Pattern Recognition, Vol. 24, No. 4, pp. 303–316, 1991
Printed in Great Britain

0031–3203/91 $3.00 + .00
Pergamon Press plc

A PROBABILISTIC HOUGH TRANSFORM

N. KIRYATI,* Y. ELDAR* and A. M. BRUCKSTEIN†

*Department of Electrical Engineering and †Department of Computer Science, Technion, Israel Institute of Technology, Haifa 32000, Israel

(*Received* 10 *January* 1990; *in revised form* 27 *June* 1990; *received for publication* 2 *August* 1990)

Abstract—The Hough Transform for straight line detection is considered. It is shown that if just a small subset of the edge points in the image, selected at random, is used as input for the Hough Transform, the performance is often only slightly impaired, thus the execution time can be considerably shortened. The performance of the resulting "Probabilistic Hough Transform" is analysed. The analysis is supported by experimental evidence.

Computer vision Feature extraction Hough Transform Monte-Carlo methods
Probabilistic algorithms

1. INTRODUCTION

The well-known Hough Transform[1] is an efficient method for the detection of predefined features in digital images. Consider the Hough Transform for straight line detection using normal parameters as suggested by Duda and Hart;[2] in order to detect lines through large collinear subsets of a planar set of edge points $P \triangleq \{(x_i, y_i)\ i = 1, \ldots, M\}$, each point is regarded as a constraint

$$\rho = x_i \cos\theta + y_i \sin\theta \tag{1}$$

on the normal parameters (ρ, θ) of the straight lines on which the point may be located. The intersection of a large number of sinusoids in the (ρ, θ) normal parameter plane corresponds to the normal parameters of a straight line through a large collinear subset of P.

The Duda and Hart[2] algorithm is a member of a family of "one to many" Hough Transforms in which one edge point is transformed to many points in the parameter space. "Many to one" and "one to one" Hough Transforms have also been suggested and analysed, especially with respect to the effects of edge point location errors on peak location in the parameter space.[3–5]

In the standard implementation of the Duda and Hart algorithm (a subset of), the (ρ, θ) plane is divided into $N_\rho \times N_\theta$ rectangular cells and represented by an accumulator array. The algorithm is performed in two stages; the first is an incrementation stage in which for each $i \in [1, \ldots, M]$ the accumulators corresponding to cells that the sinusoid (1) intersects are incremented. The second stage is an exhaustive search for maxima in the accumulator array. These represent the normal parameters of straight lines through large collinear subsets of points.

The execution of the Duda and Hart algorithm requires $O(M \cdot N_\theta)$ operations in the incrementation stage and $O(N_\rho \cdot N_\theta)$ operations in the search stage. In typical applications each dimension of the accumulator array, N_ρ and N_θ, is a few hundred cells or less, while M, the number of edge points, could be an order of magnitude higher. Thus the incrementation stage usually dominates the execution time of the algorithm. The linear dependence of the number of operations on M is an important advantage of accumulator-based "one to many" Hough Transforms in comparison with "many to one" algorithms and with algorithms that employ clustering techniques other than the accumulator method.

The large number of operations has been a major obstacle to wide-spread application of the Hough Transform. One approach to accelerate the algorithm is based on "coarse to fine" variable resolution quantization of the parameter plane. Certain theoretical limitations have, however, been observed.[6] Another approach is parallel implementation, often specifying specialized hardware which is expensive or bulky.

The incrementation stage of the Hough algorithm is often regarded and referred to as "voting": each edge point "votes" for the parameter pairs of all straight lines on which it can lie. The "winning" lines correspond to the largest collinear subsets of edge points in the image. This paper suggests that in many practical situations a limited poll can replace full-scale voting with negligible performance degradation. This means that a small subset of the edge points, selected at random, can often provide sufficient input to the Hough algorithm, allowing considerable reduction in execution time. A closely related approach has been briefly sketched in reference (7) to demonstrate a possible application of a "blackboard"-based parallel processing architecture.

The fundamental difference between the Duda and Hart algorithm and the suggested "Probabilistic Hough Transform" is that only m edge points ($m < M$) selected at random guide the incrementation process. As the number of operations at this stage is proportional to $m \cdot N_\theta$, significant computational savings with respect to the Duda and Hart algorithm will result if m could be made much smaller than M. It should be observed that if $m < N_\rho$ further computational savings can be obtained by using Gerig's[8] sinusoid-guided peak detection procedure instead of conventional exhaustive search. While exhaustive search requires about $N_\rho \cdot N_\theta$ operations, Gerig's method, which is computationally equivalent to the incrementation stage, requires approximately $m \cdot N_\theta$ operations.

The limit on the possible decrease of m is the requirement that a feature in the image that manifests itself as a significant peak in the accumulator array of the conventional algorithm should be, with high probability, detectable using just m edge points to guide accumulation in the Probabilistic Hough algorithm. The optimal choice of m, and the resulting computational savings, are problem dependent. In the next section a model is presented and analysis is carried out. The applicability of the suggested approach to real problems that differ from the model has been experimentally studied, and is reported in Section 3. Preliminary results have been presented in reference (9).

2. PROBABILISTIC HOUGH TRANSFORM ANALYSIS

Assume that the image occupies a circle of unit radius, centred at the origin. (Using square rather than circular image model would have led to cumbersome, albeit straightforward, analysis, which might have obscured more important principles.) The set of edge points P contains M points, S which lie on a "thick" straight line segment of length L and width $2b$, and $N \triangleq M - S$ which are due to noise, and are uniformly distributed in the edge map. Normal parametrization of straight lines and rectangular plane quantization are employed, such that $\theta \in [0, \pi)$ is sampled at N_θ evenly spaced values $\Delta\theta = \pi/N_\theta$ apart, and $\rho \in [-1, 1]$ is quantized into N_ρ buckets of width $\Delta\rho = 2/N_\rho$ each. Suppose that m edge points are selected at random to guide the incrementation process; let s denote the number of selected points which belong to the line segment, and let $n \triangleq m - s$ denote the number of selected points which are due to noise.

It is well known that the quantization of the parameter plane and the "thickness" of the line segment lead to spreading of the peak in the accumulator array, making it harder to detect. The worst case spreading at a given θ in terms of buckets along the ρ coordinate is:[10,11]

$$d = \left\lfloor \frac{L\sin(\Delta\theta/2) + 2b\cos(\Delta\theta/2)}{\Delta\rho} \right\rfloor + 2. \tag{2}$$

A standard peak-detection practice which will be followed here is to find the maximum sum of buckets in a sliding rectangular window which is designed to capture the full peak spread and consists of d consecutive buckets in the ρ direction.

The sliding window generally takes in contributions from both the signal and noise. At the peak, the signal contribution is the random variable s which is governed by the binomial distribution. The probability of s selected points which belong to the line is:

$$P(s) = \binom{m}{s} (S/M)^s (N/M)^{m-s}. \tag{3}$$

If m is large enough such that

$$m \cdot \frac{S}{M} \cdot \frac{N}{M} \gg 1 \tag{4}$$

then the Gaussian approximation of (3) holds, with an expected value

$$\eta_s = \frac{m}{M} S \tag{5}$$

and standard deviation

$$\sigma_s = \sqrt{m \cdot \frac{S}{M} \cdot \frac{N}{M}}. \tag{6}$$

Let a random variable n^* represent the noise contribution to the sliding window at a certain location in the accumulator array. Generally,

$$P(n^*) = \sum_{n=0}^{m} P(n) \cdot P(n^*/n) \tag{7}$$

where $P(n)$ is the probability of n selected noise edge points and $P(n^*/n)$ represents the conditional probability of n^*. $P(n)$ is binomial

$$P(n) = \binom{m}{n} (N/M)^n (S/M)^{m-n} \tag{8}$$

and assuming that (4) holds can be well approximated by a Gaussian with an expected value

$$\eta_n = \frac{m}{M} N \tag{9}$$

and standard deviation $\sigma_n = \sigma_s$. Note that since in interesting (i.e. difficult) situations N is much larger than S and $m \gg 1$, then usually $\sigma_n \ll \eta_n$.

