

Two-dimensional object recognition using partial contours

David Vernon

Most simple algorithms for generating 2D shape descriptors, whether they are based on boundary features or regional features, require that the complete object is visible. This paper introduces a simple boundary-based shape descriptor, the normal contour distance (NCD) signature, which does not require knowledge of the complete boundary and is suitable for recognition of partially occluded objects. The ability of the NCD descriptor to discriminate between standard 2D shapes has been tested and details are provided of the results obtained using 50%, 50%, 70%, 80%, 90% and 100% of the complete object boundary.

Keywords: shape recognition, shape descriptors, occluded objects

OVERVIEW OF TECHNIQUES FOR SHAPE DESCRIPTION

Many robot vision applications require the analysis and identification of relatively simple objects that are to be subsequently manipulated. While the presentation of objects to the robot in bins is most desirable, it is also, in general, beyond current capabilities (for examples of the state of the art in bin picking, see References 1 and 2). If, however, the objects are (to an approximation) two dimensional, the problem is more tenable, requiring the description and classification of partially occluded planar shapes. Shape description is a central problem in pattern recognition and, as such, it has received considerable attention in the literature. Comprehensive surveys of algorithms for the description and analysis of two-dimensional shapes may be found in References 3 and 4. In these surveys, shape descriptors are classified according to whether the descriptor is based on external

or internal properties of the shape and according to whether the descriptor is based on scalar-transform techniques or on space-domain techniques. Pavlidis also distinguishes descriptors that are information preserving from those that are information nonpreserving. The former allow for the reconstruction of the original shape while the latter techniques do not. External descriptors are typically concerned with properties based on the boundary of the shape. Scalar-transform techniques generate a vector of scalar features and would, for example, be used with decision-theoretic shape classification approaches⁵⁻¹¹. In contrast, space-domain techniques generate spatial or relational descriptions as exemplified by the structural and syntactic approaches^{7,12-17}. Thus, Pavlidis identifies four distinct types of shape descriptor on the basis of whether they correspond to scalar-transform or space-domain techniques and whether the descriptor is based on internal or external properties.

Internal scalar-transform descriptors

Internal scalar-transform techniques generate shape descriptors based on the entire shape. One of the most popular is the method of moments. The standard two-dimensional moments $m(u, v)$ of an image intensity function $g(x, y)$ are defined by

$$m(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) x^u y^v dx dy \quad u, v = 0, 1, 2, 3 \dots$$

It has been shown that these shape descriptors are information preserving¹⁸. The central moments are defined by¹⁹

$$u(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) (x - \bar{x})^u (y - \bar{y})^v dx dy$$

$$u, v = 0, 1, 2, 3 \dots$$

Department of Computer Science, Trinity College, Dublin 2, Ireland

0262-8856/87/01021-07 \$03.00 © 1987 Butterworth & Co. (Publishers) Ltd

where

$$\bar{x} = \frac{m(1, 0)}{m(0, 0)} \quad \bar{y} = \frac{m(0, 1)}{m(0, 0)}$$

That is, \bar{x} and \bar{y} are the coordinates of the centroid of the shape. Moment invariants, simple linear combinations of the normalized central moments, are more frequently used for shape description as they generate values which are invariant with position, orientation and scale changes²⁰. Shape descriptors based on moment invariants convey significant information for simple objects but fail to do so for complicated ones³. These moments may also be generated from the boundary of the object using Stokes' theorem²⁰ or using Green's theorem²¹, both of which relate the integral over an area to an integral around its boundary.

Other internal scalar-transform techniques include the two-dimensional Fourier transform of the characteristic function of the object³. The characteristic function is essentially a two-dimensional binary function having the value 1 if a point is in the object and having the value 0 otherwise. This descriptor is expensive to compute, though it is potentially information preserving, depending on the number of coefficients that are used.

