IEEE Pacific Rim Conference on Communications, Computers and Signal Processing
June 1st - 2nd, 1989

POSITION, ROTATION, AND SCALE-INVARIANT RECOGNITION OF 2-DIMENSIONAL OBJECTS USING A GRADIENT CODING SCHEME

R. Safaee-Rad*, B. Benhabib*, K.C. Smith*+, K.M. Ty+
Computer Integrated Manufacturing Laboratory*
*Department of Mechanical Engineering
+Department of Electrical Engineering
University of Toronto
Toronto, Ontario, M5S 1A4

Abstract

The 2D representation and recognition of objects, based on external space-domain descriptors, is addressed in this paper. A new shape signature is proposed for 2D shape recognition, namely the *gradient-perimeter* plot. The recognition method is , position-, rotation-, and scale-invariance. For the matching process a cost function, based on mean square error (MSE), is minimized. It is also shown that, as the basis of an alternative method, the *angle-perimeter plot* can be used to provide a feature set for 2D shapes. This method can be applied to polygonal shapes, or shapes for which polygonal approximations are available. Both methods were successfully applied to a set of standard shapes.

1. Introduction

The flexibility of robotic workcells can be significantly increased through the integration of visual sensors for 3D object recognition; thus, reducing the need for customized and often complex tooling by implementing flexible automation rather than hard automation. The research program currently underway at the University of Toronto focuses on developing an efficient and flexible 3D object-recognition method for vision-based robotic assembly. The main incentives in pursuing this research are the lack of generalized computer vision theories applicable to robotic vision, and the inefficiency and inflexibility of existing 3D object-recognition methods for industrial environments.

Volumetric methods based on the exact specification of an object, and multi-view feature representation based either on the characteristic views of an object or on the discrete view-sphere representation have been developed [1,2]. However, these techniques are complex and computationally expensive due to their need for a 3D matching process. The new 3D object-recognition technique under development is based on a pre-marking scheme, through which the matching process is performed in 2D space [3].

In the proposed 3D object-recognition technique, an object is modeled using only a small number of 2D distinct perspective views. The number of views depends partly on the degree of symmetry of the object. These are referred to as *standard-views*, each with a corresponding *standard-view-axis*. For successful recognition, the input image of an object must be one of its standard perspective views. Thus, a mobile camera is used, such that the camera's optical axis can be aligned with one of the standard-view-axes of the object in order to acquire a standard-view. The matching process is, then, performed between the acquired 2D standard-view of the object under consideration and the library of 2D standard-views of a set of objects. Based on the proposed method, then, any 2D representation and its corresponding recognition technique that is used, must be position-, rotation-, and scale-invariance.

The problem of 2D shape recognition is one of the most familiar and fundamental problems in pattern recognition. Comprehensive surveys of algorithms for description and analysis of 2D shapes have been reported in [4,5]. According to these surveys, 2D shape descriptors can be divided into four classes, two of which are based on the external (boundary) properties of the shape (external scalar-transform descriptors and external space-domain descriptors). Boundary is the basis of these two 2D shape analysis techniques, since it has been shown that the information content of a shape is concentrated along the boundary contours, and furthermore, at points (on those contours) at which the direction changes most rapidly [6].

One of the 2D representation and recognition techniques under study for integration in the proposed vision system, is based on the boundaries of shapes (an external space-domain descriptor). Boundary representation schemes generally express the curvature in some manner as a function of linear or angular variation along a blob's periphery. These functions can be referred to as signatures of planar shapes. In general, a signature is a function that can be associated with any continuous planar curve, and its definition combines local variations in the curve's tangent with the global layout of the length on the plane.

Various 2D shape signatures (with different names) have been proposed and used for pattern analysis in the machine-vision literature, such as: "K-curvature" [7], "incremental curvature" (based on a line-segment scan method) [8], "angle versus length signature" [9], "polar coding of an object" [10], "plane curve signature" [11], "the curvature K(s) of a contour" [12], "normal contour distance (NCD) signature" [13], and, "slope density function" [14]. In this paper, a new boundary based 2D shape descriptor (signature) is defined (under the constraints of position-, rotation-, and scale-invariance) and applied to a set of standard 2D shapes. The boundary is traced and chain-coded using the Freeman chain code [15].

