

# Improving the generalized Hough transform through imperfect grouping

Clark F. Olson<sup>1\*</sup>

*Jet Propulsion Laboratory, California Institute of Technology, Mail Stop 107-102, 4800 Oak Grove Drive, Pasadena, CA 91109, USA*

Received 10 February 1997; accepted 13 October 1997

---

## Abstract

This paper analyses the improvements that can be gained in the generalized Hough transform method for recognizing objects through the use of imperfect perceptual grouping techniques. In particular, we consider simple grouping techniques that determine pairs of points that are likely to belong to the same object using a criterion based on connectedness in the image edge map. It is shown that such imperfect grouping techniques can considerably improve both the efficiency and accuracy of object recognition. Experiments are described that demonstrate the improvements in performance. © 1998 Elsevier Science B.V. All rights reserved.

*Keywords:* Object recognition; Generalized Hough transform; Perceptual grouping

---

## 1. Introduction

Object recognition methods for complex problems have been plagued by both poor speed and a high incidence of false alarms. The primary source of both of these problems is clutter in the image. Such image clutter not only requires considerable computation to process, but also yields false positive instances of the objects that are sought. Solutions to these problems have been only partially successful. Two methods that have been useful are perceptual grouping and indexing. Grouping methods (e.g. [1–7]) attempt to determine which features in an image belong to the same object, while indexing methods (e.g. [8–12]) determine which sets of model features may have projected to various sets of image features. These methods can be powerful when they are used together [9].

Several studies have analysed the power of indexing, by itself, as a means to reduce the search space for object recognition [8,9,12–14]. It has been demonstrated that, when object recognition is attempted using image features that contain a localization error, indexing systems index, on average, a constant fraction of all of the possible matches for a particular set of image features. Thus, while such indexing systems can yield a large constant speedup, they do not

reduce the computational complexity of recognition in the presence of localization error.

Perceptual grouping, on the other hand, can reduce the computational complexity of object recognition. Grimson [15] has shown that the expected complexity of a particular constrained search system for object recognition is exponential in the problem size if we have spurious features, but it is low-order polynomial if the data is known to belong to a single object model. However, the requirement that all of the data belong to a single object implies the need for a perfect grouping system, unless we limit ourselves to very simple problems. This paper demonstrates that even simple, imperfect grouping techniques can both improve the computational complexity of object recognition using the generalized Hough transform [16] and improve the accuracy of recognition by decreasing the false alarm rate.

We first review an efficient variation of the generalized Hough transform that uses feature points to recognize three-dimensional objects from a single view in two-dimensional images [17]. This technique first divides the recognition problem into many small subproblems that must be examined. Randomization is used to limit the number of subproblems that must be examined. Finally, each subproblem is solved using efficient multi-dimensional histogramming techniques. The complexity of this method is  $O(mn^3)$ , where  $m$  is the number of feature points in the model and  $n$  is the number of feature points in the image.

We next discuss a simple perceptual grouping mechanism that is used in conjunction with this generalized Hough transform variation to further improve the performance of

---

\* Tel.: 001 818 354 0638; Fax: 001 818 354 8172;

e-mail: olson@robotics.jpl.nasa.gov

<sup>1</sup>This research was performed while the author was at Cornell University and was supported in part by ARPA under ARO contract DAAH04-93-C-0052 and by the National Science Foundation under PYI grant IRI-9057928.

object recognition. Many cues have been used to perform perceptual grouping of image features. Some examples include parallelism, proximity, colinearity, connectivity, convexity, symmetry and closure. In this paper, we consider a simple grouping mechanism that locates pairs of points that are likely to belong to the same object. While proximity, or even color and texture, can provide such information, we concentrate on the use of connectivity of the feature points in the image edge map.

The computational complexity of recognition using these techniques is then examined. It is shown that, under general assumptions about the performance of the grouping mechanism, the complexity of recognition using the generalized Hough transform is reduced to  $O(mn^2)$ . Further, an analysis of the rate at which false positives are found by these techniques, based on previous work [17–20] using a Bose–Einstein occupancy model, indicates that performance is considerably improved. We also discuss experiments, where these techniques have been applied to real images, that support these results. Finally, we summarize the contributions of this work. A preliminary version of this work appeared in the 1995 International Symposium on Computer Vision [21].

