Solving the Bin of Parts Problem

by

J.L. Turney, T.N. Mudge and R.A. Volz

Department of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, Michigan 48109-1109
U.S.A.

Abstract

This paper presents a technique for automatically recognizing parts that are only partly visible. This situation often occurs when parts are stored in a bin. The technique can be used to identify parts (if they are not completely hidden), locate them, and inventory them rapidly and without contact. It can also be used as part of an assembly process in collaboration with one or more robots.

[1] This work was supported in part by a grant from the US Army Research Office under contract DAAG29-84-K-0070.

To appear in the Proceedings of Vision '86, Detroit, Michigan, June 1986.

a 1986

1. Introduction

In this paper we will present a technique for solving the bin of parts problem. We will restrict our solution to parts that are 2-dimensional, untilted, and at a fixed distance from the camera. We refer to a part as 2-dimensional when two of its dimensions are much larger than the third, and we assume that the part is untilted when its smallest dimension is aligned with the viewing axis of the camera. When viewed at a fixed distance isolated parts of this type have only two possible projections, i.e., when the smallest dimension of the part is either parallel or antiparallel to the viewing axis. These projections can be used to easily recognize parts. By comparison, overlapped parts, which have an infinite number of possible projections due to their relative degrees of freedom, are much more difficult to recognize. When parts are overlapped, less of each part is exposed. As a consequence, the exposed portion of a part in the presence of noise is often sufficiently distorted to make it unrecognizable, and the recognition algorithm may fail to locate the part from its partial projection. Alternatively, the exposed area of the part may appear to belong to another part or to another section of the part, and thus the part may be either incorrectly located or its pose (position and orientation) may be incorrectly determined. Again the algorithm has failed. While it true that no algorithm can be designed to locate all the parts in the set of possible projections (in many projections some parts will be totally occluded), it is, however, desirable to design an algorithm that will maximize the number of correctly located parts while at the same time minimizing the number of incorrectly located parts.

The problem we describe has received considerable attention in the literature where it is commonly referred to as the 2-dimensional "partially occluded parts" or "bin of parts" problem. Approaches to solving the problem have been proposed by [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20]. Several of the more recent works [10, 11, 18, 20] have proposed the use of *segment* matching for recognition. In these approaches *edge points*, which are locations in the image of rapid intensity changes, are linked together to form *edge boundaries*. Edge boundaries are partitioned into segments, which are lists of contiguous edge points taken from the edge boundary. Parts in an image are located by matching segments of the edge boundaries of the parts to segments of the edge boundaries of the image. This strategy makes sense for the following reasons. Edge boundaries of 2-dimensional parts contain most of the structural information of the part, but require less storage than an image of the part. Edge boundaries are more robust to changes to illumination than are the images from which they are derived. Finally, segments are used for matching because at least some can be expected to be exposed in an image. (Figure 1a shows the boundary of a part and Fig. 1b shows the segments of the exposed part.)

2. Our Approach

In order to maximize the number of correctly located parts, the recognition algorithm must be designed to use as many segments as possible of the part's boundary for recognition. In addition, the algorithm should recognize a part with a minimum of the part exposed. Our approach works with boundaries of parts that are partitioned into fixed length segments. Segments are selected so that one segment is centered about each edge point on a boundary. This selection leads to a set of segments such as AA', BB' and CC' shown in Fig. 2a. By using overlapping segments, none of the segments is exluded from use in recognition. This adds robustness to the recognition algorithm because if one segment is partially occluded, there is a good chance that a neighboring segment will be totally exposed. To minimize the boundary used in recognition the length of the segment should be chosen to be as short as possible. On the hand, the length should be chosen to be long enough for many segments to have sufficient structure to make them distinguishable. Presently a compromise length is chosen by trial and error based on the parts to be located.

In order to minimize the number of incorrectly recognized parts the recognition algorithm must be designed to unambiguously interpret the exposed segments of the part's boundary. While a segment, by itself, may not provide sufficient recognition we have found that a configuration of two segments which is unique to the part can usually provide the required level of reliability. A configuration, as we define it, is simply two boundary segments in a fixed relative pose. The relative pose of two segments can be defined by the vectors from the mid-points of the segments to the centroid of the part's boundary from which the segments are taken (see Fig. 2b).

In previous work we used a generalized Hough transform strategy to locate parts from the set of edge boundary segments of the part [12, 13, 17]. The approach was basically a weighted template matching scheme. A weight, that we called "saliency," was associated with each segment of the part. This weight was used to emphasize segments which were important in recognizing the part and its pose while deemphasizing those that were not. Once a set of parts was selected, the weights were determined automatically in an off-line stage by comparing all of the segments of the set of parts and by assigning those that appeared frequently with little weight and those that appear infrequently with large weight. The weights were constrained to sum to 1 and an quadratic optimization algorithm was used to adjust the weights. During the on-line stage the weighted template was matched to the image. This proceeded by matching each segment of the template to each segment of the image boundary. Every pose of the part was assigned the accumulated weight of all of the weights of the segments of the template that matched segments of the image boundary at the pose. The pose with the most accumulated weight was selected as the pose of the part. Experiments with this approach showed it to be robust, but slow. Every segment of the part was compared to every segment of the image boundary. After implementing this approach it was realized that in most cases only a pair of part segments, i.e., a configuration (see [18]), was needed to recognize any part.

In [21] we redefined saliency from its meaning in [17] to refer to the probability that a configuration belongs to a particular pose of a part. Saliency was taken informally as the inverse of the frequency with which a configuration occurs within the set of parts that are to appear in the image (see [21] for a more formal treatment). We refer to the unique configurations, i.e., those that have an inverse freqency of 1, as the most *salient* configurations or for the purposes of the following discussion as simply the salient configurations.

To determine the set of salient configurations for a part, every configuration of the part is compared to every configuration of the set of parts that are appear in the image. If any configuration appears more than once in the part or appears even once in any of the other parts, it is discarded. The configurations that remain are considered the salient configurations. Membership in this set obviously depends on the set of parts to appear in the image. For example, if the obverse side of the part shown in Fig. 3a were the only part to appear in an image, configuration A and B would form a salient configuration. If, however, the reverse side is included in the set of part boundaries (see Fig. 3b), the configuration A and B strongly resembles configuration A' and B'. A and B, therefore, no longer uniquely recognize the part's pose and are discarded from the set of salient configurations. However, A and C still are a salient configuration because they are found only once in the set of part boundaries. It is obvious that enlarging the set of part boundaries will reduce the membership of the set of salient configurations for the obverse side of the part even further.