External scalar-transform descriptors

External scalar-transform descriptors are based on scalar features derived from the boundary of an object. Simple examples of such features include the perimeter length, the ratio of the major to minor axis of the minimal bounding rectangle of the shape, and the number and size of residual concavities lying within the bounds of the shape's convex hull. More sophisticated scalar-transform techniques are often based on the Fourier series expansion of a periodic function derived from the boundary. This can be expressed as the Fourier series expansion of the tangential angle of the boundary as a function of the distance around the boundary¹⁹ or as the expansion of a sequence of complex numbers derived by letting the x coordinate correspond to the real part of the number and letting the y coordinate correspond to the imaginary part²². Certain simple manipulations of this frequency domain representation can eliminate dependence on size, position and orientation so that a shape can be matched to a test set of Fourier descriptors regardless of its original size, position and orientation. This is a useful technique and has the advantage that it is easy to program and is backed by a well understood theory. Nevertheless, it suffers from the disadvantage, common to all transform techniques, of difficulty in describing local shape information²³.

Internal space-domain descriptors

Internal space-domain techniques comprise descriptors which utilize structural or relational properties derived from the complete shape. The medial axis transform (MAT) is an example of a commonly used space-domain descriptor in that it generates a skeletal line drawing from a two-dimensional figure. It is a computationally expensive technique and is sensitive to noise. Other descriptors may be derived using integral geometry; for

example, an object shape may be intersected by a number of chords in different directions and the locations and the length of the intersection may be used in various ways as a shape descriptor³. Shape decomposition is a third approach in which the original shape is expressed as the union of certain of its subsets. The shape of the subsets should be simpler than the original and, hence, amenable to simple description. This technique has been used for, among other things, the analysis of cursive script of decomposing the script into strokes^{24,25}.

External space-domain descriptors

Shape descriptors based on external space-domain techniques make explicit use of the spatial or structural organization of the boundary. One popular technique is the use of syntactic descriptors of boundary primitives, e.g. short curves, line segments and corners. These techniques have been applied to computer classification of human chromosomes¹² and to printed-board inspection¹³. Sze and Yang propose the use of a syntactically based recognition strategy with a polygonal approximation of the object boundary²⁶. Polygonal approximations are also used as shape descriptors for recognition of industrial parts²⁷⁻²⁹. References 28 and 29 also suggest shape recognition based on correlating a polar radius signature, encoding the distance from the shape centroid to the shape boundary as a function of ray angle, with a predetermined template. Descriptors based on external space-domain techniques are generally efficient, have minimal storage requirements, and are based on well developed general methodologies such as the theory of formal languages. A disadvantage of these approaches is that points which are geometrically close together can be encoded quite far apart in the boundary representation, though this is offset in the case of syntactic descriptors which are augmented by semantic information, utilizing attributed grammars¹⁷.

A cursory analysis of this brief overview will reveal that most of the descriptors mentioned (with the notable exception of some external space-domain descriptors) require that the complete shape is available for reliable and consistent shape recognition. This paper is directed specifically at alleviating this restriction.

NORMAL CONTOUR DISTANCE SHAPE DESCRIPTOR

The normal contour distance (NCD) shape descriptor is a simple external space-domain descriptor. It is essentially a one-dimensional signature in which each signature value represents an estimate of the distance from a point on an object's boundary, with a local orientation of \mathbf{m} , to the opposing boundary point which lies on a path whose direction is normal to \mathbf{m} . The NCD signature is evaluated over the length of the available boundary. Note that, if the contour represents a partial boundary, there may not be another boundary point which lies on a path which is normal to the contour and, hence, some segments of the NCD may be undefined; these signature elements are assigned a value of zero.