## 2. Recognition framework

Variants of the generalized Hough transform [16] have been applied to many object recognition problems (e.g. [22–29]) and these techniques have sometimes been called *pose clustering* in this context. The key idea in these applications is that, when many sets of matches between image features and object model features are mapped into the space of object positions that bring them into alignment, a cluster forms at the correct position of the object, if it appears in the image.

For each problem, there is some minimum number of matches between image features and object model features such that there exist only a finite number of object positions that bring the matches into alignment (under the assumption that there is no localization error and the points are in general position). Most recognition systems based on the generalized Hough transform examine each possible matching between image features and object model features of the appropriate cardinality, map them into the pose space, and

then perform some clustering procedure (usually multi-dimensional histogramming) to determine likely object positions. A verification stage often follows for each of the clusters that is found.

For the case of recognizing three-dimensional objects in two-dimensional images using feature points, the number of matches that are required to constrain the position of the object to a finite set in the errorless case is three [30,31]. If all of the matches between three object model feature points and three image feature points are considered, then  $O(m^3n^3)$  poses must be computed, where  $m$  is the number of model points and  $n$  is the number of image points.

Previous work [17] has shown that, if a geometrically precise clustering operation is used that determines the object positions that bring some number of matches between model points and image points into alignment up to some error criterion (e.g. [32,33]), then object recognition techniques based on clustering of poses can achieve the optimal accuracy for feature matching based on this criterion. Suboptimal algorithms based on multi-dimensional histogramming techniques are usually used for efficiency reasons. This work also demonstrated that the complexity of such pose clustering techniques can be reduced to  $O(mn^3)$  and the memory usage improved through the use of decomposition and randomization techniques, if an arbitrarily small chance of failure is allowed. We summarize this method here.

The premise for the decomposition of the problem into many smaller subproblems is the following observation. If we map all of the sets of three matches between object model points and image points that share some pair of matches into the parameter space and locate a model position that brings  $x - 2$  of them into alignment up to some error criterion for the individual matches, then this is equivalent to finding  $x$  individual matches that are brought into alignment by the position up to the error criterion [17]. Each subproblem that we examine thus considers only those triples of matches that share some pair of matches between points in the object model and in the image. If enough initial pairs of matches are examined, then we can locate all of the objects in the image with a small chance of failure.

Fig. 1 gives the new algorithm. Each iteration through the main loop selects a pair of *distinguished points*,  $v_1$  and  $v_2$ , that must belong to the object model for the iteration to succeed. Each of the  $2m(m - 1)$  permutations of possibly

1. **Algorithm** *recognize*( $M, I$ ) /\*  $M$  is the model point set.  $I$  is the image point set. \*/
2.   **Repeat:**
3.     Choose two random image points,  $v_1 \in I$  and  $v_2 \in I$ , with  $v_1 \neq v_2$ .
4.     **For all** pairs of model points,  $\mu_1 \in M$  and  $\mu_2 \in M$ , with  $\mu_1 \neq \mu_2$ :
5.       **For all** point matches,  $(\mu_3 \in M, \nu_3 \in I)$ , with  $\mu_1 \neq \mu_3, \mu_2 \neq \mu_3, v_1 \neq \nu_3$ , and  $v_2 \neq \nu_3$ :
6.         Determine the poses aligning the set of point matches,  $\{(\mu_1, v_1), (\mu_2, v_2), (\mu_3, \nu_3)\}$ .
7.     **End-for**
8.     Find and output clusters among these poses.
9.   **End-for**
10. **End-repeat**
11. **End**

Fig. 1. Efficient pose clustering algorithm.

matching model points,  $\mu_1$  and  $\mu_2$ , are considered as matches for the distinguished points. For each such pair of possible matches, we form all of the sets of three point matches between the object and the image that contain the pair of matches, and determine the poses that bring each triple of matches into alignment. The clusters among these poses are then determined and output.

If each possible pair of distinguished points is considered, then this algorithm has essentially the same performance as the conventional pose clustering method and requires  $O(m^3n^3)$  time. Randomization can be used to reduce the number of pairs of distinguished points that must be examined while maintaining a low rate of failure. If  $\frac{n^2}{(fm)^2} \ln \frac{1}{\delta}$  such pairs are examined, where  $f$  is the minimum fraction of the object model feature points that we require to appear in the image to recognize it reliably, then a rate of failure below  $\delta$  is achieved [17]. Approximately  $\frac{n^2(n-2)(m-2)}{f^2} \ln \frac{1}{\delta}$  poses are computed overall using this method.