2. Gradient-Perimeter Function

A number of preprocessing steps must be carried out before the boundary of an object silhouette can be chain-coded. These include the thresholding of a "good" contrasted image, followed by its edge detection using, for example, the Sobel operator. The starting point of the edge is recorded and the edge chain coded. The gradient map used for chain coding is shown in Figure 1. Since the Freeman gradient map can cause a sudden jump from 0 up to 7, or 7 down to 0, when the gradient is plotted against the distance along the contour, forming a gradient-perimeter plot, a modified code is used. The next vector code, C_{i+1}, in the modified chain code is defined as,

$$C_{i+1} = C_i - (U_i - U_{i+1} + 8K) \quad \text{for } i > 0 \tag{1}$$

where, $K = 0$ if $U_{i+1} < \text{octmod}(U_i + 3)$

$K = 1$ if $U_{i+1} > \text{octmod}(U_i + 3)$

$C_i = U_i$ for $i=1$.

U_i is the original Freeman code. The function octmod (S) is the absolute value of S, modulo 8. Applying equation (1) to the Freeman code sequence "2, 1, 0, 7, 0, 6", for example, would yield the modified code sequence "2, 1, 0, -1, 0, -2". Thus, the difference between two consecutive chain vectors is always less than 4.

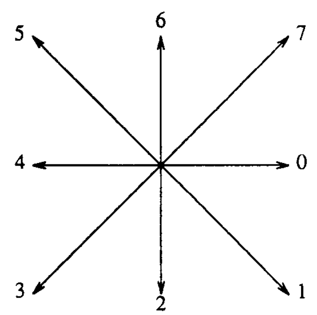

Figure 1. The gradient map.

Using this modified gradient coding scheme, if the boundary of the object silhouette is traced from left to right and top to bottom, a convex boundary would have a non-increasing gradient-perimeter function. A polygonal approximation of a boundary with none of its angles greater than 180 degrees is defined as convex boundary, whereas a boundary with any of its angles greater than 180 degrees is defined as a (partially) concave boundary.

Since only 8 discrete directions can be obtained from the boundary of the original object silhouette, the resolution of the direction vector may be seen to be too low for a good estimate of the slope at a point on the boundary of the object silhouette To increase the resolution, one can increase the number of direction vectors in the map of Figure 1, but this requires the examination of more than the nearest 8 grid neighbors of the current point on the boundary. Moreover, the complexity of the algorithm would increase greatly if more neighboring pixels were examined.

In order to retain low complexity and improve the resolution, an averaging process, which uses a set of P adjacent elements of the contour chain, is proposed . Based on the experiments carried out on a set of standard 2D shapes, P was selected to be 15. The averaging procedure for the gradient vectors is important, since it eliminates the discrete jumps in the values of the gradient vectors and provides a better estimation of the true gradient. The criterion for choosing the value of P was derived from the view that: "P should not be too large, so that important shape information such as corners would be smoothed out; however, it should also not be too small, so that it causes poor estimation of the gradient ".

The proposed averaging process is as follows: Let C_i be the gradient code, (direction code of which there are 8), of the current link, and, C_i^x and C_i^y represent the lengths of its projection on the x- and y-axes respectively, then,

$$C_i^x = | C_i | \cos\theta \tag{2}$$

$$C_i^y = | C_i | \sin\theta \tag{3}$$

where $| C_i |$ is the link length of the gradient code C_i (which can be either 1 or square root of 2), and θ is the angle measured from the positive x-axis counterclockwise.

Then, if P is an odd number, and $M = (P-1)/2$, the chain vectors C_{i-M}, C_{i-M+1}, ..., C_i, ..., C_{i+M} would be examined. The total projection on the x-axis, X_{proj}, is the sum of the P individual projections on the x-axis,

$$X_{proj} = \sum_{j=-M}^{M} C_{i+j}^x . \tag{4}$$

A similar expression can be defined for Y_{proj},

$$Y_{proj} = \sum_{j=-M}^{M} C_{i+j}^y . \tag{5}$$

The direction of the line can be determined from the angle θ_{proj} formed by the resultant "sum" vector and the positive x-axis, where θ_{proj} is equal to $\tan^{-1}(Y_{proj}/X_{proj})$. The vector code of the current link after averaging, $\overline{C}_i$, is then given by,

$$\overline{C}_i = 8 - \frac{\theta_{proj}}{\frac{\pi}{4}} + 8n. \tag{6}$$

The value of n (..., -2, -1, 0, 1, 2, ...) is chosen such that the absolute value of $(\overline{C}_i - C_i)$ would be less than 4. It should be noted that, in general, $\overline{C}_i$ does not have an integer value.