To locate a part, a segment that is a member of many salient configurations of the part is selected. Segment A in Fig. 4 is such a section. Segment A together with the segments whose centers are marked with vertical spikes all uniquely recognize the part and its pose. This segment is compared to the segments of the image boundary. If a good match is found, the rotation and translation necessary to align the two segments is computed. The rotation and translation are applied to the entire boundary of the part (see dashed boundary in Fig. 5a) and the transformed boundary is used as a guide in searching for the second segment, B, of a salient configuration. If such a segment is in turn found (see Fig. 5b), the part is assumed to be located. If not, the process of selecting an initial segment and then searching for a second segment that forms a unique configuration is repeated. We have assumed in the above that the segments in a configuration found in the image come from the same part, and not from the *accidental alignment* of segments of two or more parts. More precisely, an accidental alignment occurs when a configuration of segments from two or more parts happens to fall in a relative position that resembles a configuration of segments from a single part. We assume this occurs with very low probability.

Segments of the part and image are compared in a slope angle-arclength representation. We refer to this representation as the θ–s representation. The θ–s representation is a one dimensional functional representation, $\theta(s)$, of an edge boundary. The slope angle at each edge point is parameterized by the arclength along the boundary from an arbitrary starting point to the edge point (see Fig. 6).

In the θ–s representation segments are fit with one parameter, the relative orientation between segments, Θ. To compare a segment of a part with a segment of the image boundary, we select the sum of the squares of the differences between corresponding slope angles as a measure of the closeness of the fit. The centers of the segments are aligned (see Fig. 7) and the θ values of the part's segment, $\theta_M(s_i)$ for $i = [-n, \cdots, n]$, are least squares fit to the corresponding θ values of the image segment, $\theta_I(s_i)$ for $i = [-n, \cdots, n]$. We assume that both segments have been sampled at equal arclengths at n points on either side of their centers. The fit parameter, Θ, (see Fig. 7) is chosen to minimize the following

$$\frac{1}{2n+1} \sum_{i=-n}^{n} (\theta_M(s_i) - \theta_I(s_i) - \Theta)^2.$$

The minimum occurs when

$$\Theta = \frac{1}{2n+1} \sum_{i=-n}^{n} (\theta_M(s_i) - \theta_I(s_i)) = \bar{\theta}_M - \bar{\theta}_I,$$

in other words, when Θ is simply the difference between the mean tangent angles of the two segments. The minimum residue

$$R = \left[\frac{1}{2n+1} \sum_{i=-n}^{n} (\theta_M(s_i) - \bar{\theta}_M - (\theta_I(s_i) - \bar{\theta}_I))^2 \right]^{\frac{1}{2}},$$

is used as a measure of the similarity of the segments to decide whether the segments match. We assume that two segments match if R is less than a fixed threshold, D, where the value of D is chosen to reflect the noise anticipated in the images under consideration.

Besides its simplicity, this method of comparison has the addition advantage that it allows some latitude in matching segments. It is not always possible to obtain precise estimates of the arclength, s, along the boundary, and with the sum of squares of the differences as a measurement of similarity, small distortions in the s values do not affect the measurement as much as they would, say, a fit that minimizes the maximum difference.

Another feature of using the θ–s representation is that it allows one to easily calculate critical points. We define critical points as the extrema of the curvature of the boundary that have absolute curvatures above a fixed threshold. The curvature of a contour is the first derivative of $\theta(s)$ with respect to s, or $\frac{d\theta(s)}{ds}$ (see [22]). The locations of critical points in the contour are readily obtained by applying a 1-dimensional edge detector to the function $\theta(s)$. Figure 8a shows the critical points of a part; curvature maxima are shown as circles and curvature minima are shown as squares. Figure 8b shows the critical points of the boundary in an image. Note the correspondence of critical points.

Critical points in a boundary, if they exist, can be used to further improve the efficiency of comparison. If a segment of the part contains critical points, as is often the case, it need only be compared to segments in the image that contain similar critical points. By comparing a part segment to only those image segments with similar critical points, we substantially reduce the number of comparisons needed to locate matches between segments.

3. Results

Figures 9a-h show some of the results obtained for locating parts. Fourteen images of seven parts were used with a success rate of approximately 5.8 parts found per image (83%). Recognition times for the seven parts were on the order of 20 to 30 seconds on an Apollo 660 node. This value does not include preprocessing such as edge detection and edge linking.

The plot in Fig. 10 is perhaps more meaningful. It shows the percent of the parts recognized in the images versus the percent of the part's boundary exposed. The numbers in the figure adjacent to each point show the number of instances supporting the data. One can see that we obtained reasonable results even when as little as a third of the part's boundary was exposed. Of course, these results depend on the part chosen, but reflect the robustness of the approach.

4. Implementation

In the present implementation, pixel wide edges are extracted from the images using the Canny edge detector [23]. The slope angle, θ, at each edge point in determined from the edge strengths at locations of the edge boundaries. More precisely, the slope angle is taken to be the arctangent of the of the ratio of strength in the y direction to the strength in the x direction. The edge points are then linked to form contours which are normalized by unwrapping artificial discontinuities in the $\theta(s)$ function due to the branch cut in the arctangent function and by resampling the boundaries at unit arclengths. The θ–s representation is processed with a one dimensional Canny edge detector to located critical points of the boundary. Salient configurations are determine by individually matching the segments of the configuration to all the other segments in the set of parts. Then it is determined if two segments of a configuration ever match at the same relative position and orientation to other segments of the part or to other segments of other parts. If they do, the configuration is not considered salient.

5. Conclusion

In the example throughout the paper, we have used a bin of identical parts since this is a common mode of presentation of parts in a bin. The algorithm, however, is not limited to working with identical parts and, in fact, can be applied to any bin of parts. (See [18], where earlier versions of the algorithm were demonstrated on a set of nine different parts.) We believe that the approach that we have suggested is both robust and efficient enough to be used as a solution to the 2-dimensional partially occluded parts problem. In the future the algorithm will be modified to handled scaled 2-dimensional untilted parts and reflective parts. We also believe that our use of salient configurations is extendable to other problems, notably, the 3-dimensional occluded parts problem.