The overall time required by the algorithm is  $O(mn^3)$ , since the clustering step is performed in linear time using recursive histogramming techniques.

Further, an error analysis [17] based on work by Grimson and Huttenlocher [18–20] determined an estimate of the probability that a bin in the pose space will yield a false positive of size  $t$  as a result of the random accumulation of feature points. If  $b$  is the average fraction of pose space that brings a model point into alignment with an image point up to some error criterion, then this estimate is:

$$p \approx \left( \frac{bmn}{1 + bmn} \right)^t. \quad (1)$$

If a threshold of  $fm$  on the number of consistent matches is used to determine whether an object hypothesis has been found, then the maximum number of image feature points that can be tolerated with a false alarm rate no greater than  $\gamma$  is:

$$n_{\max} \approx \frac{f}{b \ln \frac{1}{\gamma}}. \quad (2)$$

### 3. A simple grouping mechanism

Grouping image features into sets that are likely to belong to the same object can considerably improve the efficiency of object recognition and reduce the rate at which false positives are found. It is possible for such techniques to not only distinguish sets of points that are likely to belong to the same object, but also to reduce the amount of search that is necessary in matching points between an object model and points in the image that are known to be from the object, by producing only certain subsets of the points among all possibilities. See [34] for a review of perceptual grouping techniques.

Perceptual grouping techniques have been integrated into several recognition systems (e.g. [2,35,36]) and have certainly contributed to the performance of these systems. However, the improvements that have been gained through grouping have not been fully quantified. One goal of our work is to quantify the improvements that can be gained, both in efficiency and in the rate that false positives are observed, through the use of perceptual grouping techniques.

We briefly mention here some results that exist in this area. Grimson [15] has shown that the expected complexity of a constrained search system for object recognition is exponential in the number of image features, in general, but when the data is known to belong to a single object, the expected complexity becomes a low-order polynomial. Clemens and Jacobs [9] analyse the speedup that can be achieved by indexing techniques alone and by indexing combined with grouping. They argue that indexing is useful only when combined with grouping techniques, owing to the combinatorial explosion in the total number of possible image groups as the size of the groups is increased. Davies [37] analyses the appropriate cardinality for the sets of image features that are mapped into the parameter space in the Hough transform. This analysis indicates that when grouping can be used effectively, the sets should be just large enough to constrain the possible positions of the model being sought to a finite set.

We analyse a simple grouping mechanism that determines pairs of feature points that are likely to belong to the same object. The criterion that is used to determine whether two feature points should be grouped, is whether they are connected in the image edge map<sup>2</sup>. Since the objects examined in our experiments are largely polyhedral, we use the additional heuristic that the edges connecting the feature points should be straight. A second heuristic, requiring that the points be some minimum distance apart in the image to form a group, is used for two reasons. First, this excluded many unmodeled groups that were formed as a result of areas of high texture. Second, groups that are close together produce unstable results in the pose estimation process, and are thus less likely to be useful.

Of course, this grouping mechanism does not yield perfect results. Pairs of points that are not modeled as relevant object groups will be found and pairs of points that belong to completely different objects may be found owing to occluding contours and other image phenomena. This is acceptable, since we use a recognition method that does not require perfect grouping. These imperfect grouping techniques still lead to considerable improvement in the performance of recognition.

The first step in our grouping mechanism is to determine the feature points in the image. This is accomplished through the use of a fast interest operator to detect corner points present in the image [38]. Next, the image edges are

<sup>2</sup> We use the term connected loosely. A method is used that allows small gaps between coincident edges to be bridged.

detected using a variant of the Canny edge detector [39]. The straight lines present in the edge map are determined using efficient Hough transform techniques [40]. Finally, pairs of corners that lie close to a straight line in the image are determined and stored.

Fig. 2 shows an example of this grouping process applied to an image of a stapler. In this case, 38 groups were found, incorporating 34 of the 63 image feature points.

**4. Computational complexity**

We shall now consider the computational complexity of object recognition using the generalized Hough transform variant described earlier, combined with the simple perceptual grouping mechanism from the previous section. Rather than examining arbitrary pairs of image points as the distinguished points and arbitrary pairs of model points to match them, we consider as the distinguished points only those pairs of points that have been grouped together. We consider only those pairs of model points that are similarly joined by an object edge as their possible matches.

