlined above is very successful in locating single instances of an object in a field in the presence of noise [1]. However, the method can sometimes converge to an incorrect result for multiple objects in the circumstance where the distance between the objects is comparable with or less than the chord lengths of the objects. It is easy to see how this can arise.

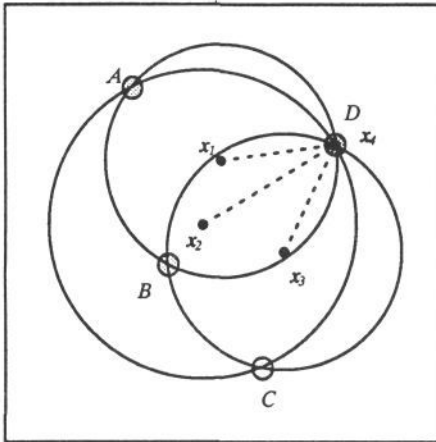


Fig 3: predictions using chords only

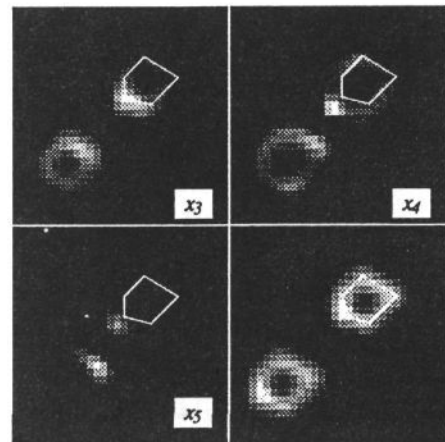


Fig 4: maps after 4 iterations for two similar polygonal objects (a) top = x_3 , x_4 (b) bottom = x_5 , original map

Fig. 3 depicts 3 points x_1 , x_2 , x_3 at known positions. The circles represent the predictions for a 4th point x_4 , given x_1 , x_2 , x_3 . In this case the predictions combine in a fashion analogous to a voting scheme as used by Hough transforms [9]. In our case, the belief in the location of x_4 is also weighted by the edge evidence, which may be stronger at A, B, C than at D resulting in incorrect convergence of the updating scheme. Fig. 4 shows an example with two similar polygonal objects whose vertices are labelled x_1 to x_5 clockwise from the bottom. Only one polygon icon is drawn. The initial maps are bottom right. The maxima in the top left, top right and bottom left diagrams should correspond to vertices x_3 , x_4 and x_5 but clearly do not. (compare this with Fig. 7).

This problem can be overcome by developing the CLD to remove the reflectional symmetry ambiguity.

1.2 The Revised CLD Representation

The modification is illustrated in Fig. 5. We have introduced angles θ_{ij} which describe the the angle that the normal to the boundary at x_i has to be rotated anti-clockwise to indicate the direction to x_j . The choice of object-related direction is arbitrary – in fact we use the direction of the image gradient at x_i in our experiments. We record the distributions $P(\theta_{ij})$ as part of the model.

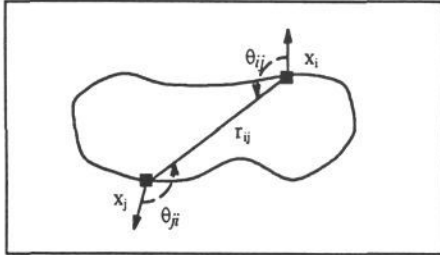


Fig. 5: CLD with direction information

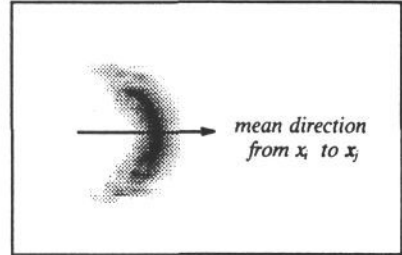


Fig. 6: $P(x_j | x_i)$

The conditional probability maps $P(x_i | x_j)$ become reduced annuli as depicted in Fig. 6, where the angular dispersion is determined by the variance in θ_{ij} . These new maps produce far more constrained predictions and result in faster and more stable convergence.