[1] A. P. Ambler, H.G. Barrow, C. M. Brown, R. M. Burstall, and R. J. Popplestone, "A Versatile Computer-Controlled Assembly System ," *Proceedings of the 3rd International Joint Conference on Artificial Intelligence*, pp. 298-307, August 1973.

[2] J. W. McKee and J. K. Aggarwal, "Computer Recognition of Partial Views of Curved Objects," *IEEE Transactions on Computers*, vol. C-26, no. 8, pp. 790-800, August 1977.

[3] W. A. Perkins, "A Model-based Vision System for Industrial Parts," *IEEE Transactions on Computers*, vol. C-27, no. 2, pp. 126-143, February 1978.

[4] J. D. Dessimoz, M. Kunt, and J. M. Zurcher, "Recognition and Handling of Overlapping Industrial Parts," *9th International Symposium on Industrial Robots*, pp. 357-366, March 1979.

[5] D. H. Ballard, "Generalizing the Hough Transform to Detect Arbitrary Shapes," *Pattern Recognition*, vol. 13, no. 2, pp. 111-122., 1981.

[6] W. S. Rutkowski, "Recognition of Occluded Shapes Using Relaxation," *Computer Graphics and Image Processing*, vol. 19, pp. 111-128, 1982.

[7] R. B. Kelly, H. A. S. Martins, J. R. Birk, and J. D. Dessimoz, "Three Vision Algorithms for Acquiring Workpieces from Bins," *Proceedings of the IEEE*, vol. 71, no. 7, pp. 803-820, July 1983.

[8] Jakub Segen. "Locating Randomly Oriented Objects from Partial View," *Proceedings of the Society of Photo-optical Instrumentation Engineers Cambridge Symposium on Optical and Electro-optical Engineering*, November 1983.

[9] R. C. Bolles and R. A. Cain, "Recognizing and Locating Partially Visible Objects: The Local-Feature-Focus Method," in *Robot Vision*, Pugh, Ed, 1984.

[10] N. Ayache and O. D. Faugeras, "A New Method for Recognition and Position of 2-D Objects," *Proceedings 7th International Conference on Pattern Recognition,*, vol. 2, pp. 1274-1277, Aug. 1984..

[11] W.E. Grimson and T. Lozano-Perez, "Model-based recognition and localization from sparse range or tactile data," *International Journal of Robotics Research*, vol. 3, no. 3, 1984.

[12] J. L. Turney. T. N. Mudge, R. A. Volz, and M. D. Diamond, "Experiments in occluded parts recognition using the generalized Hough transform," *Proceedings of the Conference on Artificial Intelligence*, April 1983.

Turney, Mudge & Volz

[13] J. L. Turney, T. N. Mudge, and R. A. Volz, "Experiments in Occluded Parts Recognition," *Proceedings of the Society of Photo-optical Instrumentation Engineers Cambridge Symposium on Optical and Electro-optical Engineering*, November 1983.

[14] P. Rummel and W. Beutel, "Workpiece Recognition and Inspection by a Model-Based Scene Analysis System," *Pattern Recognition*, vol. 17, no. 1, pp. 141-148, 1984.

[15] C. K. Cowan, D. M. Chelberg, and H. S. Lim, "ACRONYM Model Based Vision in the Intelligent Task Automation Project," *1st Conference on AIA*, pp. 176-183, December 1984.

[16] B. Bhanu and O. D. Faugeras, "Shape Matching of Two-Dimensional Objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 2, pp. 137-156, March 1984.

[17] J. L. Turney, T. N. Mudge, and R. A. Volz, "Recognizing Partially Occluded Parts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, July, 1985.

[18] J. L. Turney, T. N. Mudge, and R. A. Volz, "Recognizing Partially Hidden Objects," *IEEE International Conference on Robotics and Automation*, March, 1985.

[19] M. W. Koch and R. L. Kashyap, "A Vision System to Identify Occluded Industrial Parts," *IEEE Conference on Pattern Recognition and Image Processing*, pp. 55-60, 1985.

[20] T.F. Knoll and R. Jain, "Recognizing partially visible objects using feature indexed hypothesis," *IEEE J. Robotics & Automation*, 1985.

[21] T. N. Mudge, J. L. Turney, and R. A. Volz, "Automatic Generation of Salient Features for the Recognition of Partially Occluded Parts," *Robotica (to appear)*.

[22] M. M. Lipschutz, in *Schaum's Outline of Theory and Problems of Differential Geometry*. New York: McGraw-Hill, p. 73, 1969.

[23] J. F. Canny, "Finding Lines and Edges in Images," *Master's Thesis, MIT*.