With an edge tracing algorithm, the modified edge-gradient-coding method, and the above gradient-code-smoothing process, the function of the edge gradient versus the distance along the contour, namely the *gradient-perimeter function*, can be plotted. The distance is defined in terms of the number of links and link lengths, (either 1 or square root of 2), in the Freeman chain code, and measured from the starting point of the chain to the point of interest.

In particular, provided that the edge of the object's silhouette is traced from left to right, and from top to bottom, the gradient-perimeter plot of a circle is a straight line with a constant negative slope. That for r-sided convex polygon is a descending stare case having r+1 horizontal segments separated by sudden descending jumps. If the boundary tracing starts from one of the vertices, there are only r horizontal segments.

Figure 2 and Figure 3 show the gradient-perimeter plots of a rectangle and a triangle, respectively. In these figures, it can be seen that the plots have 4 and 3 descending jumps corresponding to 4 and 3 straight edges of these two shapes, respectively.

3. Matching of Gradient-Perimeter Functions

The gradient-perimeter function described in section 2 is an effective descriptor for shape recognition. A matching process based on the gradient-perimeter function is as follows: Given the gradient-perimeter functions of an unknown shape and a reference shape , the first step in the scale-invariant recognition process would be to normalize the two perimeter lengths, such that, they have the same

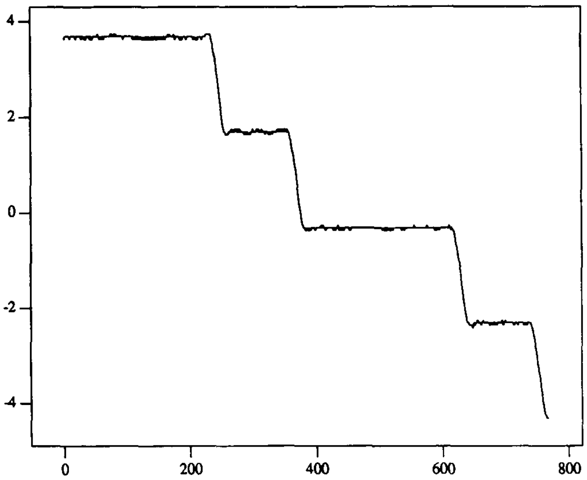

Figure 2. The gradient-perimeter plot of a rectangle.

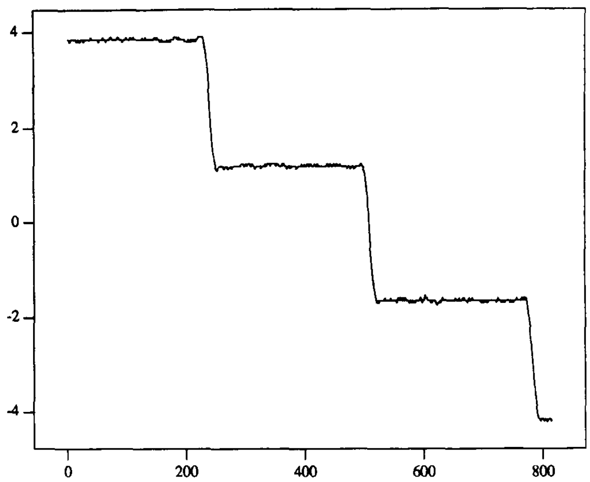

Figure 3. The gradient-perimeter plot of a triangle.

(total) number of links, N. The next step is to compare mutually shifted versions of the gradient-perimeter functions of the unknown and a reference shapes. The minimum mean square error (MSE) [18] between the two mutually shifted gradient-perimeter functions is used to decide whether the unknown shape and the reference shape match.