Since this is a boundary-based descriptor, it assumes

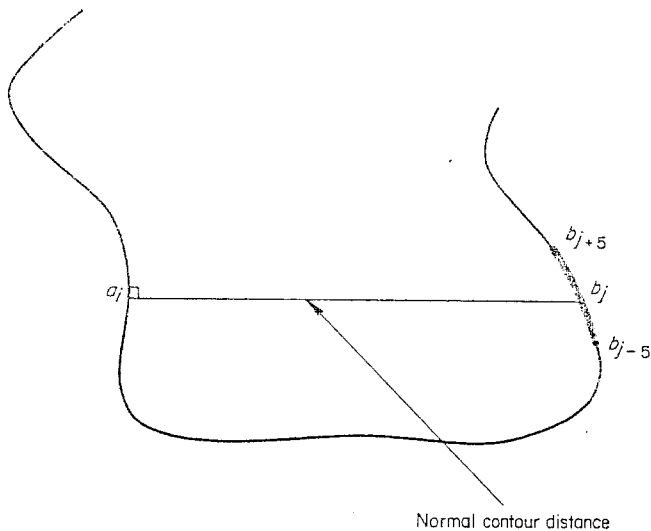


Figure 1. Estimation of normal contour distance

that a boundary representation of the shape exists, from which it can generate the NCD signature. This is achieved, at present, using a dynamic contour-following algorithm which tracks the local maximum gradient (derived using the Prewitt edge detector³⁰) and is capable of bridging short gaps and linking edge segments to yield a continuous segmented (partial) boundary. The contour follower ultimately generates a boundary chain code (BCC) representation³¹ of the shape.

The BCC is, however, a nonuniformly sampled representation of the object boundary since it is based on a square lattice with interpixel distances of 1 in the horizontal and vertical directions and $\sqrt{2}$ in the diagonal directions. Thus, the number of BCC links in a given boundary segment will vary if the overall orientation of the boundary segment changes. This is a severe drawback of the representation. It can be seen that an NCD which is generated from such a BCC will have a signature value for each link and will be inherently dependent on the orientation of the original shape. In order for an NCD to be rotation invariant, it is necessary for it to be generated from a uniformly sampled boundary representation of the object. Uniformly sampled representations have been discussed recently in the literature³² but the suggested algorithm is computationally expensive. A simpler algorithm for transforming a non-uniformly sampled BCC into a uniformly sampled BCC is presented in the appendix. This algorithm, while not as general as that suggested in Reference 32, has proven to be adequate for the purposes of constructing rotation-invariant NCDs.

Once a (partial) resampled boundary has been generated, the corresponding NCD descriptor can then be computed. The following pseudocode summarizes the NCD generation algorithm (refer also to Figure 1).

```

Initialize the NCD to zero
Evaluate the coordinates at every point on the boundary

FOR all points,  $a_i$ , on the boundary
DO
    evaluate the local boundary orientation  $m_i$ 1
    FOR all points  $b_j$  in a limited range of the boundary2
    DO

```

```

/*search for a point that lies on the path which is*/
/*normal to the local boundary orientation. */

```

```

evaluate the orientation  $m_n$  of a line joining the points  $a_i$  and  $b_j$ 

```

```

IF  $m_i * m_n = -1 \pm$  a tolerance of 0.53

```

```

THEN

```

```

    evaluate the normal contour distance  $|a_i, b_j|$  ... this is the NCD
    signature value corresponding to point  $a_i$ 4

```

```

    determine the new limited boundary range:  $b_{i-5}$  to  $b_{i+5}$ 

```

Notes

- 1 The local contour orientation is estimated by computing the slope of the line joining the point a_i and the point a_{i+5} . While the inherent smoothing in this approach overcomes the noisy nature of BCC contours, it also introduces distortion in the NCD as discontinuities in a boundary will also be smoothed; the interval of five pixels was chosen (empirically) as the best tradeoff between these two conflicting properties.
- 2 The looping construct governing the search for a point giving a normal direction to the local boundary orientation actually terminates when the first acceptable point is found; the entire range is not searched exhaustively. If no point lying on the normal path is found within this limited range, then a search is made on the complete boundary.
- 3 Owing to the numerical accuracy and the discrete representation of the boundary, it is necessary to allow a tolerance on the product of slopes when determining the normal path. A tolerance value of 0.5 was chosen since it provided acceptable accuracy in estimating the normal direction while ensuring that the normal would, indeed, be detectable.
- 4 The normal contour distance $|a_i, b_j|$ is the euclidean distance between these two boundary points, given by $[(a_{ix} - b_{jx})^2 + (a_{iy} - b_{jy})^2]^{\frac{1}{2}}$.

DISCUSSION OF RESULTS

The development of the NCD shape descriptor was motivated by a need to accomplish 2D object recognition using partial contours. In order to quantify the recognition ability of the NCD descriptor, a series of tests was carried out using a set of standard shapes which have been proposed by Rosen and Gleason³³. This set comprises a disc, a square, a rectangle and an equilateral triangle (see Figure 2).