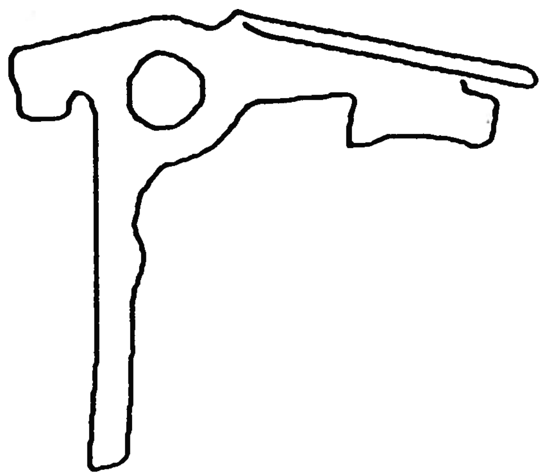

Fig. 1a

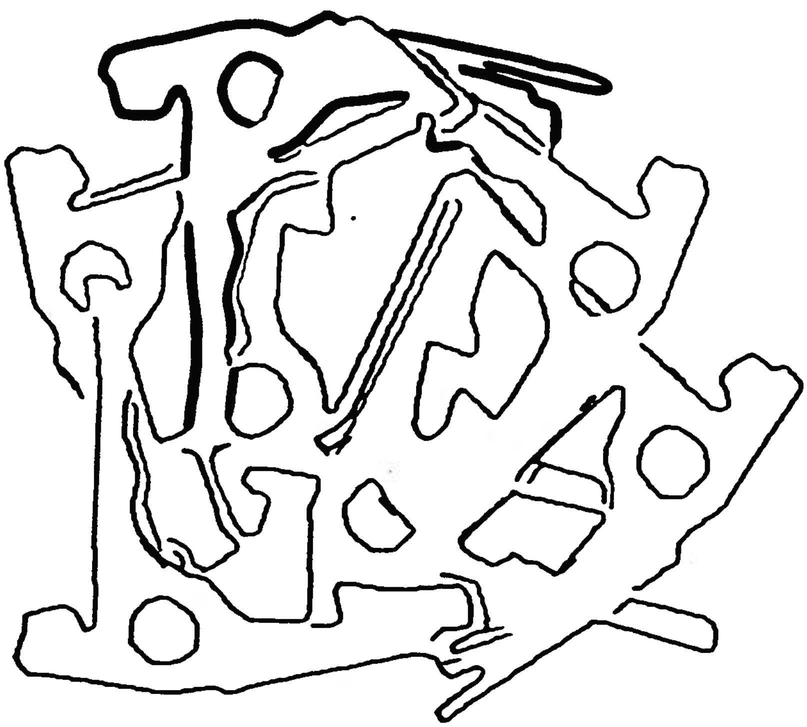

Fig. 1b

Segments of exposed part boundary.

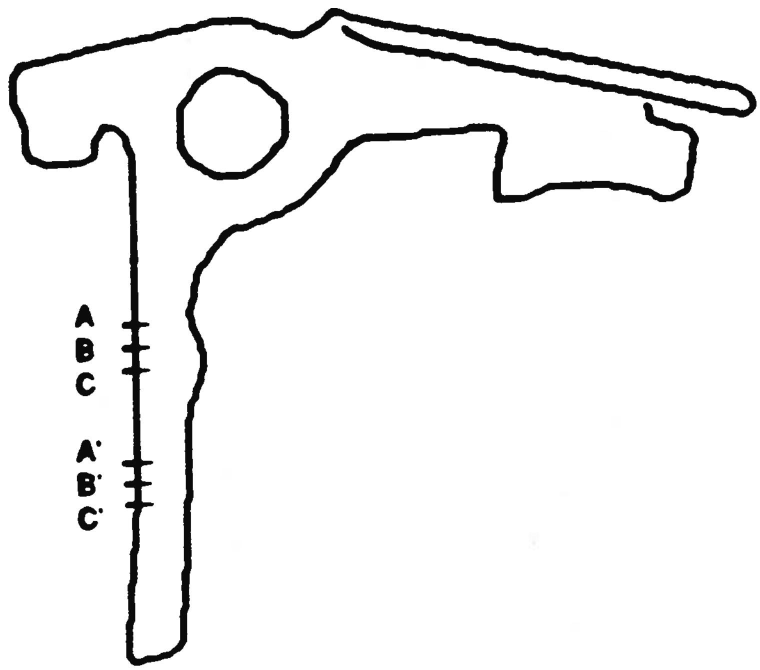

Fig. 2a. Selected segments.

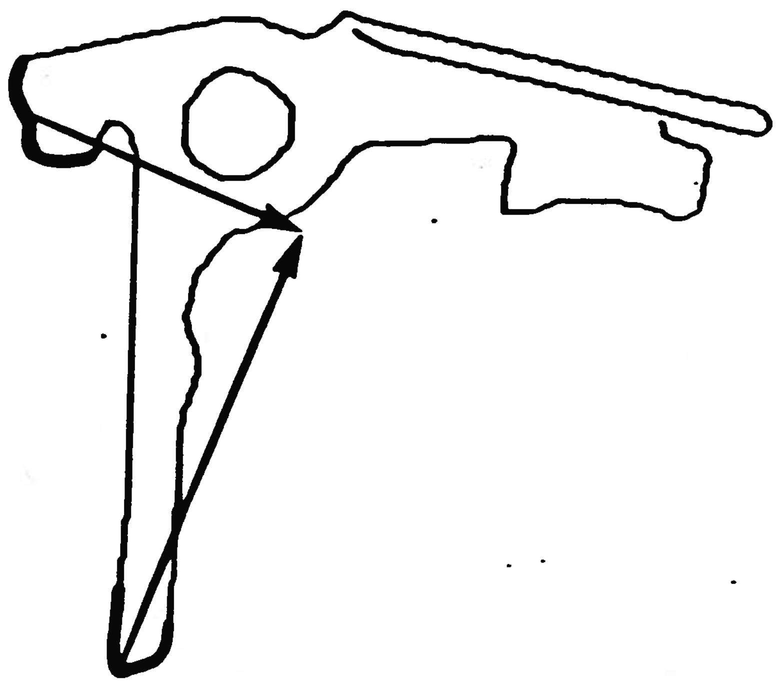

Fig. 2b. Configuration of segments.

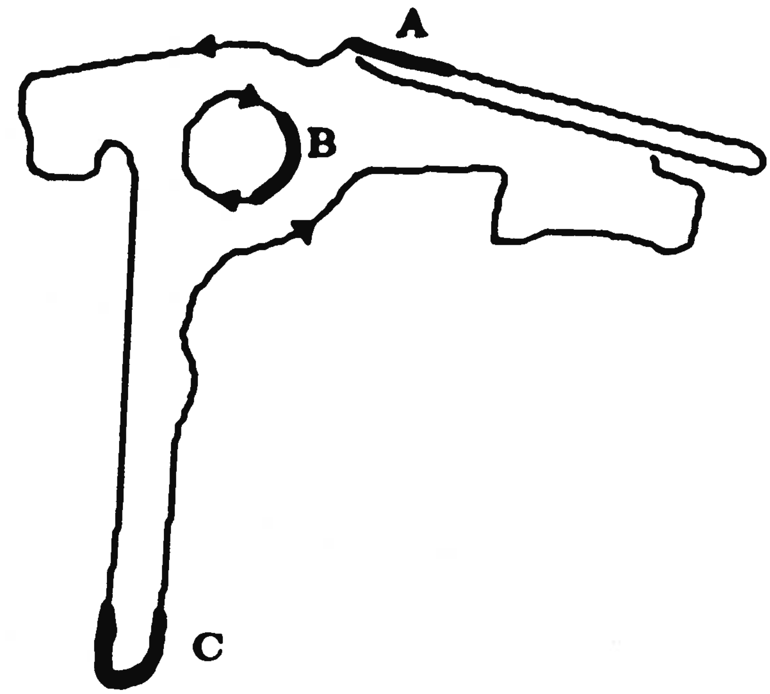

Fig. 3a. Observe side.