It is known[12] that uniform noise in the image leads to non-uniform noise in the accumulator array. In the model studied here, it is easy to show that for a sliding window centred at a certain location (ρ_0, θ_0) in the accumulator array,

$$P(n^*/n) = \binom{n}{n^*} p^{n^*} (1-p)^{n-n^*} \tag{10}$$

where p is the probability that a "noise" edge point, selected at random, contributes to the sliding window at (ρ_0, θ_0). For $d\Delta\rho \ll 1$,

$$p \approx \frac{2d\Delta\rho}{\pi}\sqrt{1-\rho_0^2} \tag{11}$$

i.e. p is the fraction of the image area which is projected onto a segment of length $d\Delta\rho$, concurrent with the diameter of the image at any orientation θ_0, and at average distance ρ_0 from the centre along that diameter, as shown in Fig. 1. In the most noisy neighbourhoods in the accumulator array

$$p \approx 2d\Delta\rho/\pi. \tag{12}$$

If np is not much larger than 1 the Poisson approximation to the binomial distribution holds:

$$P(n^*/n) \approx e^{-np}\frac{(np)^{n^*}}{n^*!}. \tag{13}$$

Furthermore, since $\sigma_n \ll \eta_n$ and since $P(n^*/n)$ is not extremely sensitive to small variations in n, it is safe to approximate $P(n)$ in (7) as an impulse function at $n = \eta_n$. Thus

$$P(n^*) \approx e^{-\eta_n \cdot p}\frac{(\eta_n p)^{n^*}}{n^*!} \tag{14}$$

where η_n is given by (9). The approximation of $P(n)$ by a constant also implies that the random variables s and n^*can be considered "almost" independent.

Peak detection difficulties in the Hough Transform, i.e. high counts at accumulators distant from the genuine peak location, are generally due to random noise and due to bias,[13] which is an interference generated by features in the image. Since our image model is noise-dominated, i.e. the image contains just noise and a single straight line segment, bias in the accumulator array is limited to that which is caused by projections of the line segment itself at various angles. This bias is characterized by its so-called butterfly pattern, and can actually be utilized to improve the location accuracy of the detected line.[11,14] Hence, in the present noise-dominated model, the main issue is whether or not the true peak would be distinct with respect to peaks generated by random noise.

It is desired to detect the peak by establishing a threshold T, such that a sliding window at a certain location would be said to be in the neighbourhood of the peak if its value exceeds the threshold, and would otherwise be associated with noise. The detection probability is

$$P_D = \text{Prob}(s + n^*) > T \tag{15}$$

and the false alarm probability per window location is

$$P_{FA} = \text{Prob}(n^*) > T. \tag{16}$$

Since $P(n^*)$ depends on ρ_0, P_D and P_{FA} are generally ρ_0-dependent as well. Thus the optimal threshold T is a function of ρ_0.

The above analysis is now used to relate P_D, P_{FA}, T and m in a particular example. Let $M = 5000$ be the total number of edge points, $N = 4000$ be the number of points due to random noise and $S = 1000$ be the number of points generated by a line segment of length $L = 0.8$ and width $2b = 0.02$, which is typical to edges obtained from real images by edge detection. Since noise level is highest at $\rho_0 \approx 0$, for worst case analysis a line that is nearly concurrent with a diameter of the circular image is assumed. The dimensions of the accumulator array are $N_\rho = N_\theta = 500$.

Using (2) we obtain that the worst case spreading of the peak in buckets along the ρ coordinate is $d = 7$. Applying (11), we find

$$p \approx 0.018\sqrt{1-\rho_0^2}. \tag{17}$$

Under the assumption that mp is not much larger than 1, the noise in the accumulator array at $\rho_0 \approx 0$ is governed by the Poisson distribution, and the probability of false alarm per window location is

$$P_{FA} \approx \sum_{n^*=T}^{\infty} e^{-0.015m}\frac{(0.015m)^{n^*}}{n^*!}. \tag{18}$$

The probability distribution of $s + n^*$ can be approximated, since s and n^* are "almost" independent, by convolution of the individual distributions. Assuming that $0.16m$ is large, $P(s)$ is approximately Gaussian with

$$\begin{cases} \eta_s = 0.2m \\ \sigma_s = 0.4\sqrt{m}. \end{cases} \tag{19}$$

Since $\sigma_{n^*} \approx 0.3\sigma_s$ and $\eta_{n^*} = 0.075\eta_s$, $P(s + n^*)$ can be roughly approximated by a Gaussian with

$$\begin{cases} \eta_{s+n^*} = 0.215m \\ \sigma_{s+n^*} = 0.418\sqrt{m}. \end{cases} \tag{20}$$

To obtain $P_D \approx 98\%$ the threshold is set at

$$T \approx \eta_{s+n^*} - 2\sigma_{s+n^*} \tag{21}$$

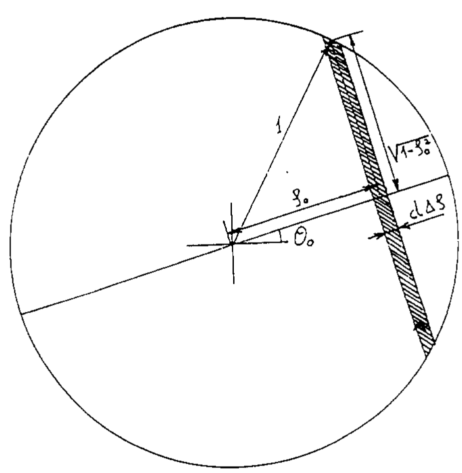


Fig. 1. The geometric interpretation of (11).