Let f(p) and g(p) be the two normalized gradient-perimeter functions, and the following be the cost function, T_q,

$$T_q = \sum_{p=0}^{N-1} \frac{[\, f(p) - g(p+q)]^2}{N} - \left[\sum_{p=0}^{N-1} \frac{f(p) - g(p+q)}{N} \right]^2 \quad (7)$$

where q is the amount of shift along the perimeter. Note that, both f(p) and g(p+q) are periodic functions with a period N, due to being the gradient-perimeter functions of closed contours. The first term (MSE) of the cost function can be considered as the total error power, and the second term as the total d.c. power, which accounts for the rotation-invariant aspect of the algorithm. If a shape is rotated, and the gradient-perimeter function starts at the same point on the contour, there is a d.c. bias between the original shape and the rotated shape gradient-perimeter functions. Thus, the d.c. bias removal is important in the rotation-invariant recognition process.

Minimization of the cost function provides the value for the parameter q. This value contains the information about the mapping of the starting point of the gradient-perimeter function of the unknown shape to the starting point of the reference gradient-perimeter function. As a result, the angle of rotation of the unknown relative to the reference shapes can be determined. Once the minimum and maximum values of the cost function have been determined for the unknown shape and a reference shape, this procedure must be repeated for the next and all the remaining reference shapes. The reference shape that yields the minimum cost function is considered as a possible match.

In general, a good match is achieved if the total error power is approximately equal to the d.c. power. The maximum value of MSE, on the other hand, can indicate the degree of geometrical orthogonality of the reference and unknown shapes.

The relative cost function for two similar rectangles is plotted in Figure 4. The two local minimum values are achieved at q=83 and q=332, when the normalized perimeter length is N=500 links. These values suggest that the shape has an axis of symmetry.

4. Angle Measurement Using the Chain Code

It has been shown that most of the information concerning the shape of a contour is carried by the curvature extrema. One reason for this phenomenon is that these extrema divide the outline into its major parts, through which the identity of the shape can be determined [6,16,17]. This is the main incentive for the detection of the corners of a 2D shape (based on its polygonal approximation), and the estimation of its angles. With the modified chain code proposed in this paper, it is possible to detect the corners, to determine the relationships between them, and estimate their angles. Based on these local information, it is possible to develop an alternative method for shape recognition.

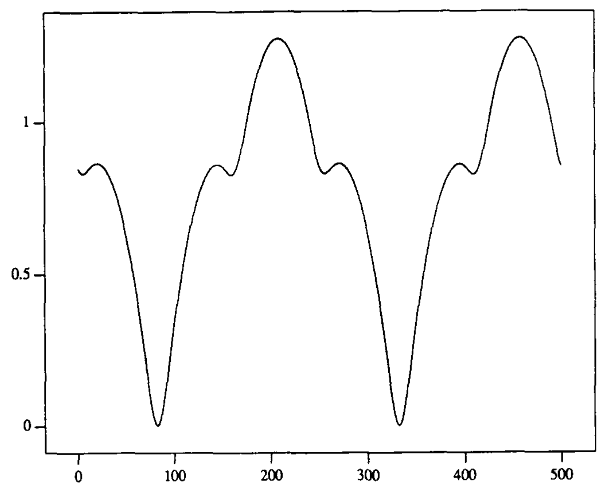

Figure 4. The relative cost function (MSE term) for two similar rectangles.

Let $\overline{C}_{in}$ and $\overline{C}_{out}$ be the average directions of the incoming and outgoing chain vectors, respectively, at the current point i on the perimeter. The direction $\overline{C}_{in}$ is a function of $\{C_{i-K}, C_{i-K+1}, ..., C_{i-1}\}$, and $\overline{C}_{out}$ is a function of $\{C_{i+1}, C_{i+2}, ..., C_{i+K}\}$. The parameter K specifies the start and the end points of the averaging window for the incoming and outgoing vectors. Based on the results of experiments carried out on a set of standard 2D shapes, K was selected to be 9. Applying the method described in section 2, the average incoming and outgoing angles, $(\theta_{in}, \theta_{out})$, can be estimated, where the angle at the current point i can then be defined as,

$$\theta_i = 180 - (\theta_{out} - \theta_{in}). \qquad (8)$$

A graph of the angles measured along a contour versus the distance along a contour, called an *angle-perimeter plot*, is a useful presentation. Figure 5 shows an angle-perimeter plot. A vertex exists where either a local minimum or a local maximum occurs in the angle-perimeter plot. Note that, the angle-perimeter plot in Figure 5 has four local minima, and consequently corresponds to a convex shape. Since each of the minima is approximately 90 degrees, it can be concluded that the 2D shape (which is that of a rectangle) has four vertices with four approximately 90-degree angles.