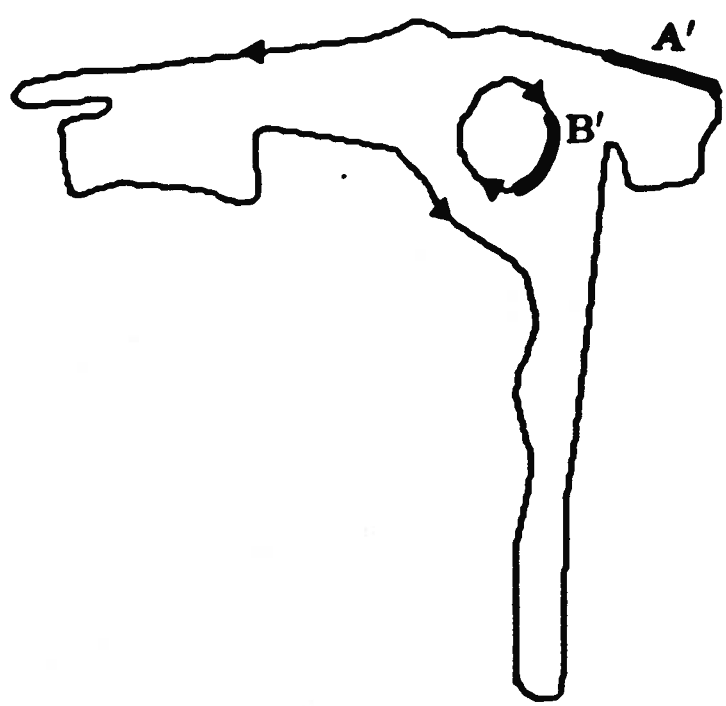

Fig. 3b. Reverse side.

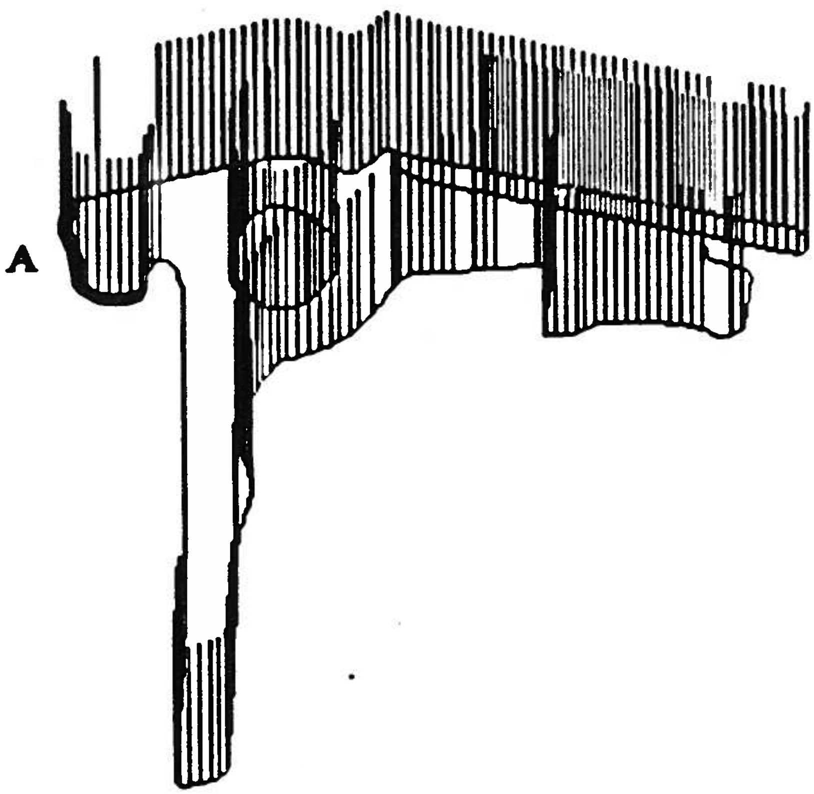

Fig. 4. Configurations associated with a segment.