The testing strategy which was employed may be summarized as follows. An NCD template of each of the four shapes is learned by the vision system. The measure of similarity between each of these templates and four test shapes is evaluated and the test shape is classified as belonging to the class of shapes for which it exhibits the maximum similarity. This is repeated for ten sets of test shapes and the entire procedure is repeated five times, using a newly learned template on each occasion. Since an estimate of the descriptor's performance with partial contours is required, each classification is performed using NCD descriptors which have been generated using 50%, 60%, 70%, 80%, 90% and 100% of the original test shape BCC. The classification

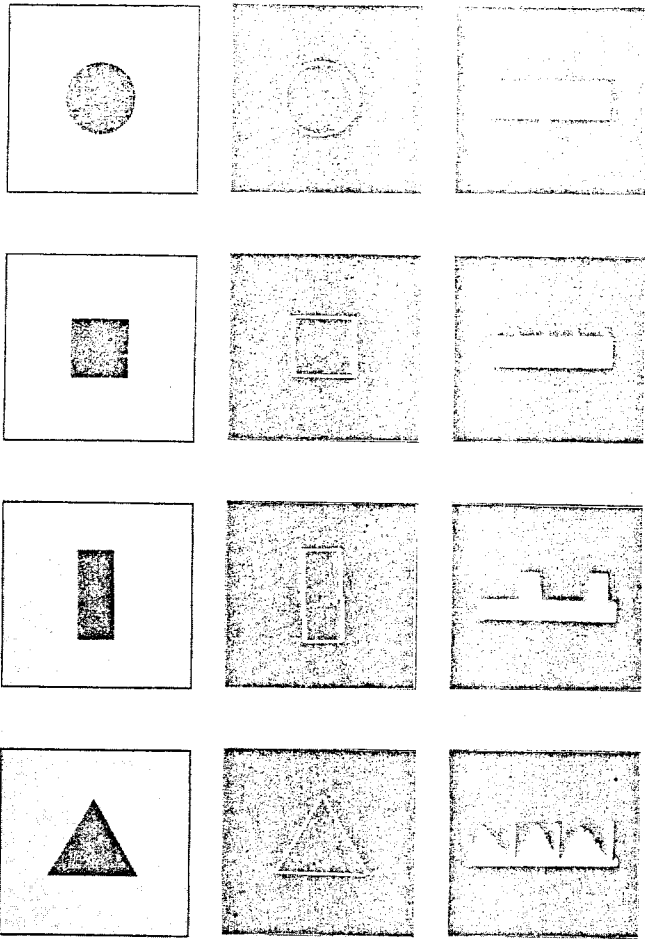


Figure 2. Test shapes, extracted boundaries and NCD signatures

Table 1. Classification of test shapes using 50% of boundary

Test	Template			
	Disc	Square	Rectangle	Triangle
Disc	0	0	11	39
Square	0	0	0	50
Rectangle	0	0	7	43
Triangle	0	0	10	40

Table 2. Classification of test shapes using 60% of boundary

Test	Template			
	Disc	Square	Rectangle	Triangle
Disc	0	0	0	50
Square	0	0	0	50
Rectangle	0	0	31	19
Triangle	0	0	5	45

results obtained for each of the 50 tests performed with 50%–100% partial contours are summarized in Tables 1–6, respectively.

Note that, since an arbitrarily learned template need not necessarily be ideally representative of the object, each learning procedure is repeated five times and a measure of the similarity between the descriptors corres-

Table 3. Classification of test shapes using 70% of boundary

Test	Template			
	Disc	Square	Rectangle	Triangle
Disc	0	1	0	49
Square	0	1	0	49
Rectangle	0	0	45	5
Triangle	0	0	0	50

Table 4. Classification of test shapes using 80% of boundary

Test	Template			
	Disc	Square	Rectangle	Triangle
Disc	33	17	0	0
Square	0	45	0	5
Rectangle	0	0	50	0
Triangle	0	0	0	50

Table 5. Classification of test shapes using 90% of boundary

Test	Template			
	Disc	Square	Rectangle	Triangle
Disc	49	1	0	0
Square	0	50	0	0
Rectangle	0	0	50	0
Triangle	0	0	0	50

Table 6. Classification of test shapes using 100% of boundary

Test	Template			
	Disc	Square	Rectangle	Triangle
Disc	50	0	0	0
Square	0	50	0	0
Rectangle	0	0	50	0
Triangle	0	0	0	50

ponding to each shape is evaluated. The descriptor which exhibits the best overall performance (i.e. the highest average similarity measure) is selected as the shape template.