Figure 6 shows the angle-perimeter plot for a triangle. The plot has 3 minima corresponding to three angles, all less than 180 degrees, the sum of which is approximately equal to 165 degrees. The 15 degrees error, introduced by the application of the proposed method of angle measurement, is less than 10%.

The fluctuations seem to occur around the 180-degree value can be considered as noise. To detect a vertex, two threshold limits were set. A local minimum or a local maximum must be smaller than 160 degrees or greater than 200 degrees, respectively, to be considered as a vertex. Thus using this method, a polygon can be described in terms of the total number of vertices, with the number of angles smaller than 180 degrees indicating local convexity, and with the number of angles greater than 180 degrees indicating local concavity.

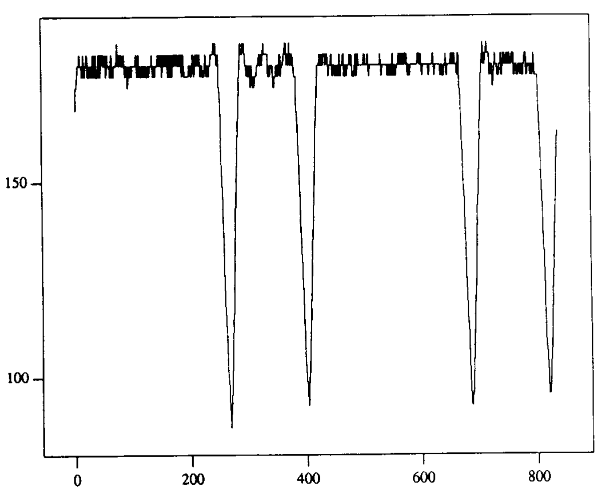

Figure 5. The angle-perimeter plot of a rectangle.

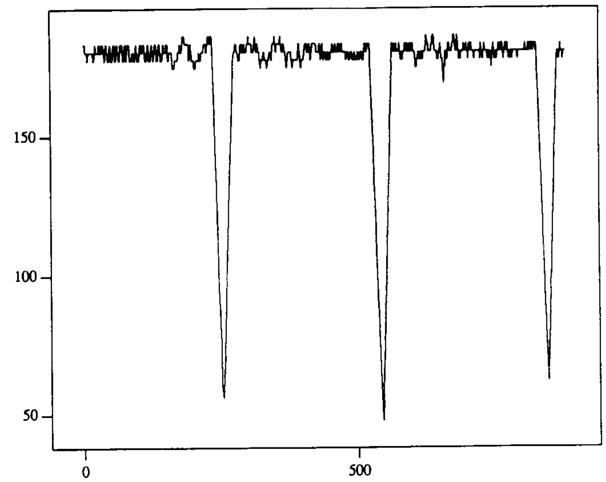

Figure 6. The angle-perimeter plot of a triangle.

Since the position of a vertex can be computed from the coordinates of the starting point and the sequence of the chain vectors, the position of each vertex can be expressed in x- and y-coordinates. From this data, the distance between any two vertices can be computed. Based on this information, the relationships between the vertices and between lengths of the polygon sides can be determined.

5. Results and Discussion

For the recognition of an unknown 2D shape, its gradient-perimeter function must be compared with the gradient-perimeter functions of all the reference shapes. The minimum difference value between these functions would yield the matching shape. In addition, the relative angle of rotation can also be determined using the second term of the cost function, once the shape is recognized.

This method was applied to a set of 2D standard shapes, where the results are given in Table 1. In this table, the "reference shapes" refer to known shapes, each with a specific orientation; whereas, the "test shapes" refer to unknown shapes, with unknown orientations. The minimum MSE and maximum MSE values are evaluated by shifting the normalized gradient-perimeter function for the test shape against the reference shape. These two values, as well as their ratio are given in Table 1.