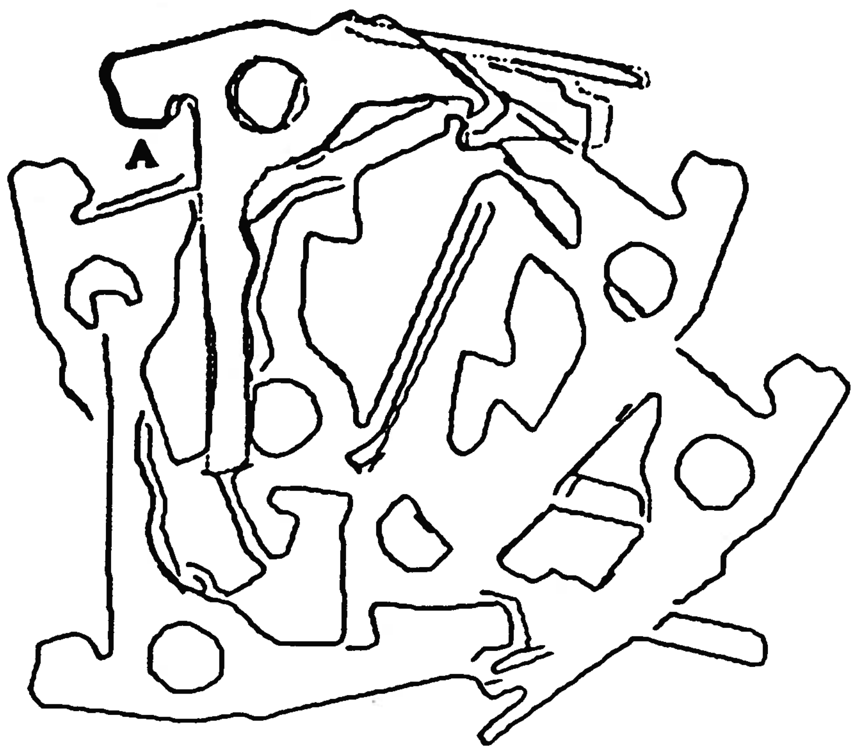

Fig. 5a

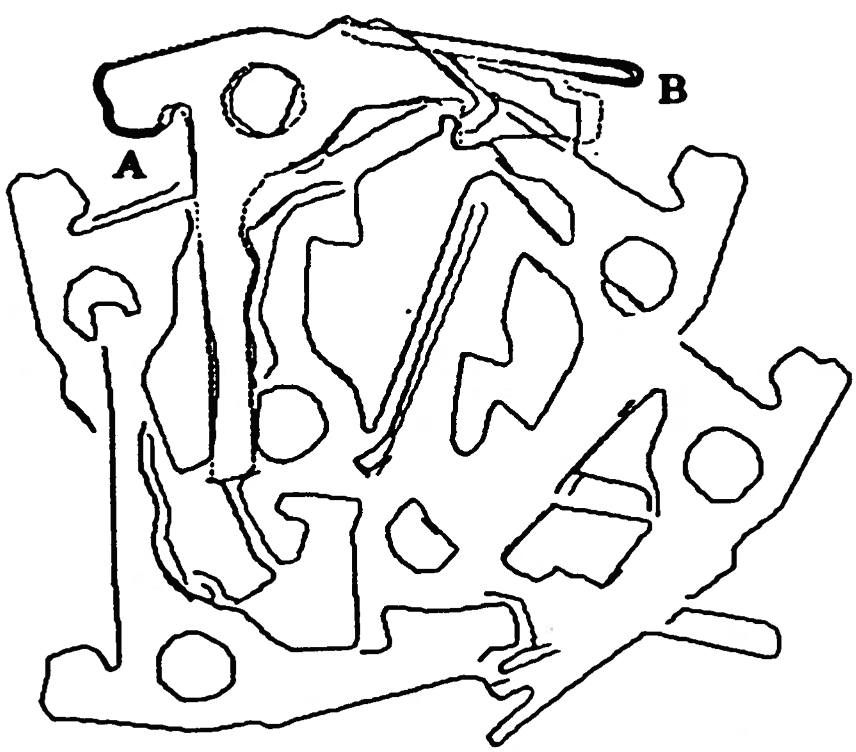

Fig. 5b

Locating a part

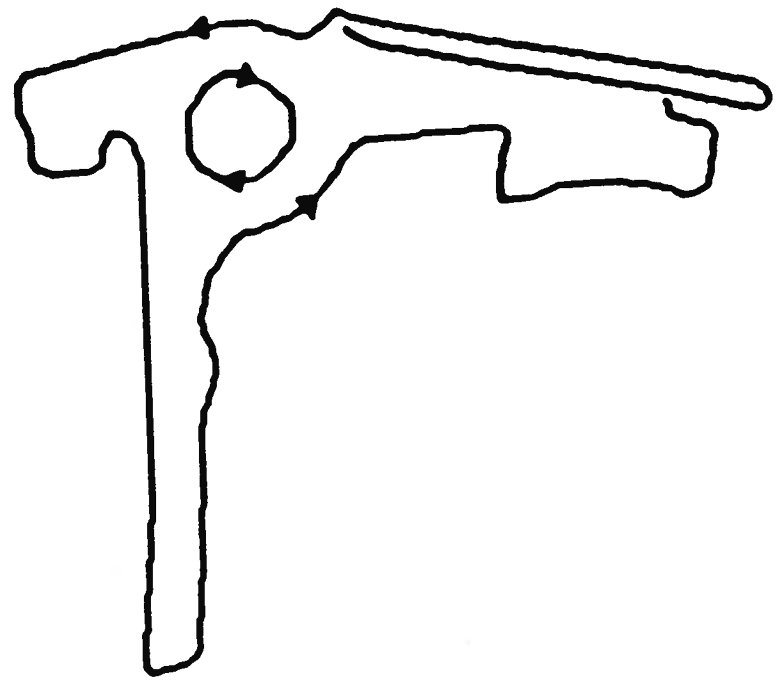

Fig. 6a

Fig. 6b

Boundary of part and its θ – s representation