2 EXPERIMENTAL RESULTS

2.1 Nearby Objects

It is easy to see that the situation depicted in Fig. 3 is far less likely to occur when conditional maps incorporating angle statistics are used. The results for the same polygon pair as in Fig. 4 are shown in Fig. 7.

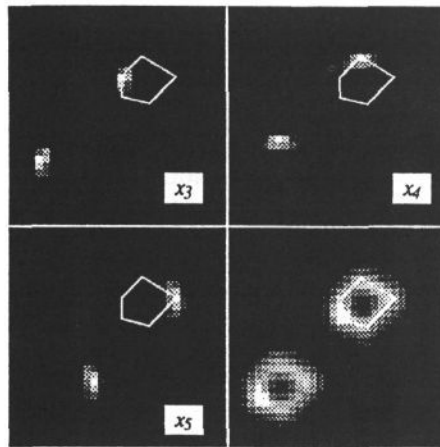


Fig 7: maps after 4 iterations for two similar polygonal objects (a) top = x_3 , x_4 (b) bottom = x_5 , original map (revised scheme)

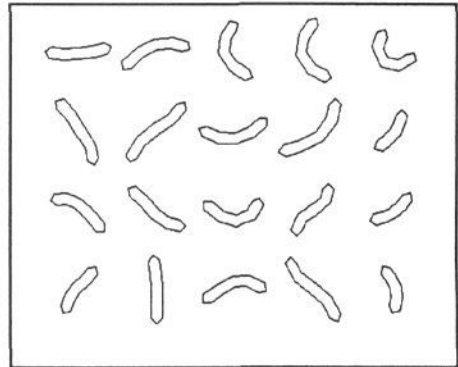


Fig 8: a family (can) of worms

In this case the local maxima in the maps correspond closely to the polygon vertices. There is one maximum for each polygon vertex. Note also the improvement in the rate of convergence using the revised scheme – the maxima are much better localised.

2.2 Flexible Objects

We wish to demonstrate our claim that we can locate objects whose expected shape is allowed to vary. To this end we have generated a set of axially symmetric ribbons (worms) whose axes can bend and be of different lengths but whose widths are fixed. Twenty examples taken from this set are shown diagrammatically in Fig. 8 to indicate the kind of variation present. The CLD in this experiment uses 12 points, one at each end of the worm and 5 pairs equally spaced along its length.

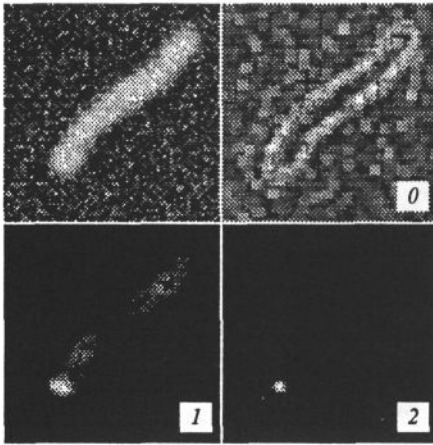


Fig 9(a): noisy worm + $P(x_i)$ after 0,1,2 iterations

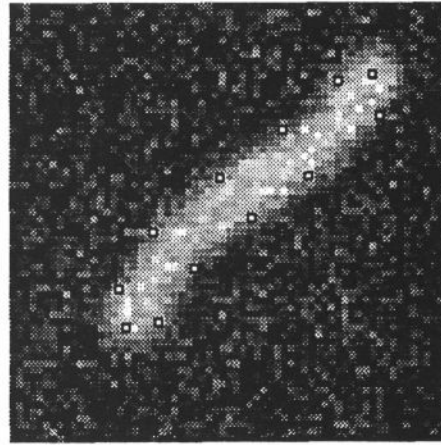


Fig 9(b): noisy worm with located points superimposed

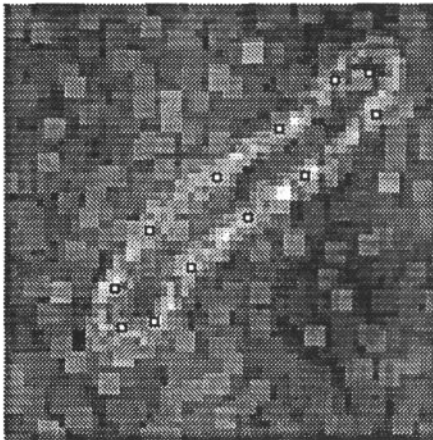


Fig 9(c): initial map data with located points

Fig. 9(a) shows a typical worm with 20% noise added. Also shown are the states of the maps for point x_i after 0,1 and 2 iterations of the updating scheme. Fig. 9(b) shows