The similarity measure used in these tests is a commonly used variation of the template matching metric based on the standard euclidean distance between two vectors³⁴. This variation uses the absolute difference of the template elements rather than the square of their difference³⁵ and is defined by

$$S(m) = \sum_i |g(i) - t(i - m)|$$

The summation is evaluated for all i , such that $i - m$ is in the domain of definition of the template. Alternative template matching strategies, for example the generalized Hough transform, could be adopted and have been successfully utilized elsewhere³⁶.

Figure 2 illustrates the test shapes, the corresponding boundary representation, and the associated NCD shape descriptor.

Two points should be noted. First, the NCD descriptor obtained from the square and the disc are (as one would expect) very similar, except for the spikes which are due to the contour discontinuities at the corners of the square. Despite this similarity, the square has been correctly classified in every instance using 100% and 90% of the complete contour to generate the NCD, and with a 45% success rate using 80% of the contour. Similarly, the disc was classified with 100% accuracy using an NCD generated from the complete boundary and with 98% and 66% accuracy using NCCs generated with 90% and 80% of the boundary, respectively. Bearing in mind the similarity of the NCD signatures for these two shapes, these results are satisfactory. The results obtained with the rectangle and triangle (which have distinctive NCD signatures) are significantly better. In the case of the triangle, it was classified correctly with 100% accuracy for NCDs generated from partial contours which are greater than or equal to 70% of the complete boundary and with 90% accuracy in the case of an NCD generated from 60% of the complete boundary. Similarly, the rectangle NCD demonstrated 100% success with 100%, 90% and 80% partial contours respectively and a success rate of 90% with a 70% partial contour.

The second point to note is that these partial contour tests generate NCD descriptors on the basis of one single partial contour, implying that the nonzero component of the NCD signature is minimal. Consider, for example, the instance of an NCD generated from 50% of the disc contour. In this case, there are no opposing contours (except in ideal circumstances, at the very ends of the arc) and the NCD descriptor will comprise zero values exclusively. This restraint would be alleviated somewhat by the use of two (or more) partial contours which correspond to opposite sides of the object, a situation which is quite likely when dealing with partially occluded objects.

CONCLUSIONS

The 2D shape descriptor which has been introduced in this paper encodes, for every boundary point, the distance to an opposing point on the object boundary along a path which is normal to the local contour orientation. Though the NCD signature is based on both local and global information, and is therefore quite robust, the signature can be generated from (several) partial contours. This contrasts with the requirements of most other simple shape description algorithms (radii signatures, for example) and, as such, it is suitable for recognition of partially occluded objects. Results obtained using the NCD with partial contours are encouraging and subsequent work is currently being directed towards the use of NCD partial-contour classification in aiding the segmentation of occluded objects. The use of knowledge-

based approaches such as the hypothesis generation and verification technique suggested in Reference 37, is also being investigated.

ACKNOWLEDGEMENTS

This research was supported by the European Strategic Program for Research and Development in Information Technology (ESPRIT) under project P419: Image and Movement Understanding.