or

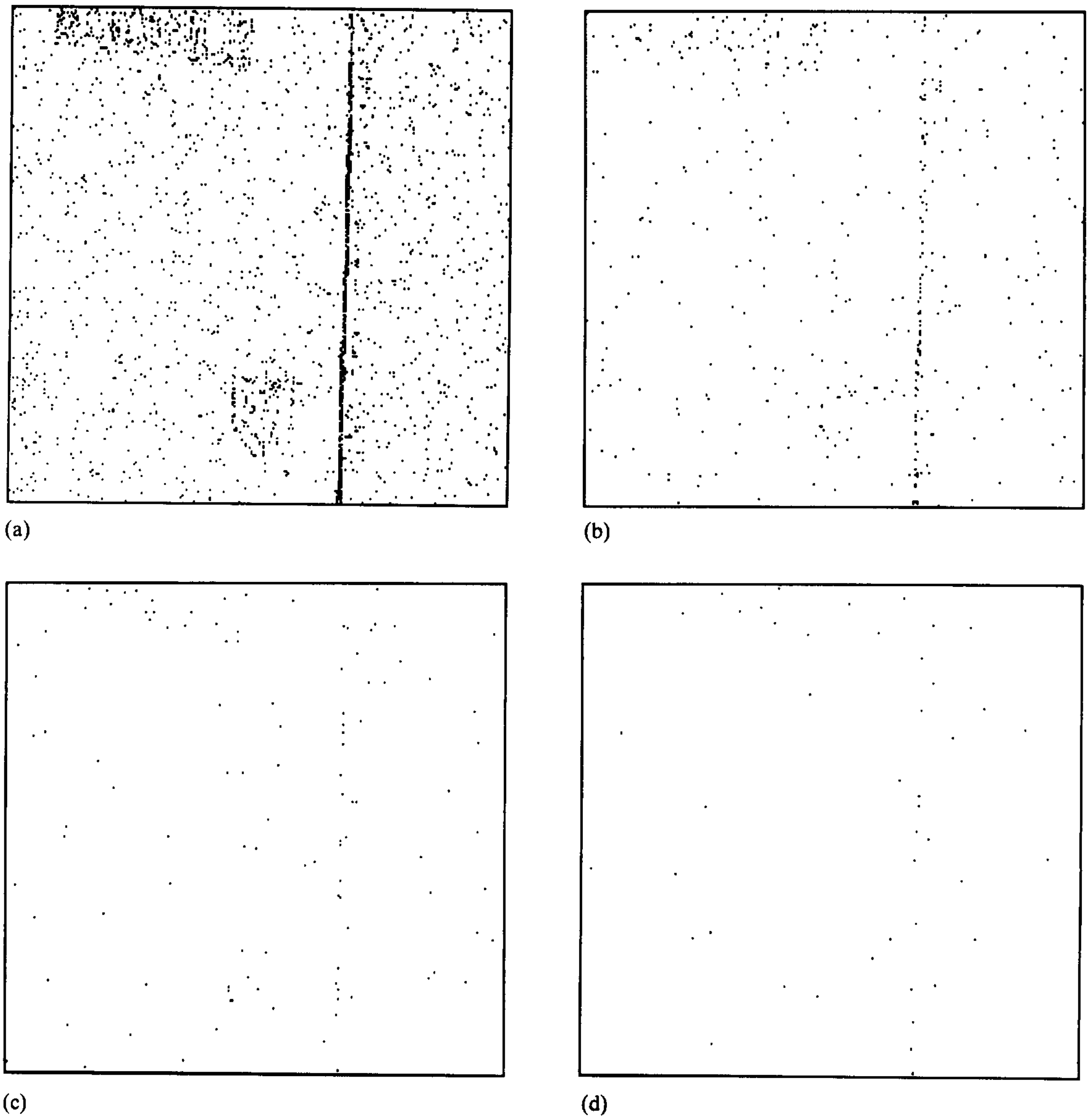
$$T \approx 0.215m - 0.84\sqrt{m}. \tag{22}$$

To illustrate, assume that $m = 100$, i.e. a random sample of just 2% of the edge points is used. To obtain 98% detection probability, the threshold is set at $T = 13$. In this case the per-window false alarm rate at the most noisy area in the accumulator array is in the order of 10^{-8}, which is negligible. The behaviour of P_{FA} as a function of m for $P_D = 98\%$ shows a threshold effect: for $m = 50$ the per-window false alarm rate is about 10^{-3} at noisy areas, implying the existence of dozens of false alarms. It is interesting to note that if the sliding window method had not been used and peaks had been detected by per-accumulator comparison, it would have been necessary to increase m to more than 400 to obtain performance similar to that which is obtained with the sliding window method at $m = 100$.

3. EXPERIMENTS

The theoretical analysis just presented predicts that the improvement of the performance of the Probabilistic Hough Transform as the sample size increases exhibits a threshold effect. Thus, even in the presence of significant noise as well as errors in the coordinates of the data points, large computational savings should be possible.

The analysis has been based on a certain image model, as well as on a specific formulation of the algorithm. A series of experiments was designed in order to verify the applicability of the suggested



(a) (b) (c) (d)

Fig. 2. (a) An original edge map containing 2150 points, 270 of which lie on the straight edge. (b), (c) and (d) show a random selection of 20%, 5% and 2% of the edge points, respectively.

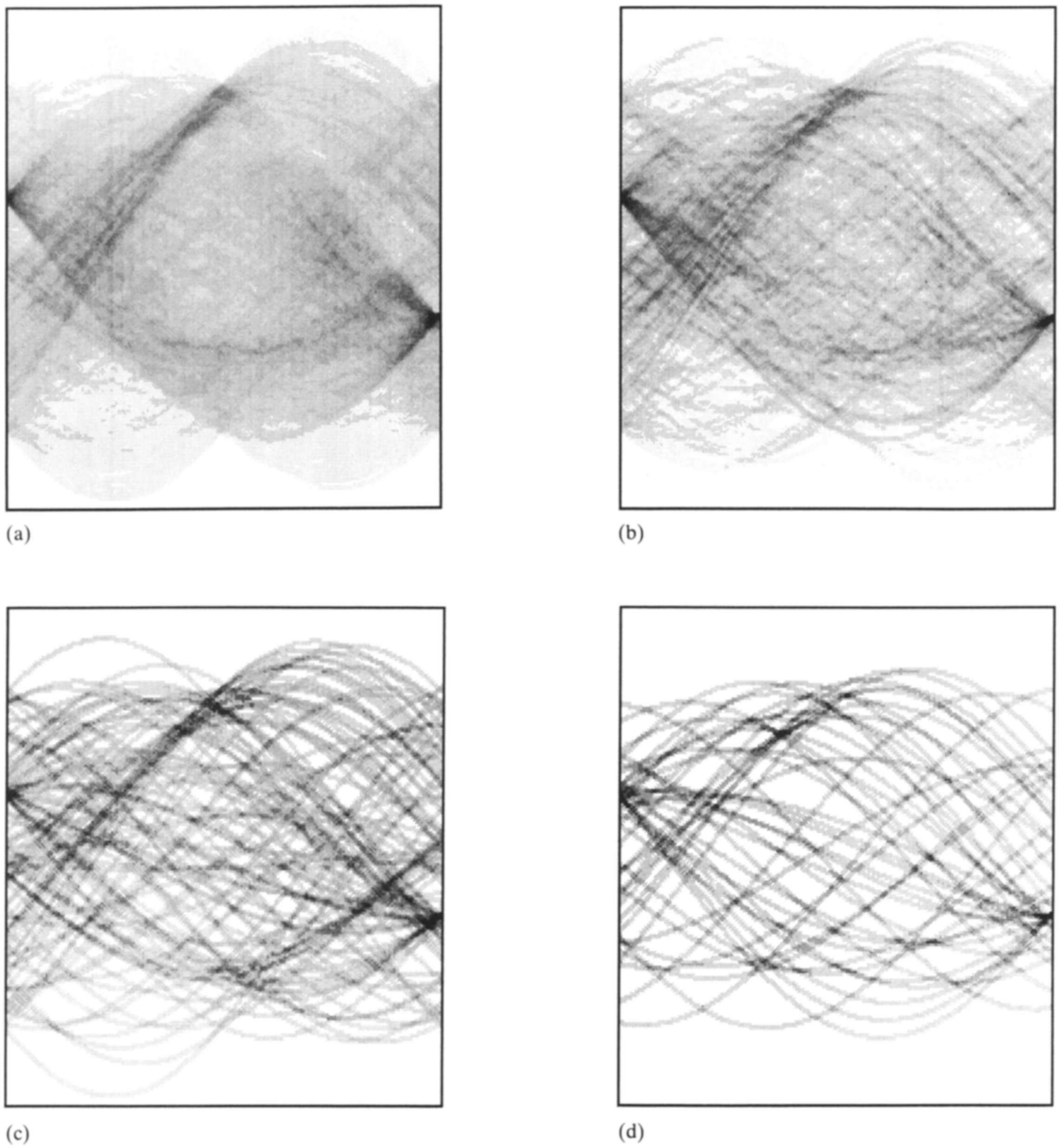


Fig. 3. The accumulator maps that correspond to the edge maps of Fig. 2.

approach in real problems, that deviate in several important ways from the theoretical assumptions.