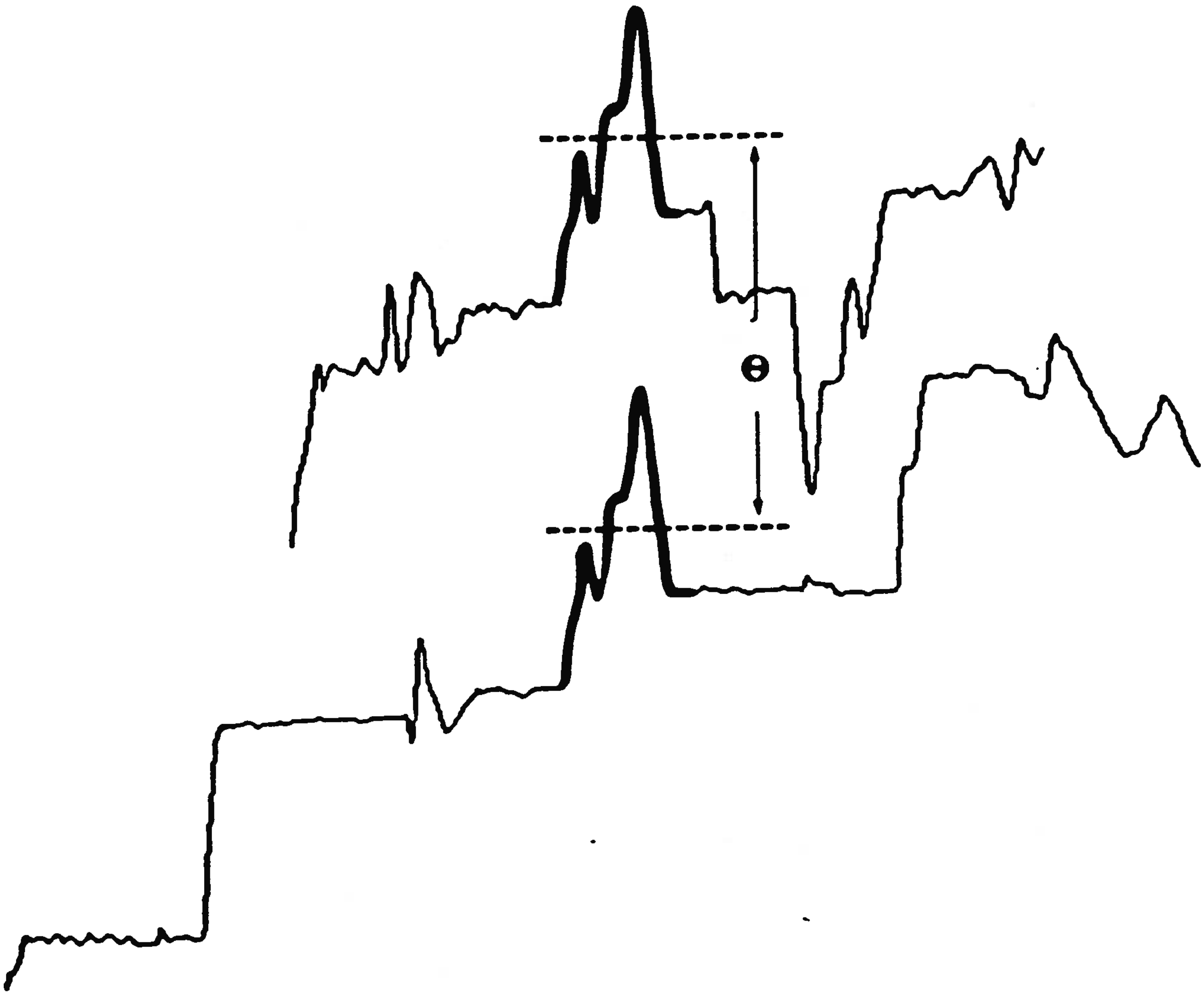

Fig. 7. Comparing segments in θ – s space.

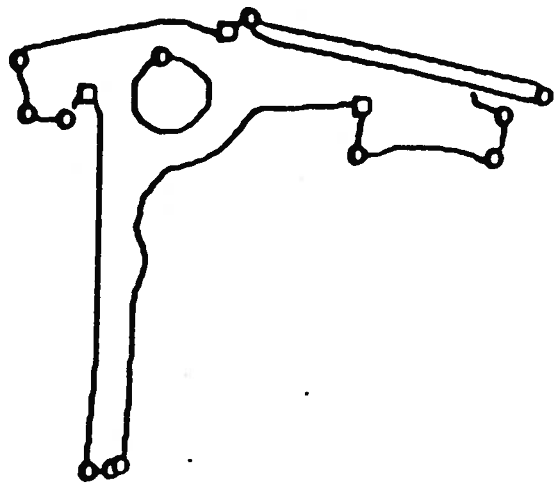

Fig. 8a

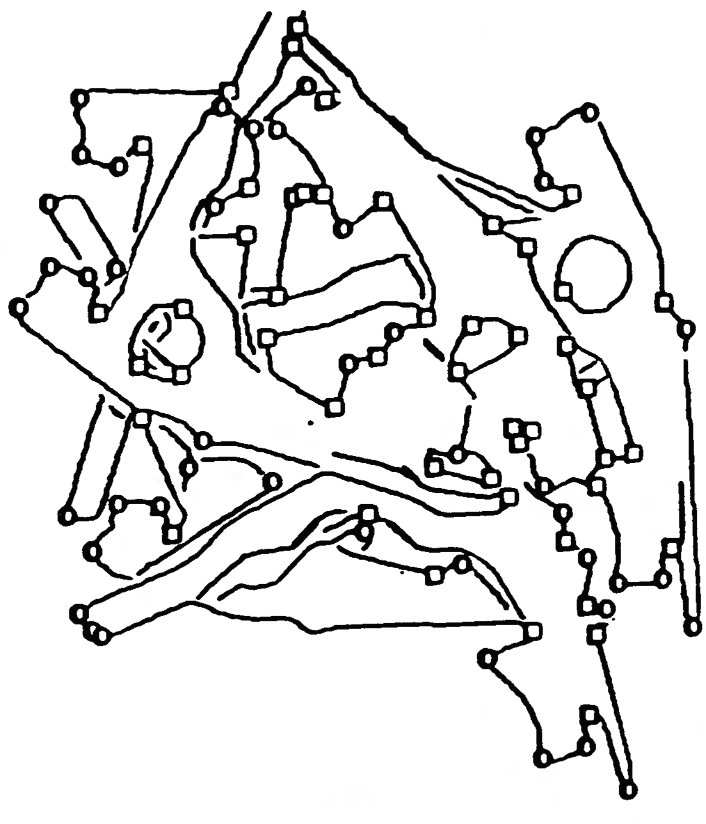

Fig 8b

Critical points in the part and image boundaries.