The first 9 entries in Table 1 correspond to cases where the reference and test shapes are the same. The remaining entries have different reference and test shapes. "Ellipse.n" has the ratio (major-axis-length/minor-axis-length) equal to n. In general, minimum MSE values would be close to zero only when the test and the reference shapes are identical (as expected), or when the two shapes are quite similar. This is especially true in the case of ellipses with n values larger than 4, due to the fact that their boundaries are not well represented on the rectangular sampling grid.

For a reliable match, the ratio (minimum MSE/ maximum MSE) must also be close to zero. Two exceptions to this rule are: ellipses with n values larger than 4, and identical circles. The reason for obtaining a comparatively large MSE ratio value for identical circles, is due to the universal symmetry of circles. Based on these results, it can be concluded that the minimum MSE is a good feature for recognition purposes, especially for well-resolved images.

Reference Object	Test Object	Minimum MSE	Maximum MSE	Ratio of Min. MSE to Max MSE
cam	cam	0.0065	0.3704	0.0175
circle	circle	0.0012	0.0168	0.0731
ellipse.2	ellipse.2	0.0061	0.7983	0.0076
ellipse.4	ellipse.4	0.0142	1.8787	0.0075
ellipse.8	ellipse.8	0.0088	2.7095	0.0032
ellipse.20	ellipse.20	0.0184	3.2851	0.0056
polygon	polygon	0.0064	5.0938	0.0013
rectangle	rectangle	0.0022	1.2735	0.0017
triangle	triangle	0.0238	1.6463	0.0145
cam	circle	0.1070	0.1140	0.9387
cam	rectangle	0.2002	0.7680	0.2607
cam	triangle	0.4738	0.7821	0.6058
circle	ellipse.2	0.1806	0.2077	0.8694
ellipse.2	ellipse.4	0.0784	1.2423	0.0631
ellipse.4	ellipse.8	0.0167	2.3260	0.0072
ellipse.8	circle	0.8756	0.9619	0.9102
ellipse.8	ellipse.20	0.0036	3.0309	0.0012
polygon	ellipse.4	1.6471	3.7464	0.4396
polygon	rectangle	1.4946	3.0354	0.4924
rectangle	triangle	0.6848	1.0749	0.6371

Table 1. Results of matching algorithm.

For polygons and general 2D shapes for which polygonal approximations are generated, another effective method for shape description would be to create a feature vector incorporating angles of vertices and lengths of sides. This information is easily extracted from the chain code as described in section 4. The angles, the lengths of the sides, the relationships between the angles, and the relationships between the sides of a polygon, provide sufficient information on its shape, and thus for its recognition, being invariant to position and orientation.

As well, scale-invariance can be achieved by comparing the ratios of the side lengths. The match can be obtained by determining the nearest neighbor between the test-shape feature vector and the reference-shape feature vectors in the feature space (using the nearest neighbor technique). The matching process of a test shape with a reference shape requires the number of vertices of both the test and the reference shapes to be equal. Thus, it would be most logical that the recognition process be made hierarchical.

The basic problem that all boundary-based recognition techniques face, is variability or "noise". As noted earlier, there are various levels of errors inherent in the results obtained: for example, there is an uncertainty of up to 20 degrees in the angle-measurement process; for ellipses with axes ratios larger than 4, the process leads to minimum MSE values which are lower than expected, leading to probable misclassification.

In general, various factors may contribute to the measurement errors of local curvature of the boundary, and also to the measurement errors of the distance along the contour. There are four main factors that need investigation and improvement in order to increase the accuracy of the measured values: the edge-detection technique and thresholding method, the edge tracing algorithm [19], the size selection of the coding matrix [20], and, the smoothing of the chain-code sequence (with the selection of a rational method for choosing a value for P rather than an empirical one - see section 2). Some of these aspects have been addressed in the literature, but their relevance to the approach presented here must be further pursued.

6. Conclusions

In this paper, a new signature is proposed for 2D shape recognition, namely the gradient-perimeter plot. The method is developed under the constraint of being position-, rotation-, and scale-invariance. For the matching process a cost function which has the property of rotation-invariance is defined. The second term of the proposed cost function accounts for the rotation-invariant aspect of the method. The scale-invariance of the method is achieved by normalizing the length of the shape's perimeter length.

The method has been applied to a set of standard 2D shapes. The experimental results were consistent and satisfactory.