In order to evaluate the performance of the algorithm, an error measure must be chosen. In all the following experiments the chosen measure is the average number of false alarms, i.e. peaks in the accumulator array that are higher than the peaks that correspond to the "true" features. This measure represents the average number of random arrangements that would erroneously be detected whenever the threshold is adjusted to be low enough to enable detection of the genuine features. The average number of false alarms is useful in predicting the computational effort required for implementing a post-processing verification algorithm. Each point in the following graphs (showing the average number of false alarms as a function of the fraction of data points used to guide accumulation) is an average based on the results of 100 experiments, each with its own random selection of data points.

All experiments were carried out using real 256×256 pictures, taken with a CCD camera. The corresponding edge maps were generated by simple edge detection procedures. In some cases synthetic noise or features had to be added.

The first experiment was meant to verify the analysis, using a simple image which roughly fits the theoretical model's assumptions. Figure 2(a) shows an

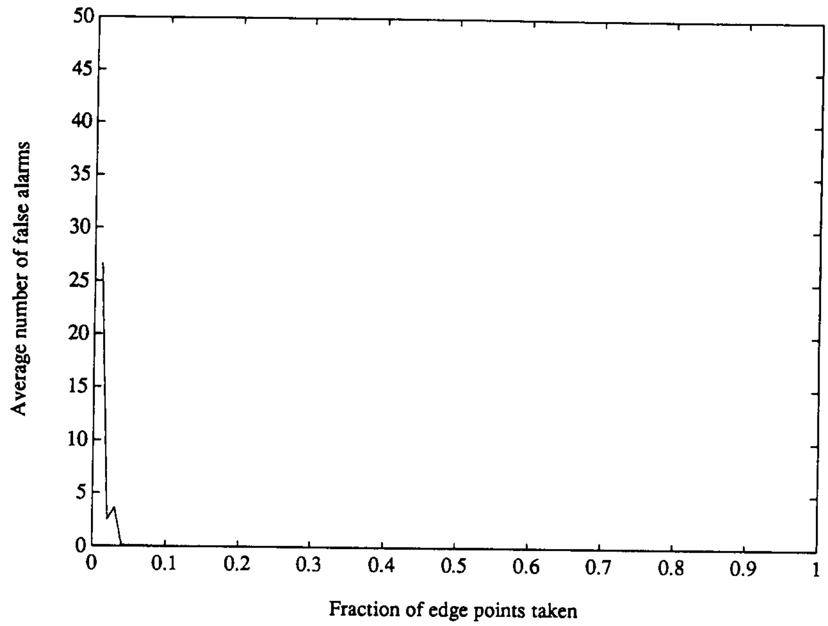


Fig. 4. The average number of false alarms (based on 100 experiments) as a function of the fraction of edge points selected to guide the accumulation. A threshold effect is evident.

edge map obtained from a real image, to which uniformly distributed noise has been synthetically added. The main feature in the image is a straight edge which consists of about 270 edge points out of a total of about 2150 in the edge map. There are also a few smaller patterns and very many noise points that constitute the main "distractor". This essentially corresponds to the noise-dominated model assumed in the theoretical analysis. When all the edge points are input to the Duda and Hart algorithm (with an 180 × 200 accumulator array) the accumulation is as shown in Fig. 3(a), and the straight edge is easily detected. Figures 2(b), (c) and (d) show, respectively, a random selection which consists of 20%, 5% and 2% of the edge points; the respective accumulator maps are shown in Figs 3(b), (c) and (d). Figure 4 shows the average number of false alarms as a function of the fraction of edge points selected to guide the accumulation. It clearly demonstrates the threshold effect predicted by theory which makes significant computational savings possible.

In the edge map of Fig. 5, which is a synthetic

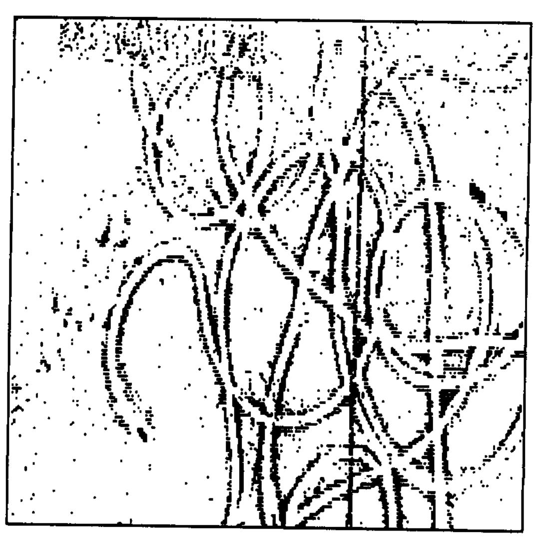


Fig. 5. An edge map (6998 points), containing a straight edge of 420 points, a distracting curve, and noise.

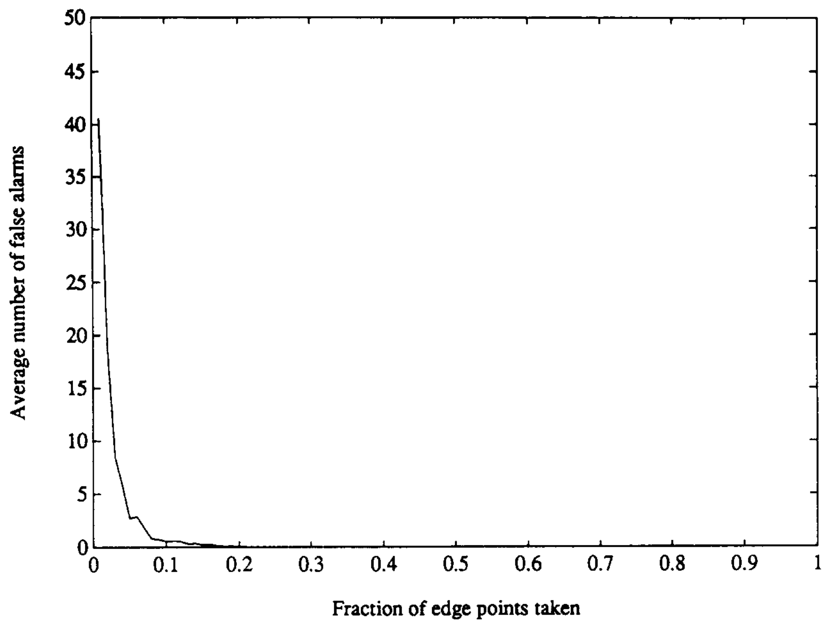
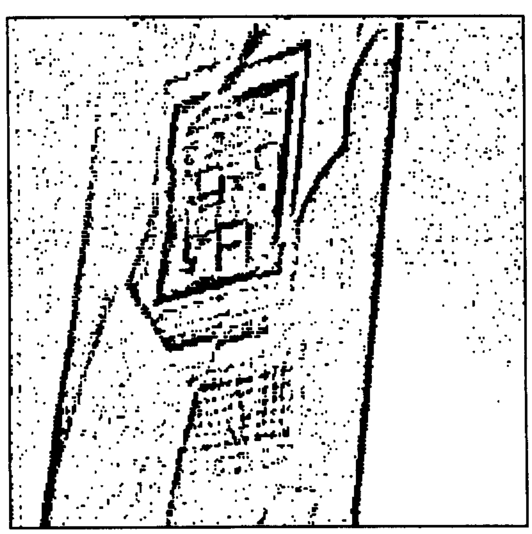


Fig. 6. The average number of false alarms as a function of the sample size, when the Probabilistic Hough Transform is applied to the edge map shown in Fig. 5. The sharp threshold indicates that large computational savings are possible.