the worm at a larger scale. The 12 located points are superimposed. Fig. 9(c) also shows the initial state of the maps $P(x_i)$, which were generated via a morphological edge operator [10]. It is an indication of the power of the method that the points have been located so well considering that no prior integration of the obviously poor edge data has been made.

2.3 Overlapping Objects

We have investigated the behaviour of the revised scheme by applying the model to images of overlapping worms. Fig. 10(a) shows an example in which 4 possibilities arise for the position of x_i . Figs. 10(b) and 10(c) show 2 solutions obtained by selecting the south and west candidate positions for x_i and continuing the iterations. The other 2 solutions are similar and differ only in the labelling of the points.

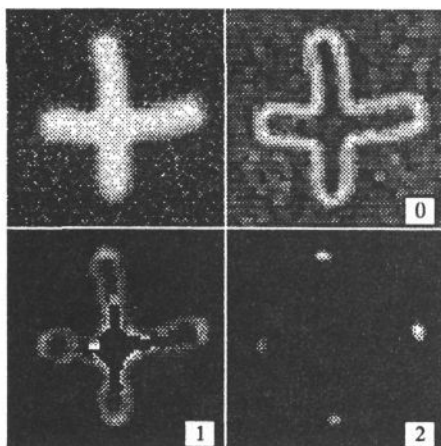


Fig 10(a): crossed worms and maps for point x_i after 0,1,2 iterations

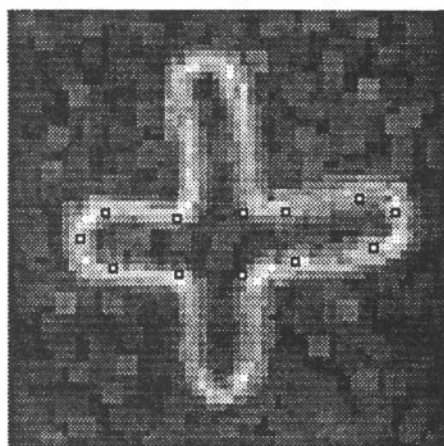


Fig 10(b): first solution

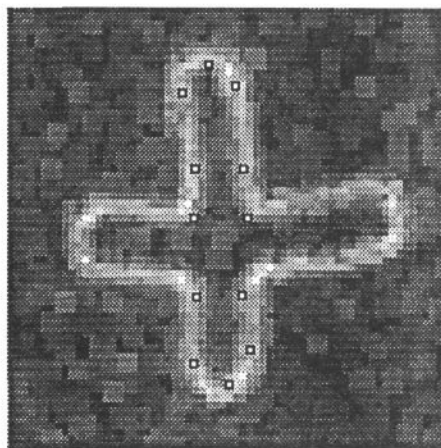


Fig 10(c): second solution

Fig. 11 shows a second case where one of the two solutions fails to include one of the extreme ends of the worm. Probable causes for this behaviour are that (a) the true

distributions for the model parameters are nearer to uniform over an interval than normal as assumed by the model, and so the predictions are weighted against examples at the edges of the distribution as is the case here, and (b) the object edge directions are corrupted in regions of overlap, giving rise to misleading predictions. We have yet to verify whether either of these possibilities is responsible.