The number of groups that are found by perceptual grouping systems is typically linear in the number of features. For our grouping techniques based on connectivity in the image edge map, we are guaranteed to find no more than a linear number of such groups if the edges form polygonal shapes in the image. Let us thus say that we have  $\alpha_i n$  image groups and  $\alpha_m m$  model groups. (For trihedral objects, we have  $\alpha_i \leq 3/2$  and  $\alpha_m \leq 3/2$ .)

Now, we can determine the number of iterations of the algorithm that are necessary to achieve a low rate of failure. If  $f$  is the fraction of the model points that appear in the image, then the expected fraction of the total number of

model groups that appear in the image is no less than  $f^2$ . The actual fraction should, in fact, be larger than  $f^2$ , since the pairs of points that are grouped are more likely to be either both occluded or unoccluded than random pairs of points. So, we expect at least  $f^2 \alpha_m m$  model groups to appear among the  $\alpha_i n$  image groups. The probability that a single iteration will fail to examine a group from the model is thus no more than  $1 - \frac{f^2 \alpha_m m}{\alpha_i n}$ . The probability that  $k$  iterations will all be unsuccessful is:

$$p \leq \left( 1 - \frac{f^2 \alpha_m m}{\alpha_i n} \right)^k \tag{3}$$

We can set this probability to be less than some arbitrarily small constant  $\delta$  and solve for the number of iterations necessary to achieve this accuracy:

$$\left( 1 - \frac{f^2 \alpha_m m}{\alpha_i n} \right)^k \leq \delta, \tag{4}$$

$$k \geq \frac{\ln \delta}{\ln \left( 1 - \frac{f^2 \alpha_m m}{\alpha_i n} \right)} < \frac{\alpha_i n}{f^2 \alpha_m m} \ln \frac{1}{\delta} \tag{5}$$

For each iteration, we consider each of the model groups as a possible match for the image group and for each such match, we determine and cluster the poses corresponding to all matches between three image points and three model points that match the image group to the model group. The total number of poses computed is approximately  $\frac{\alpha_i n(n-2)(m-2)}{f^2} \ln \frac{1}{\delta}$  a speedup of approximately  $\frac{n}{\alpha_i}$  over the method that did not use perceptual grouping. The complexity of the algorithm is  $O(mn^2)$ .

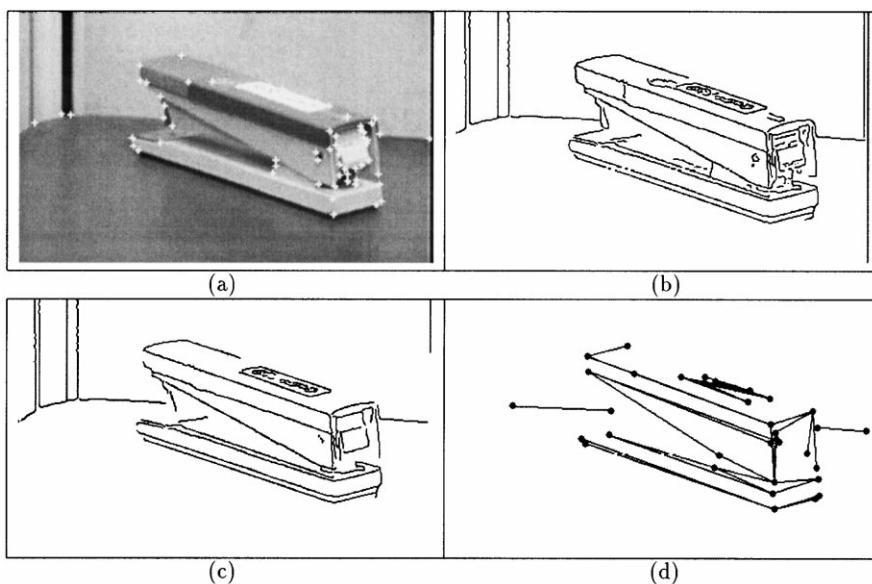


Fig. 2. Example of the grouping process. (a) Corners found in the image; (b) edges detected in the image; (c) lines detected in the edge image using Hough transform techniques; and (d) groups detected in the image.