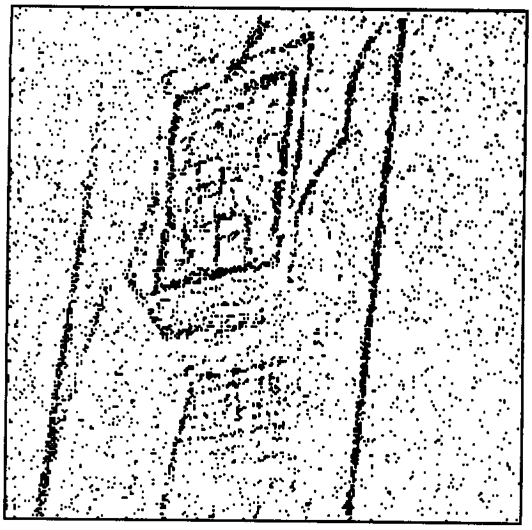
combination of two real edge maps, the dominant distracting feature is an irregular curve; the straight edge consists of merely 420 of the 6998 edge points. Figure 6 shows the average number of false alarms in this case, as a function of sample size. The threshold effect is clear.

The next experiment was designed to test the suggested approach when the distracting features are themselves straight line segments. Figure 7(a) shows an edge map which contains two significant straight edges, one longer than the other, and several shorter straight segments. Figures 7(b) and (c) are noisy

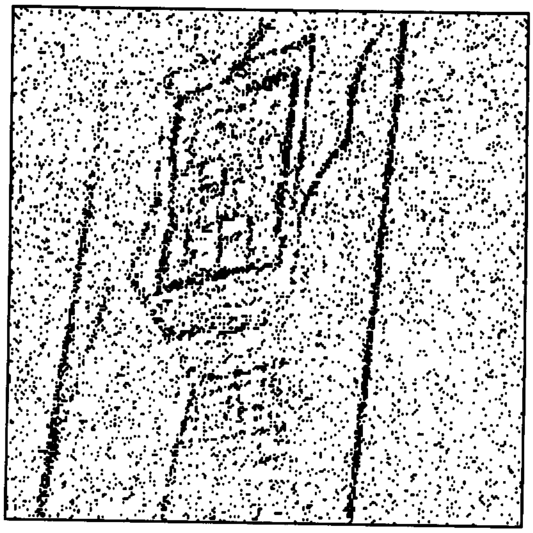


(a)

Fig. 7. (continued on next page).



(b)



(c)

Fig. 7. (a) An edge map containing two significant straight edges, one longer than the other, several shorter straight edges and some random noise. Out of a total of 4000 edge points, 326 points lie on the longer straight edge and 234 points lie on the other. (b) A similar edge map with 2000 uniformly distributed noisy edge points added. (c) An edge map similar to (a) with 4000 noisy edge points added.

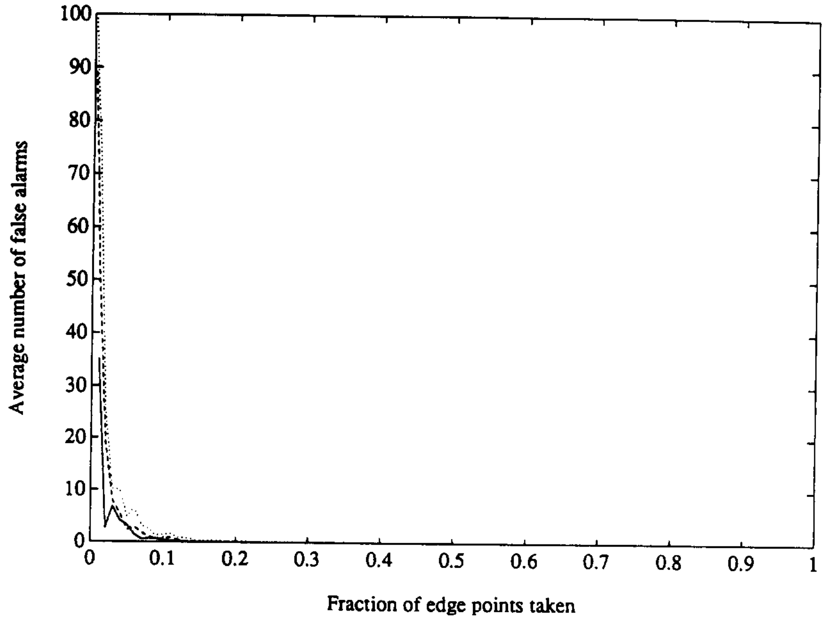


Fig. 8. The average number of false alarms as a function of the fraction of edge points taken. The continuous line corresponds to Fig. 7(a). The dashed line and the dotted line correspond to Fig. 7(b) and 7(c) respectively. The robustness of the suggested approach to the presence of distracting straight edges as well as significant random noise is evident.

versions of the original edge map of Fig. 7(a). Figure 8 shows the average number of false alarms (with respect to the longest edge) as a function of sample size, and exhibits a strong threshold effect.

Figure 9 shows the average number of false alarms with respect to the *two* most significant edges in Fig. 7(a), i.e. the number of peaks in the accumulator array larger than at least one of the two peaks that correspond to the longest straight edges. For comparison the graph of the average number of false alarms with respect to the (single) longest edge is provided as well. While the threshold effect is still evident, it is clear that the sample size must be increased if the false alarms rate should not change while the number of features to be detected grows. This means that as the complexity of the task which the Hough algorithm is expected to perform increases, more of the available data should be used. This observation seems to be in line with the conclusions of Grimson and Huttenlocher.[15]

To test whether the threshold effect that makes the suggested probabilistic approach attractive depends on the specific Hough algorithm used, circle detection experiments were carried out. The three dimensional ($128 \times 128 \times 60$) accumulator array, which is needed since the radius is assumed to be unknown, and the use of directional edge information (with incrementation guided by an error model) make this Hough algorithm quite different to the algorithm for straight line detection.

Figure 10 shows a textured circle covered by a distracting curve; the corresponding edge map is shown in Fig. 11. Figure 12 shows the average number of false alarms as a function of sample size, and exhibits a clear threshold.

In Fig. 13(a) the feature to be detected is again the circle, the distractors being straight edges. Figures 13(b) and (c) are noisy versions of Fig. 13(a), and Figs 14(a), (b) and (c) are the respective edge maps. The average number of false alarms in the three cases as a function of sample size is shown in Fig. 15, and again shows a sharp threshold effect, indicating the applicability of the suggested probabilistic approach to the Hough algorithm for circle detection.

4. DISCUSSION

The theoretical analysis and the various experiments demonstrate that even in the presence of distracting features, significant noise and errors in the coordinates of the data points, large computational savings are often possible by using the suggested Probabilistic Hough Transform. The results also show another aspect of the robustness of conventional Hough algorithms.