REFERENCES

- 1 Ikeuchi, K, Nishihara, H K, Horn, B K, Sobolvarro, R and Nagata, S 'Determining grasp configurations using photometric stereo and the PRISM binocular stereo system' *Int. J. Robotics Res.* Vol 5 No 1 (1986) pp 46-65
- 2 Kelley, R B, Martins, H A S, Birk, J R and Dessimoz, J D 'Three vision algorithms for acquiring workpieces from bins' *Proc. IEEE* Vol 71 No 7 pp 803-821
- 3 Pavlidis, T 'A review of algorithms for shape analysis' *Comput. Graphics Image Process.* Vol 7 (1978) pp 243-258
- 4 Pavlidis, T 'Algorithms for shape analysis of contours and waveforms' *IEEE Trans. Pattern Anal. Mach. Intell.* Vol PAMI-5 No 6 (1980) pp 584-592
- 5 Agin, G 'An experimental vision system for industrial application' *Technical Report 103* SRI International, Menlo Park, CA, USA (1975)
- 6 Gleason, J G and Agin, G J 'A modular vision system for sensor-controlled manipulation and inspection' *Technical Note 178* SRI International, Menlo Park, CA, USA (1979)
- 7 Fu, K 'Digital pattern recognition' *Commun. Cybern.* Vol 10 (1980)
- 8 Chen, M J and Miligram, D L 'A development system for machine vision' *IEEE Computer Society Conf. on Pattern Recognition and Image Processing* (1982) pp 512-517
- 9 Pot, J, Coiffet, P and Rives, P 'Comparison of five methods for recognition of industrial parts' in Aleksander, I (ed.) *Artificial vision for robots* Kogan Page, London, UK (1983) pp 43-57
- 10 Batchelor, B G and Cotter, S M 'Detection of cracks using image processing algorithms implemented in hardware' *Image Vision Comput.* Vol 1 No 1 (1983) pp 21-29
- 11 Connors, R W, McMillin, C W, Lin, K and Vasquez-Espinosa, R E 'Identifying and locating surface defects in wood: part of an automated lumber processing system' *IEEE Trans. Pattern Anal. Mach. Intell.* Vol PAMI-5 No 6 (1983) pp 573-583
- 12 Mendelsohn, M L, Mayall, B H and Prewitt, J M S 'Approaches to the automation of chromosome analysis' in *Image processing in biological science* University of California Press, Berkeley, CA, USA (1968) pp 119-136
- 13 Bjorklund, C M and Pavlidis, T 'On the automatic inspection and description of printed wiring boards' *Proc. Int. Conf. Cybernetics Society* Princeton, NJ, USA (1977) pp 690-692

- 14 Chin, R T and Harlow, C A 'Automated visual inspection: a survey' *IEEE Trans. Pattern Anal. Mach. Intell.* Vol PAMI-4 No 6 (1982) pp 557-573
- 15 Fu, K 'Pattern recognition for automatic visual inspection' *Computer* Vol 15 No 12 (1982) pp 34-40
- 16 Gaglio, S, Morasso, P and Tagliascio, V 'Syntactic techniques in scene analysis' in Aleksander, I (ed) *Artificial vision for robots* Kogan Page, London, UK (1983) pp 58-74
- 17 You, K C and Fu, K S 'A syntactic approach to shape recognition using attributed grammars' *IEEE Trans. Syst. Man, Cybern.* Vol SMC-9 No 6 (1979) pp 334-345
- 18 Hu, M K 'Visual pattern recognition by moment invariants' *IRE Trans. Inf. Theory* Vol IT-8 (1962) pp 179-187
- 19 Gonzalez, R C and Wintz, P *Digital image processing* Addison-Wesley, Reading, MA, USA (1977)
- 20 Wiejak, J S 'Moment invariants in theory and practice' *Image Vision Comput.* Vol 1 No 2 (1983) pp 79-83
- 21 Tang, G Y 'A discrete version of Green's theorem' *IEEE Trans. Pattern Anal. Mach. Intell.* Vol PAMI-4 No 3 (1982) pp 242-249
- 22 Hall, E L *Computer image processing and recognition* Academic Press, New York, USA (1979)
- 23 Rice, J R *The approximation of functions* Addison-Wesley, Reading, MA, USA (1969)
- 24 Frischkopf, L S and Harmon, L D 'Machine reading of cursive script' *Proc. Symp. on Information Theory* Butterworths, London, UK (1961) pp 300-316
- 25 Eden, M 'Hand writing and pattern recognition' *IRE Trans. Inf. Theory* Vol IT-8 (1962) pp 160-168
- 26 Sze, T-W and Yang, Y-H 'Goal directed segmentation' *IEEE Computer Society Conf. on Pattern Recognition and Image Processing* (1982) pp 504-510
- 27 Ayache, N, Faverjon, B, Boissonnat, J D and Bollack, B 'Manipulation automatique de pieces industrielles en vrac planaire' *Proc. 1st Image Symp.* Biarritz (1984) pp 869-875
- 28 Juvin, D and Dupreyrat, B 'ANIMA (analysis of images): a quasi real-time system' *IEEE Computer Society Conf. on Pattern Recognition and Image Processing* (1981) pp 358-361
- 29 Juvin, D and de Cosnac, B 'ANIMA 2: un système générale de vision pour la robotique' *Proc. 1er Colloque Image* Biarritz, France (1984) pp 165-169
- 30 Prewitt, J M S 'Object enhancement and extraction' in Lipkin, B and Rosenfeld, A (eds) *Picture processing and psychopictorics* Academic Press, New York, USA (1970) pp 75-149
- 31 Freeman, H 'Computer processing of line-drawing images' *ACM Comput. Surv.* Vol 6 No 1 (1961) pp 57-97
- 32 Shahraray, B and Anderson, D J 'Uniform resampling of digitized contours' *IEEE Trans. Pattern Anal. Mach. Intell.* Vol PAMI-7 No 6 (1985) pp 674-681
- 33 Rosen, C A and Gleason, G J 'Evaluating vision system performance' in Pugh, A (ed) *Robot vision* IFS (Publications) Ltd, UK (1983) pp 97-103
- 34 Duda, R O and Hart, P E *Pattern classification and scene analysis* (1970)
- 35 Nevatia, R *Machine perception* Prentice-Hall, Englewood Cliffs, NJ, USA (1982)
- 36 Turney, J L, Mudge, T N and Volz, R A 'Recognising