## 5. Rate of false positives

The previous section demonstrated that the computational complexity of object recognition using the generalized Hough transform can be reduced through the use of imperfect perceptual grouping techniques. We shall now see that the use of such techniques can also reduce the rate at which false positive instances of objects are found.

We compute an estimate of this false alarm rate, building on previous work [17–20] that has used the Bose–Einstein occupancy model (see e.g. [41]) to estimate the probability that a position in the pose space brings a large number of feature matches into alignment for various recognition problems. The previous analysis can be modified for the case at hand by considering the probability that a position in the pose space is both consistent with a match between pairs of grouped features and consistent with sufficient additional matches to result in false positive.

In order to simplify the analysis, we assume that the locations of the individual model and image features can be modeled as being independent. In this case, the probability,  $P_{fp}(t)$ , that a false positive of size  $t$  occurs for some position in the pose space can be estimated as the product of the probability,  $P_g$ , that the pose brings one of the model groups into alignment with one of the image groups, and the probability,  $P_r(t-2)$ , that the pose brings  $t-2$  additional model feature points into alignment with image feature points.

$$P_{fp}(t) = P_g \cdot P_r(t-2). \quad (6)$$

Let's first consider  $P_g$ . We approximate the distribution of the transformed model features in the image (over the space of possible model positions) as being uniform. That is, we model the probability that a model point is projected to any point in the image as a constant regardless of which point it is in the image. Points outside of the image are neglected. This implies that the probability,  $P_1$ , that a pose brings a pair of model features into alignment with a random pair of image features to within an error of  $\epsilon$  in an image that is  $W$  pixels wide and  $H$  pixels high is:

$$P_1 = \left( \frac{\pi \epsilon^2}{WH} \right)^2. \quad (7)$$

If we have  $\alpha_n n$  image groups and  $\alpha_m m$  model groups, then

$P_g$  can be approximated as follows:

$$P_g \approx (1 - (1 - P_1)^{\alpha_m \alpha_n}) < \frac{\alpha_m \alpha_n \pi^2 \epsilon^4}{W^2 H^2}. \quad (8)$$

Now, we must consider,  $P_r(t-2)$ , the probability that  $t-2$  additional matches are brought into alignment by the pose. Previous work [17] has indicated that this can be approximated by

$$P_r(x) \approx \left( \frac{bmn}{1 + bmn} \right)^x, \quad (9)$$

where  $b \approx \pi \epsilon^2 / WH$  is the average fraction of the pose space that aligns a single model point with a single image point and  $x = t-2$  is the size of the false positive. We thus have:

$$P_r(t-2) \approx \left( \frac{mn}{\frac{WH}{\pi \epsilon^2} + mn} \right)^{t-2}. \quad (10)$$

Substituting in Eq. (6) yields:

$$P_{fp}(t) \approx \left( \frac{mn}{\frac{WH}{\pi \epsilon^2} + mn} \right)^{t-2} \frac{\alpha_m \alpha_n \pi^2 \epsilon^4}{W^2 H^2}. \quad (11)$$

A comparison of the estimated probability of a false positive with the parameters  $m = 30$ ,  $n = 150$  and  $WH = 65536$  for the case with no grouping (Eq. (1)) and for the case with grouping (Eq. (11)) can be found in Table 1. The probability of a false positive is much lower when grouping is used. False positives occur approximately 230 times more frequently when grouping is not used for  $\epsilon = 3$  and approximately 49 times more frequently for  $\epsilon = 5$ .

## 6. Experiments

A recognition system using these techniques has been implemented on a SPARCstation<sup>®</sup>20. The basic algorithm is similar to the one given in Fig. 1. The only differences are that instead of choosing two random image points on line 3 of the algorithm, we choose a random pair of image points that have been grouped, and instead of considering all possible matching pairs of model points on line 4 of the algorithm, we examine only the pairs of object model points that have been grouped a priori. This grouping of the model points was done offline, by hand, for these experiments, by pairing the feature points that were connected by an edge in

Table 1

Estimated probability of a false positive for cases with no grouping and with grouping. Parameters used:  $m = 30$ ,  $n = 150$ ,  $WH = 65536$ ,  $\alpha_i = \alpha_m = 1.5$