Fig. 9a

Fig. 9b

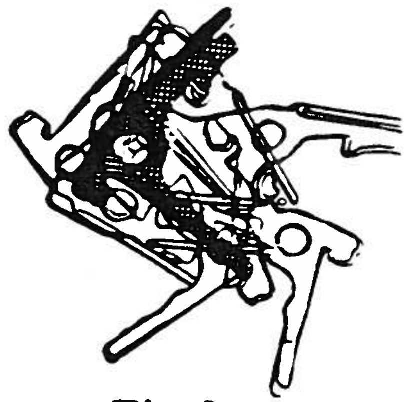

Fig. 9c

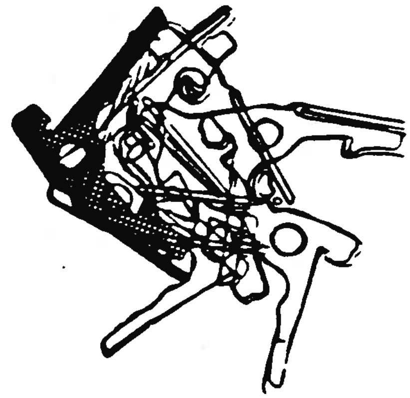

Fig. 9d

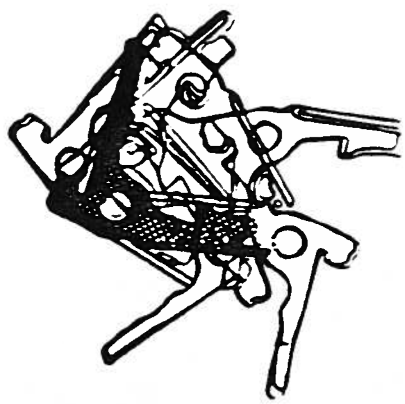

Fig. 9e

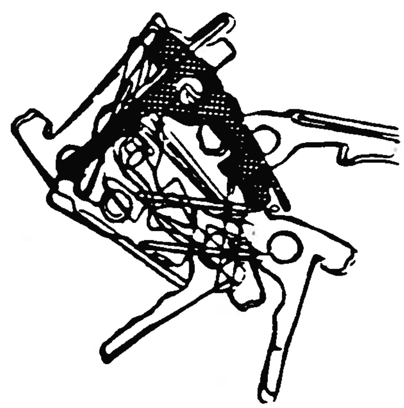

Fig. 9f

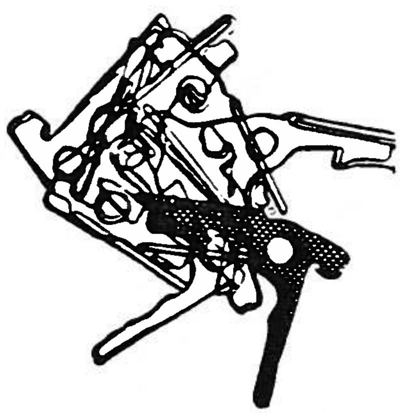

Fig. 9g

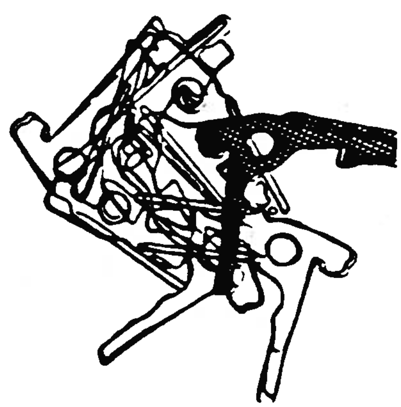

Fig. 9h

Parts found.

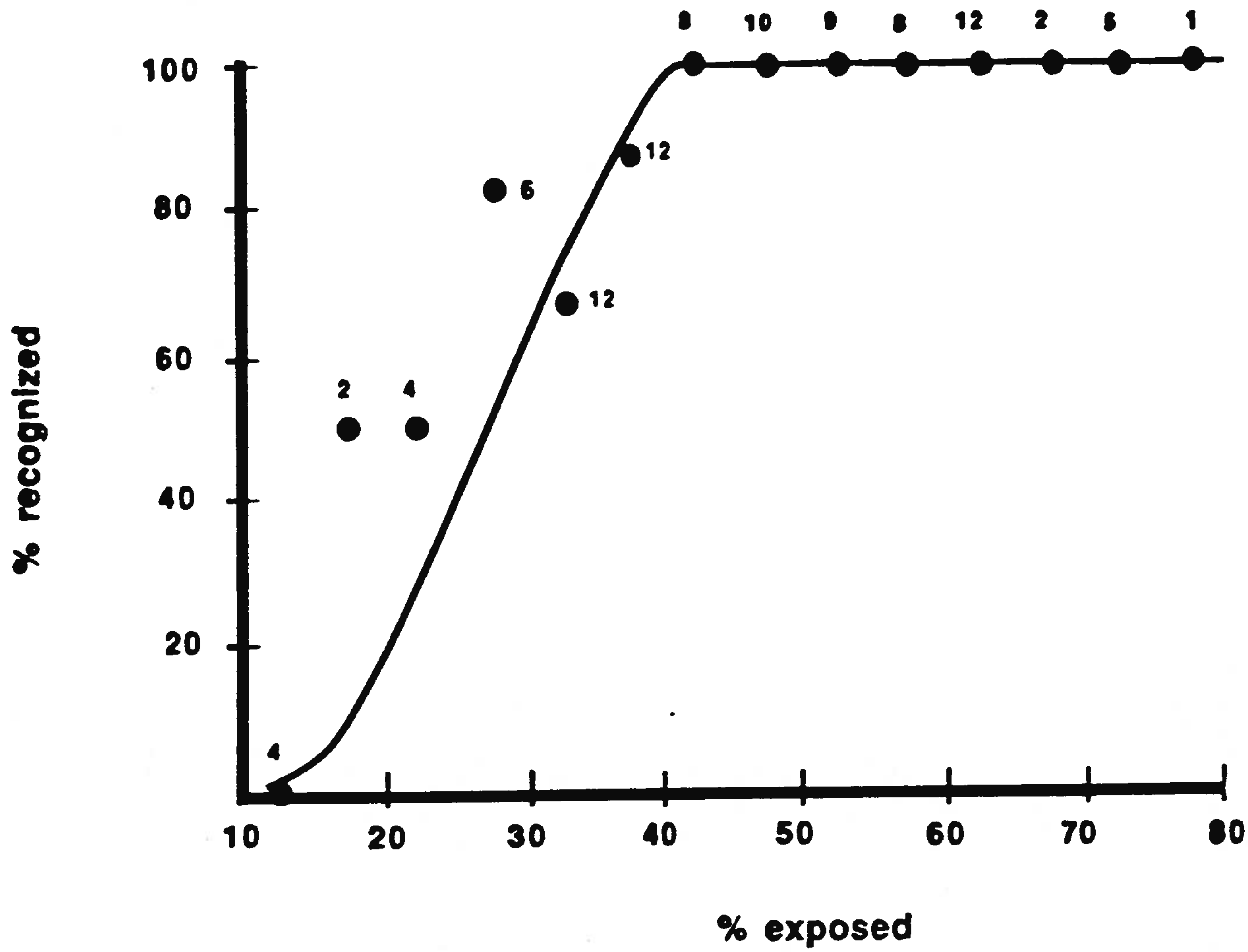

Fig. 10. Percent of parts recognised vs. percent of part exposed.