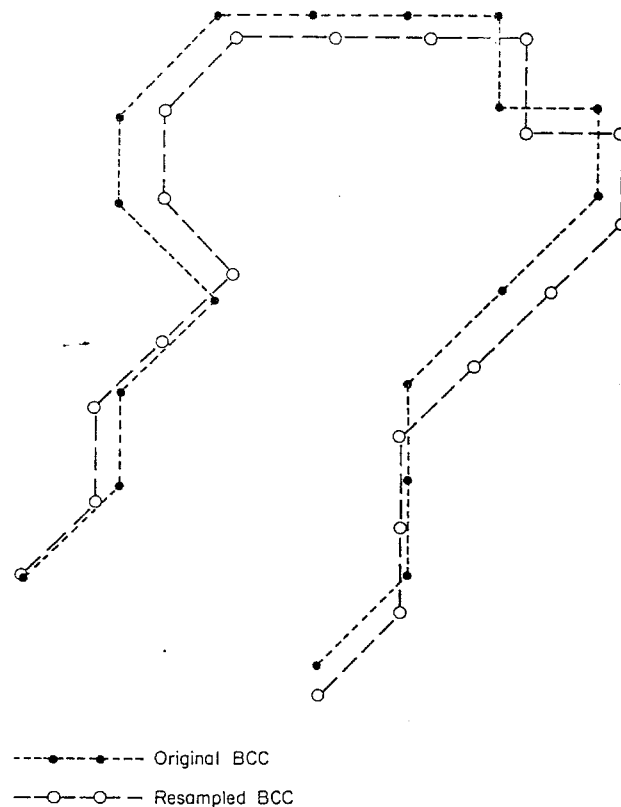


Figure 3. Resampling a BCC

partially occluded parts' *IEEE Trans. Pattern Anal. Mach. Intell.* Vol PAMI-7 No 4 (1985) pp 410-421

- 37 Knoll, T F and Jain, R C 'Recognising partially visible objects using feature indexed hypotheses' *IEEE J. Robotics Automation* Vol RA-2 No 1 (1986) pp 3-13

APPENDIX: AN ALGORITHM FOR RESAMPLING BOUNDARY CHAIN CODES

A BCC is a nonuniformly sampled function, i.e the distance between the sample points along the boundary may be either 1 or $\sqrt{2}$ depending on whether the neighbouring boundary points are horizontal/vertical neighbours or diagonal neighbours, respectively. Thus, a BCC is dependent on the orientation of the object boundary on two distinct bases:

- Each link encodes the absolute direction of the boundary at that point.
- The link length (1 or $\sqrt{2}$) varies with the boundary direction.

Any shape descriptor which is derived from this non-uniformly sampled BCC is inherently sensitive to changes in orientation. To alleviate this rotational variance, it is necessary to remove the dependence on link length, ensuring that the link lengths of the BCC are all equal, specifically by resampling the BCC in uniformly spaced intervals.

Such a sampling system will generate noninteger coordinate values; the actual image pixel values can be obtained for shape reconstruction, which is of interest