$f$	No grouping		Grouping	
	$\epsilon = 3$	$\epsilon = 5$	$\epsilon = 3$	$\epsilon = 5$
1.0	$3.86 \times 10^{-6}$	$6.08 \times 10^{-3}$	$1.67 \times 10^{-8}$	$1.24 \times 10^{-4}$
0.75	$8.74 \times 10^{-5}$	$2.18 \times 10^{-2}$	$3.77 \times 10^{-7}$	$4.45 \times 10^{-4}$
0.50	$1.97 \times 10^{-3}$	$7.80 \times 10^{-2}$	$8.50 \times 10^{-6}$	$1.59 \times 10^{-3}$

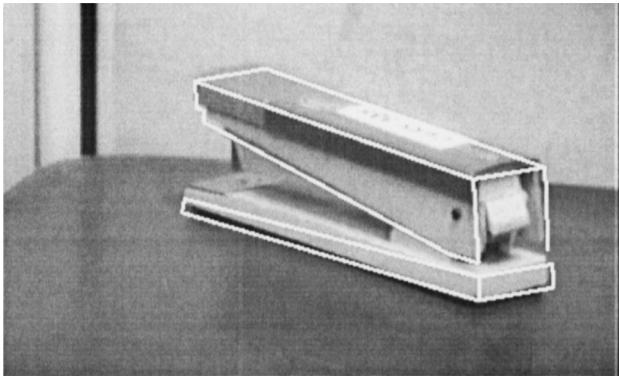


Fig. 3. Recognized position of the stapler.

the object model. The method that is used to determine the poses in line 6 of the algorithm is the technique of Huttenlocher and Ullman [31], which uses the weak-perspective approximation to the perspective projection. The clusters among the poses are detected using recursive histogramming techniques [17].

This system was tested on several of the same images as the original pose clustering system to verify the improved performance. For example, Fig. 3 shows the recognized position of the stapler from Fig. 2. Fig. 4 gives an additional example of the recognition process. Each object that appears in the images is correctly recognized.

The system required between 61 and 100 s per object in these examples, depending on the complexity of the object model and the image, which is an tremendous improvement over the previous implementation. However, this improvement is not solely a result of the grouping. Not only was a faster CPU used, but the implementation of the clustering method was considerably improved. A true measure on the

performance improvement yielded by grouping is the number of poses that are computed and clustered in each case. Table 2 gives these numbers for the objects recognized in Fig. 4. When the grouping techniques are used to improve the performance, the number of poses that are computed and clustered is reduced to less than 1% of the number of poses that are computed and clustered when grouping is not used in this image.

No significant false positives were found when the grouping techniques were used. Some instances occurred where a pair of distinguished points from the model produced hypotheses that included several correct and incorrect feature matches, but the same groups produced superior matches involving predominantly correct matches in each case.

## 7. Summary

This paper has examined the benefits that perceptual grouping techniques can yield in the object recognition process. While it is known that perfect grouping techniques can reduce the computational complexity of recognition tasks, the benefits of imperfect grouping techniques have not been quantified in detail previously. We have taken a step in this direction by quantifying the benefits that can be obtained in performing object recognition using the generalized Hough transform through the use of imperfect grouping techniques.

It has been demonstrated that even very simple, imperfect grouping techniques can yield a considerable improvement in the both the speed and accuracy of object recognition. We have concentrated on grouping techniques that determine