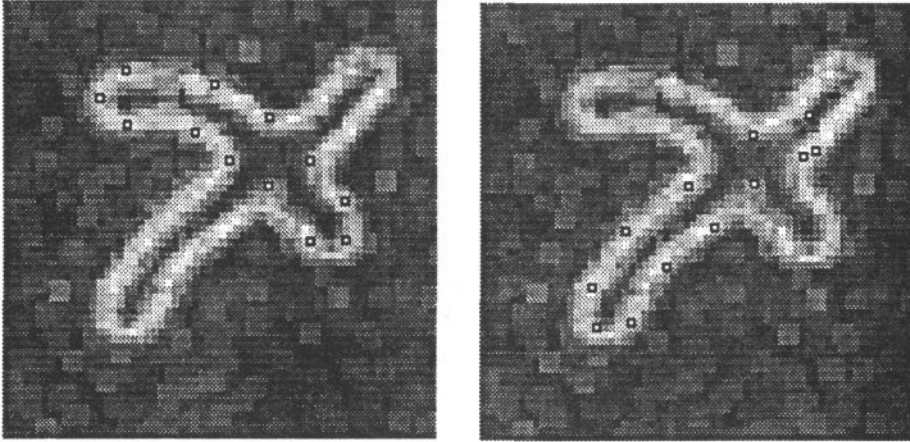


Fig. 11: a second example showing a failure to locate an extreme end in one case

3 DISCUSSION

As the figures above show, the revised CLD representation shows encouraging behaviour in locating objects whose shape is difficult to model explicitly, both in the presence of noise and clutter. Some further work is required to evaluate robustness when occlusion is present, but the results are promising. The method copes with both rigid and flexible objects. As expected, convergence is faster for rigid objects because of the more constraining predictions. Experiments (not described here due to lack of space) indicate that location performance increases with the number N of points in the model, and that digitisation errors can occur if the inter-point distances r_{ij} are small, typically 5 pixels or less. These factors limit the size of the smallest object that can be located.

The main drawback of the method is that it is slow – $O(N^2 a^2 b^2)$ where N is the number of points in the model, a is the typical prediction mask size and b is the region of interest size in pixels. On a SUN3/160 the 12-point model above on a $64 * 64$ region typically takes hours per iteration (66 convolution-type predictions). The predictions can be expensive because the mask size is determined by the size of the object and can be large. Although we can propose a number of ad-hoc tricks to reduce this complexity, we are unlikely even on a modern workstation to achieve execution times which are practical for a working inspection system.

Despite this we can fruitfully apply the technique to multiple objects. By choosing object-defining points or derived reference points (the x_0 in Fig. 1) for several

objects in a scene, we can capture inter-object spatial relationships using a CLD and exploit the arrangement in a top-down (predictive) way to limit search regions. The ability of the CLD to capture variability is being investigated mathematically with a view to applications using other search techniques which use shape generation. Early work in this direction is described in a companion paper [11] submitted to BMVC91.

4 ACKNOWLEDGEMENT

This work has been funded by DTI/SERC as project ref. IED3/1/2114 "VISAGE: Visual Inspection System Application Generation Environment".

5 REFERENCES

- [1] Taylor, C.J., Cooper, D. H., *Shape Verification Using Belief Updating* Proceedings: British Machine Vision Conference BMVC90 (Oxford); pp 61-66, 1990.
- [2] Mokhtarian, F., Mackworth, A. *Scale-based description and recognition of planar curves and two dimensional shapes.* IEEE PAMI Vol. 8 p 34-43 ,1986.
- [3] Brady, M. *Criteria for Representations of Shape* Human and Machine Vision. Academic Press , 1983.
- [4] Pearl, J. *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufman (Publishers) ,1988.
- [5] Bolles, R. C. *Robust Feature Matching Through Maximal Cliques* SPIE , Bellingham, Wash. Vol 182 pp 140-149. , 1979.
- [6] Chin,R.T.,Dyer, C.R. *Model-Based Recognition in Robot Vision* Computing Surveys Vol 18 No 1 , 1986.
- [7] 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.
- [8] 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.
- [9] Ballard, D., Brown, C. *Computer Vision.* Prentice Hall , 1982.
- [10] Maragos, P. *Tutorial on Advances in Morphological Image Processing and Analysis* Optical Engineering Vol 28 No 7 pp 623-632 , 1987.
- [11] Cootes, T.F., Cooper, D. H., Taylor, C.J., Graham, J. *A Trainable Method of Parametric Shape Description* Proceedings: British Machine Vision Conference BMVC91 (Glasgow) , 1991.