An alternative approach, based on the Random Sample Consensus (RANSAC) paradigm,[16] is suggested in reference (7). To detect straight edges, pairs of data points are consecutively selected at

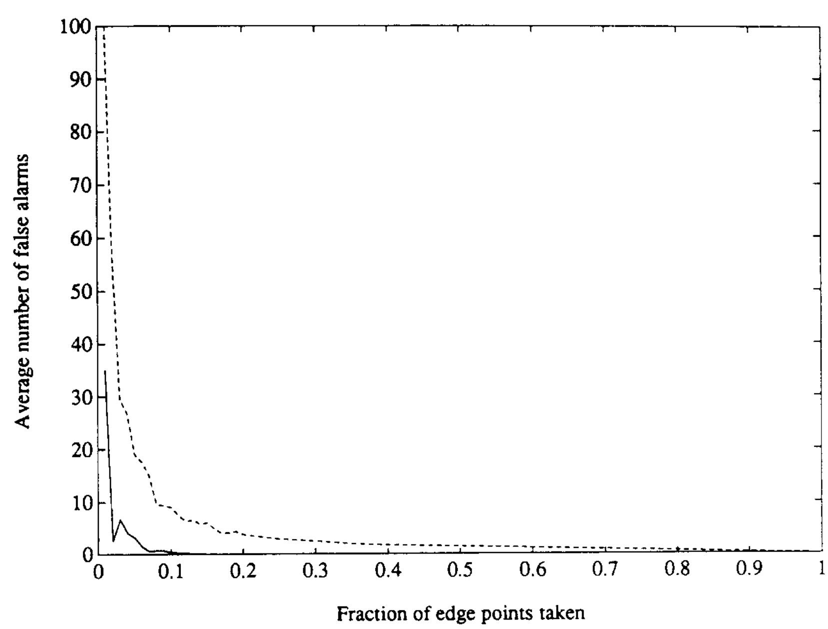


Fig. 9. The dashed line shows the average number of false alarms with respect to the two most significant edges in Fig. 7(a). For comparison, the solid line shows the false alarms with respect to the single longest edge. While it is clear that the sample size must be increased if a certain fixed false alarms rate is to be maintained, the threshold effect remains evident.

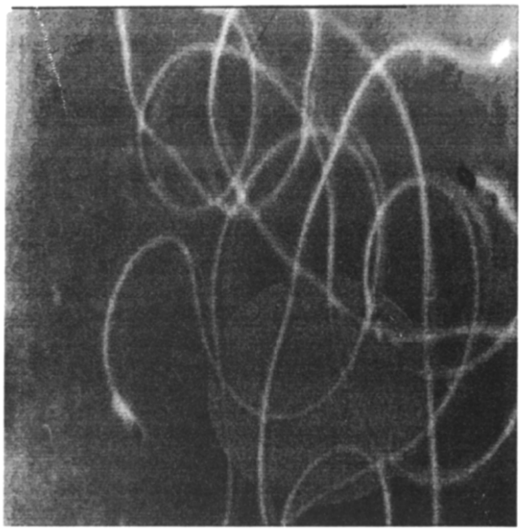


Fig. 10. A grey scale image showing a textured circle covered by a distracting curve.

Fig. 11. An edge map obtained from the image shown in Fig. 10 by the Marr-Hildreth edge detection algorithm. Out of a total of 6231 edge points, 105 lie on the circle's edge.