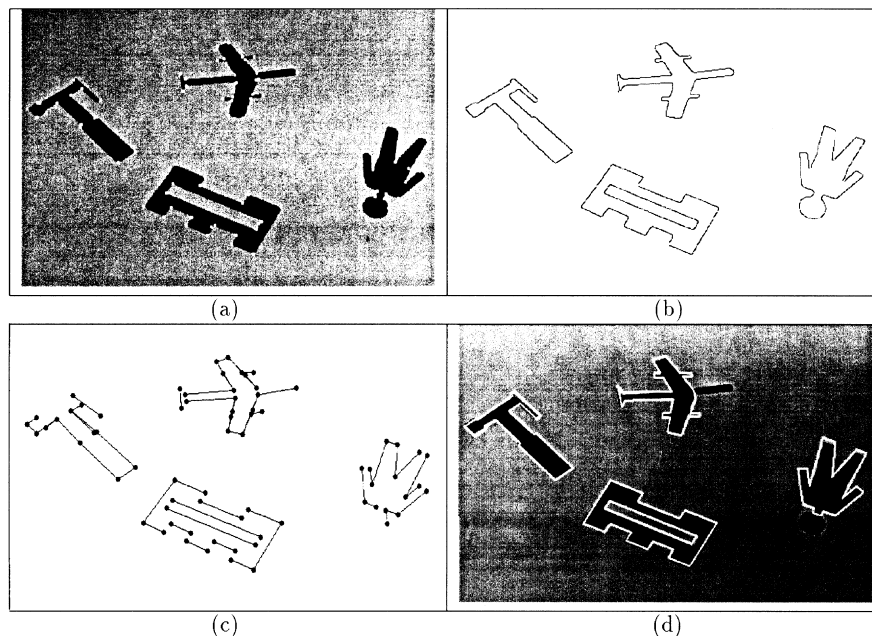


Fig. 4. Recognition of two-dimensional figures. (a) Corners found in the image; (b) edges detected in the image; (c) groups formed in the image; and (d) recognized positions of the figures.

Table 2  
Number of poses computed with no grouping ( $N_{ng}$ ) and with grouping ( $N_g$ )

Object	$N_{ng}$	$N_g$	$\frac{N_{ng}}{N_g}$
Widget	$2.56 \times 10^8$	$2.52 \times 10^6$	101.5
Plane	$3.70 \times 10^8$	$3.54 \times 10^6$	104.7
Hammer	$2.56 \times 10^8$	$2.52 \times 10^6$	101.5
Person	$2.11 \times 10^8$	$2.09 \times 10^6$	100.9

pairs of points that are likely to belong to the same model by examining which feature points are connected by edges in the image edge map.

Our analysis shows that the computational complexity of recognition using a variant of the generalized Hough transform is reduced from  $O(mn^3)$  to  $O(mn^2)$  through the use of these grouping techniques. Further, the use of these techniques considerably reduces the probability of detecting a false positive object instance. Finally, we have implemented a recognition system that uses these techniques and performed experiments with this system that have supported the analysis and the conclusions that we reached.