Four sources of error have been identified. The effectiveness of the method presented above can be increased, if errors or their effects can reduced.

As the basis of an alternative method, the angle-perimeter plot can also be used to provide a feature set for 2D shapes. This method can be applied to polygonal shapes or shapes for which polygonal approximations are available. The accuracy of the measured features is highly crucial in the recognition process, in general, and the above-mentioned error sources have to be addressed in this second method as well.

References

[1] Basel, P.J., and Jane, R.C., "Three-Dimensional Object Recognition", *Computer Surveys*, Vol. 17, No. 1, pp. 75-145, March 1985.

[2] Horn, B.K.P., ***Robot Vision***, MIT Press, 1986.

[3] Safaee-Rad, R., Benhabib, B., Smith, K.C., and Zhou, Z., "Pre-Marking Methods for 3D Object Recognition", *IEEE, Int. Conf. on Systems, Man, and Cybernetics*, Nov. 1989. (Submitted)

[4] Pavlidis, T., "Survey: A Review of Algorithms for Shape Analysis", *Computer Graphics and Image Processing*, Vol. 7, pp. 243-258, 1978.

[5] Pavlidis, T., "Algorithms for Shape Analysis of Contours and Waveforms", *IEEE, Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-2, No. 4, pp. 301-312, July 1980.

[6] Attneave, F., "Some Informational Aspects of Visual Perception", *Psychol. Rev.*, Vol. 61, pp. 183-193, 1954.

[7] Geisler, W., "A Vision System for Shape and Position Recognition of Industrial Parts", *Proc., 2nd Int. Conf. on Robot Vision and Sensory Controls*, pp. 253-262, F.R. of Germany, Nov. 1982.

[8] Freeman, H., "Use of Incremental Curvature for Describing and Analyzing Two-Dimensional Shape", *IEEE, Proc., Conf. on Pattern Recognition and Image Processing*, pp. 437-444, Chicago, Aug. 1979.

[9] Yang, H.S., and Sengupta, S., "Intelligent Shape Recognition for Complex Industrial Tasks", *IEEE, Control Systems Magazine*, pp. 23-30, June 1988.

[10] Nehr, G., and Martini, P., "The Coupling of a Workpiece Recognition System With an Industrial Robot", ***Robot Vision***, Alan Pugh, pp. 83-96, 1983.

[11] O'Rourke, J., "The Signature of a Plane", *SIAM, J. Comput.*, Vol. 15, No. 1, pp. 34-51, Feb. 1986.

[12] Dessimoz, J.D., " Visual Identification and Location in a Multi-Object Environment by Contour Tracking and Curvature Description", *Proc., 8th Int. Symp. on Industrial Robots*, pp.764-777, F.R. of Germany, May-June 1978.

[13] Vernon, D., "Two-Dimensional Object Recognition Using Partial Contours", *Image and Vision Comput-*

ing, Vol. 5, No. 1, pp. 21-27, Feb. 1987.

[14] Ballard, D.H., and Brown C.M., ***Computer Vision***, Prentice-Hall, 1982.

[15] Freeman, H., " Computer Processing of Line-Drawing Images", *Computing Surveys*, Vol. 6, No.1, pp. 57-97, March 1974.

[16] Biederman, I., "Human Image Understanding: Recent Research and a Theory", *Comput. Vis. Graphics Image Process.*, Vol. 32, pp. 29-73, 1985.

[17] Resnikoff, H., ***The Illusion of Reality: Topics in Information Science***, Springer-Verlag, New York, 1985.

[18] Oppenheim, A., and Schafer, R., ***Digital Signal Processing,*** Prentice-Hall, Englewood Cliffs, New Jersey, 1975.

[19] Bennett, J.R., and Mac Donald, J.S., "On the Measurement of Curvature in a Quantized Environment", *IEEE, Trans. on Computers*, Vol. C-24, No. 8, pp. 803-820, Aug. 1975.

[20] Freeman, H., and Saghri, A., "Generalized Chain Codes for Planar Curves", *Proc., 4th International Joint Conference on Pattern Recognition*, Kyoto, Japan, pp. 701-703, Nov. 1978.

Acknowledgement

This research was supported in part by the Manufacturing Research Corporation of Ontario (MRCO).