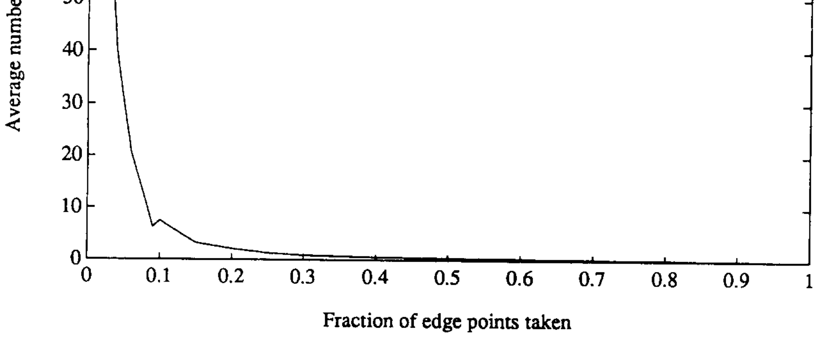
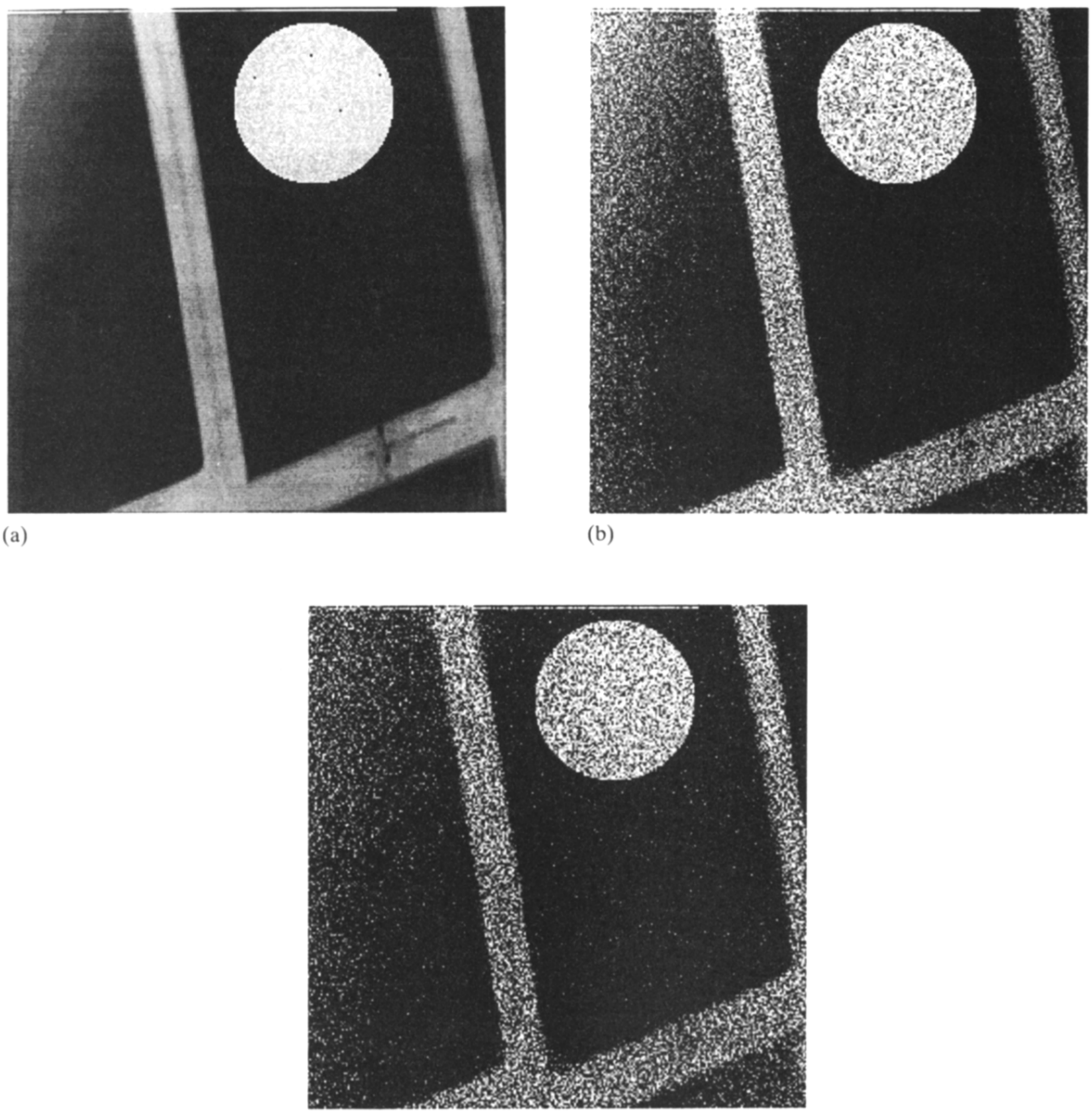
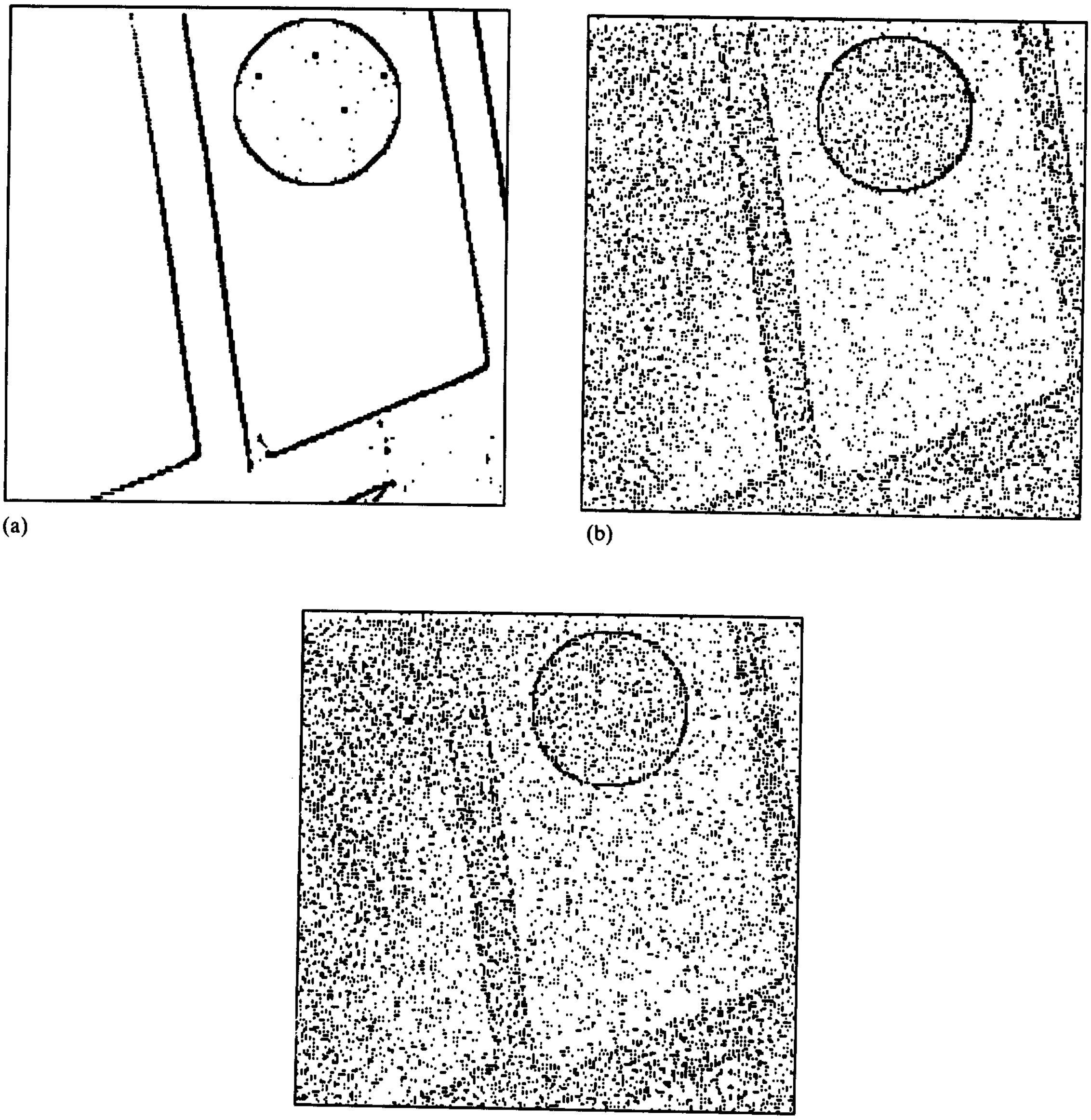


Fig. 12. The average number of false alarms as a function of sample size when a probabilistic circle detection Hough algorithm was applied to the edge map of Fig. 11. Directional edge information was used. The clear threshold effect indicates that significant computational savings can be obtained even in such a problem, which is rather difficult for edge based circle detection algorithms.



(a) (b) (c)

Fig. 13. (a): A grey level image that contains a circular feature, several straight edges and noise. (b) and (c) are similar, with different levels of additive Gaussian noise.



(a) (b) (c)

Fig. 14. The edge maps that correspond to the images shown in Fig. 13. About 10% of the points lie on the circle's edge in Fig. 14(a), about 3% in Fig. 14(b) and about 1.5% in Fig. 14(c). Note that directional edge information has also been severely distorted by the addition of noise to the original image.

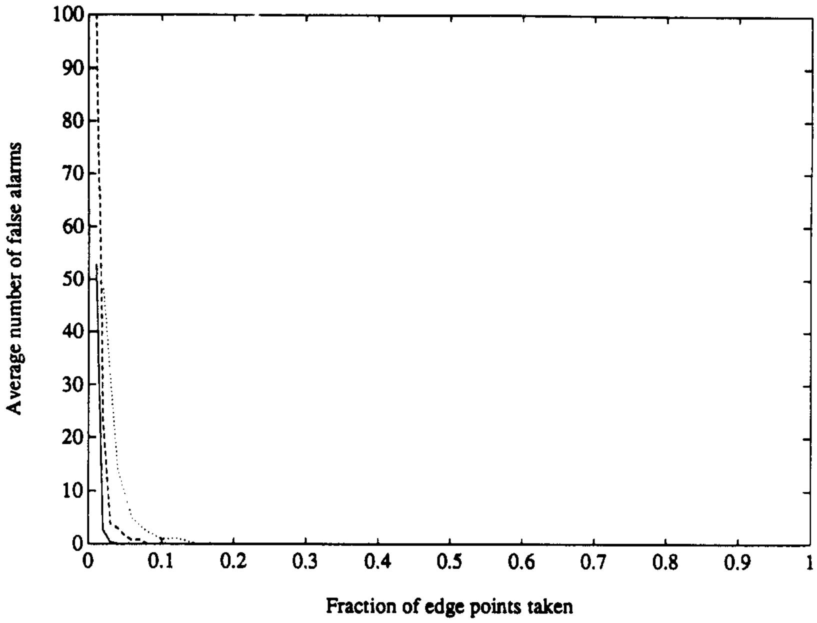


Fig. 15. The average number of false alarms as a function of sample size. The dotted, dashed and solid lines correspond to the edge maps shown in Figs 14(c), (b) and (a), respectively. The sharp threshold demonstrates against the robustness of the suggested approach.

random, and the goodness of fit of the line that they define is verified by evaluating its distance from all other edge points, until an acceptable fit is reached. That technique, possibly with minor modifications aimed at lowering the computational cost of the verification process, can yield good results, especially when the errors in the coordinates of edge points are very small, and when the S/N level is not too low. Reference (17) regards the conventional Hough Transform, which emphasizes evidence collection, and techniques similar to RANSAC, that stress hypotheses generation and verification with minimum data gathering, as two ends of a spectrum, and suggests, in the context of model-based vision, a dynamic compromise between those extremes. See also reference (18). The Probabilistic Hough Transform in conjunction with a verification algorithm can be regarded as a possible compromise. It is interesting to note that the results of a recent study[15] concerning the robustness of Ballard's Generalized Hough Transform[19] imply that in demanding applications of that algorithm to complex high dimensional problems all available data should preferably be used.