## References

- [1] A. Witkin, J. Tenenbaum, On the role of structure in vision, in: J. Beck, B. Hope, A. Rosenfeld, (Eds.), *Human and Machine Vision*, Academic Press, New York, 1983, pp. 481–543.
- [2] D.G. Lowe, *Perceptual Organization and Visual Recognition*. Kluwer, Dordrecht, 1985.
- [3] A. Sha'ashua, S. Ullman, Structural saliency: the detection of globally salient structures using a locally connected network, in: *Proceedings of the International Conference on Computer Vision*, 1988, pp. 321–327.
- [4] N. Ahuja, M. Tuceryan, Extraction of early perceptual structure in dot patterns: integrating region, boundary, and component gestalt, *Computer Vision, Graphics, and Image Processing* 48 (1989) 304–356.
- [5] D.P. Huttenlocher, P.C. Wayner, Finding convex edge groupings in an image, *International Journal of Computer Vision* 8 (1) (1992) 7–27.
- [6] R. Mohan, R. Nevatia, Perceptual organization for scene segmentation and description, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14 (6) (1992) 616–635.
- [7] D.W. Jacobs, Robust and efficient detection of salient convex groups, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (1) (1996) 23–37.
- [8] Y. Lamdan, J.T. Schwartz, H.J. Wolfson, Affine invariant model-based object recognition, *IEEE Transactions on Robotics and Automation* 6 (5) (1990) 578–589.
- [9] D.T. Clemens, D.W. Jacobs, Space and time bounds on indexing 3-d models from 2-d images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13 (10) (1991) 1007–1017.
- [10] D.W. Jacobs, Space efficient 3d model indexing, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1992, pp. 439–444.
- [11] D. Weinshall, Model-based invariants for 3-d vision, *International Journal of Computer Vision* 10 (1) (1993) 27–42.
- [12] C.F. Olson, Probabilistic indexing for object recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (5) (1995) 518–522.
- [13] W.E.L. Grimson, D.P. Huttenlocher, On the sensitivity of geometric hashing, in: *Proceedings of the International Conference on Computer Vision*, 1990, pp. 334–338.
- [14] W.E.L. Grimson, D.P. Huttenlocher, D.W. Jacobs, A study of affine matching with bounded sensor error, *International Journal of Computer Vision* 13 (1) (1994) 7–32.
- [15] W.E.L. Grimson, The combinatorics of object recognition in cluttered environments using constrained search, *Artificial Intelligence* 44 (1-2) (1990) 121–165.
- [16] D.H. Ballard, Generalizing the Hough transform to detect arbitrary shapes, *Pattern Recognition* 13 (2) (1981) 111–122.
- [17] C.F. Olson, Efficient pose clustering using a randomized algorithm, *International Journal of Computer Vision* 23 (2) (1997) 131–147.
- [18] W.E.L. Grimson, D.P. Huttenlocher, On the sensitivity of the Hough transform for object recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12 (3) (1990) 255–274.
- [19] W.E.L. Grimson, D.P. Huttenlocher, On the verification of hypothesized matches in model-based recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13 (12) (1991) 1201–1213.
- [20] W.E.L. Grimson, D.P. Huttenlocher, T.D. Alter, Recognizing 3d objects from 2d images: an error analysis, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1992, pp. 316–321.
- [21] C.F. Olson, On the speed and accuracy of object recognition when using imperfect grouping, in: *Proceedings of the International Symposium on Computer Vision*, 1995, pp. 449–454.
- [22] G. Stockman, S. Kopstein, S. Benett, Matching images to models for registration and object detection via clustering, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 4 (3) (1982) 229–241.
- [23] T.M. Silberberg, L.S. Davis, D.A. Harwood, An iterative Hough procedure for three-dimensional object recognition, *Pattern Recognition* 17 (6) (1984) 620–629.
- [24] J.L. Turney, T.N. Mudge, R.A. Volz, Recognizing partially occluded parts, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 7 (4) (1985) 410–421.
- [25] T.M. Silberberg, D.A. Harwood, L.S. Davis, Object recognition using oriented model points, *Computer Vision, Graphics, and Image Processing* 35 (1986) 47–71.
- [26] M. Dhome, T. Kasvand, Polyhedra recognition by hypothesis accumulation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9 (3) (1987) 429–438.
- [27] G. Stockman, Object recognition and localization via pose clustering, *Computer Vision, Graphics, and Image Processing* 40 (1987) 361–387.
- [28] D.W. Thompson, J.L. Mundy, Three-dimensional model matching from an unconstrained viewpoint, in: *Proceedings of the IEEE Conference on Robotics and Automation*, vol. 1, 1987, pp. 208–220.
- [29] S. Linnainmaa, D. Harwood, L.S. Davis, Pose determination of a three-dimensional object using triangle pairs, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 10 (5) (1988) 634–647.
- [30] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Communications of the ACM* 24 (1981) 381–396.
- [31] D.P. Huttenlocher, S. Ullman, Recognizing solid objects by alignment with an image, *International Journal of Computer Vision* 5 (2) (1990) 195–212.
- [32] T. A. Cass, Polynomial-time object recognition in the presence of clutter, occlusion, and uncertainty, in: *Proceedings of the European Conference on Computer Vision*, 1992, pp. 834–842.
- [33] T. M. Breuel, Fast recognition using adaptive subdivisions of transformation space, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1992, pp. 445–451.
- [34] S. Sarkar, K.L. Boyer, Perceptual organization in computer vision: a review and a proposal for a classificatory structure, *IEEE Transactions on Systems, Man, and Cybernetics* 23 (2) (1993) 382–399.
- [35] D. W. Jacobs, GROPER: a grouping based recognition system for two dimensional objects, in: *IEEE Workshop on Computer Vision*, 1987, pp. 164–169.

- [36] P. Havaladar, G. Medioni, F. Stein, Perceptual grouping for generic recognition, *International Journal of Computer Vision* 20 (1-2) (1996) 59–80.
- [37] E. R. Davies, A framework for designing optimal Hough transform implementations, in: *Proceedings of the IAPR International Conference on Pattern Recognition*, vol. 3, 1992, pp. 509–512.
- [38] W. Förstner, Image matching, in: *Computer and Robot Vision*, vol. 11, chap. 16, Addison-Wesley, Reading, MA, 1993.
- [39] J. Canny, A computational approach to edge detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8 (6) (1986) 679–697.
- [40] C. F. Olson, Decomposition of the Hough transform: Curve detection with efficient error propagation, in: *Proceedings of the European Conference on Computer Vision*, vol. 1, 1996, pp. 263–272.
- [41] W. Feller, *An Introduction to Probability Theory and Its Applications*. Wiley, New York, 1968.