SUMMARY

The well-known Hough Transform is a standard method for the detection of predefined features in images. Its usual implementation consists of a voting stage, in which each feature point is mapped into a curve (hypersurface) in a parameter space, and an exhaustive search for peaks. The parameter space is usually quantized and represented by an accumulator array.

One of the major obstacles to wide-spread application of the Hough Transform has been the large number of operations, which is esentially due to the dependence of the number of operations in the voting stage on the (large) number of edge points in the image. Parallel processing using standard or specialized hardware is an excellent way to reduce computation time, but usually requires costly or bulky equipment. "Coarse to fine" multiresolution processing in the parameter space had been suggested, but several theoretical and practical limitations have been observed.

This paper suggests an alternative approach to fast Hough Transform computation. Its essence is the replacement of full scale voting in the incrementation stage of the algorithm by a limited poll of a small number of edge points, selected at random. This leads to large computational savings. The key to successful application of the "Probabilistic Hough Transform" is the dependence of the algorithm's performance (in terms of detection probability and false alarm rate) on the fraction of the data that is used. The expected number of peaks in the parameter space that are higher than the "true" peak is especially important if a post-processing verification algorithm is considered.

The dependence of the algorithm's performance on the fraction of edge points used has been analysed for the case of straight line detection. An image model that consists of considerable background noise and one straight edge segment with significant edge point location errors has been assumed. The analysis shows that the dependence of the performance on the fraction of edge points used exhibits a sharp threshold effect, implying that considerable computational savings are possible with negligible performance degradation.

The applicability of this interesting result in the processing of real images (that deviate in many ways from the model used in the analysis) has been investigated experimentally. The threshold effect predicted by the theoretical analysis appeared even when the image contained several distracting features in addition to various kinds of noise. This means that large computational savings can often be achieved by using the Probabilistic Hough Transform.

REFERENCES

1. J. Illingworth and J. Kittler, A survey of the Hough transform, *Comput. Vision Graphics Image Process.* **44**, 87–116 (1988).
2. R. O. Duda and P. E. Hart, Use of the Hough transformation to detect lines and curves in pictures, *Comm. Ass. Comp. Mach.* **15**, 11–15 (1972).
3. S. D. Shapiro, Transformations for the computer detection of curves in noisy pictures, *Comput. Graphics Image Process.* **4**, 328–338 (1975).
4. S. D. Shapiro, Properties of transforms for the detection of curves in noisy pictures, *Comput. Graphics Image Process.* **8**, 219–236 (1978).
5. A. Iannino and S. D. Shapiro, A survey of the Hough transform and its extensions for curve detection, *Proc. IEEE Conf. Pattern Recognition Image Process. PRIP'78*, pp. 32–38, Chicago (1978).
6. N. Kiryati and A. M. Bruckstein, Antialiasing the Hough transform, *Proc. IAPR 6th Scandinavian Conf. Image Anal.*, pp. 621–628, Oulu, Finland (1989). Accepted for *Comput. Vision Graphics Image Processing: Graphical Models and Image Processing (CVGIP: GMIP)*.
7. M. A. Fischler and O. Firschein, Parallel guessing: a strategy for high-speed computation, *Pattern Recognition* **20**, 257–263 (1987).
8. G. Gerig, Linking image-space and accumulator space: a new approach for object recognition, *Proc. IEEE 1st Int. Conf. Comput. Vision*, pp. 112–117, London (1987).
9. N. Kiryati, Y. Eldar and A. M. Bruckstein, Probabilistic Hough transform, *Proc. 7th Israeli Conf. Artif. Intell., Vision Pattern Recognition*, pp. 475–482, Tel-Aviv (1989).
10. T. M. Van Veen and F. C. A. Groen, Discretization errors in the Hough transform, *Pattern Recognition* **14**, 137–145 (1981).
11. W. Niblack and D. Petkovic, On improving the accuracy of the Hough transform: theory, simulations and experiments, *Proc. IEEE Conf. Comput. Vision Pattern Recognition CVPR'88*, pp. 574–579 (1988).
12. M. Cohen and G. T. Toussaint, On the detection of structures in noisy pictures, *Pattern Recognition* **9**, 95–98 (1977).
13. C. M. Brown, Inherent bias and noise in the Hough transform, *IEEE Trans. Pattern Anal. Mach. Intell.* **5**, 493–505 (1983).
14. V. F. Leavers and J. F. Boyce, The Radon transform and its application to shape parametrization in machine vision, *Image Vision Comput.* **5**, 161–166 (1987).
15. W. E. L. Grimson and D. P. Huttenlocher. On the sensitivity of the Hough transform for object recognition, A.I. Memo No. 1044, Artificial Intelligence Lab., MIT (1988).
16. M. A. Fischler and R. C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Comm. Ass. Comp. Mach.* **24**, 381–395 (1981).
17. Y. Kuno, K. Ikeuchi and T. Kanade, Model-based vision by cooperative processing of evidence and hypotheses using configuration spaces, *SPIE* **938**, Digital and Optical Shape Representation and Pattern Recognition (1988).
18. R. Sitaraman and A. Rosenfeld, Probabilistic analysis of two stage matching, *Pattern Recognition* **22**, 331–343 (1989).
19. D. H. Ballard, Generalizing the Hough transform to detect arbitrary shapes, *Pattern Recognition* **13**, 111–122 (1981).

About the Author—Nahum Kiryati was born in Haifa, Israel, on 25 July 1958. He received the B.Sc. degree in Electrical Engineering from Tel Aviv University, Tel Aviv, Israel, in 1980, the post-B.A. degree in the Humanities from Tel Aviv University in 1986, and the M.Sc. degree in Electrical Engineering from the Technion, Israel Institute of Technology, Haifa, Israel, in 1988. He is presently with the Technion, studying towards the D.Sc. degree and working as an instructor. His fields of interest are image processing and computer vision.

About the Author—Yuval Eldar was born in Haifa, Israel, on 8 September 1963. He received the B.Sc. degree in Computer Engineering from the Technion, Israel Institute of Technology, Haifa, Israel, in 1989. He is currently a student in the Technion. His fields of interest are biological and computer vision.

About the Author—Alfred M. Bruckstein was born in Sighet, Transylvania, Romania, on 24 January 1954. He received the B.Sc. and M.Sc. degrees in Electrical Engineering, from the Technion, Israel Institute of Technology, in 1977 and 1980, respectively, and the Ph.D. degree in Electrical Engineering from Stanford University, Stanford, California, in 1984.

From October 1984 until June 1989 he has been with the Faculty of Electrical Engineering at the Technion, Haifa. Presently he is with the Faculty of Computer Science there. His research interests are in computer vision, image processing, estimation theory, signal processing, algorithmic aspects of inverse scattering, point processes and mathematical models in neurophysiology.

Professor Bruckstein is a member of SIAM, MAA and AMS.