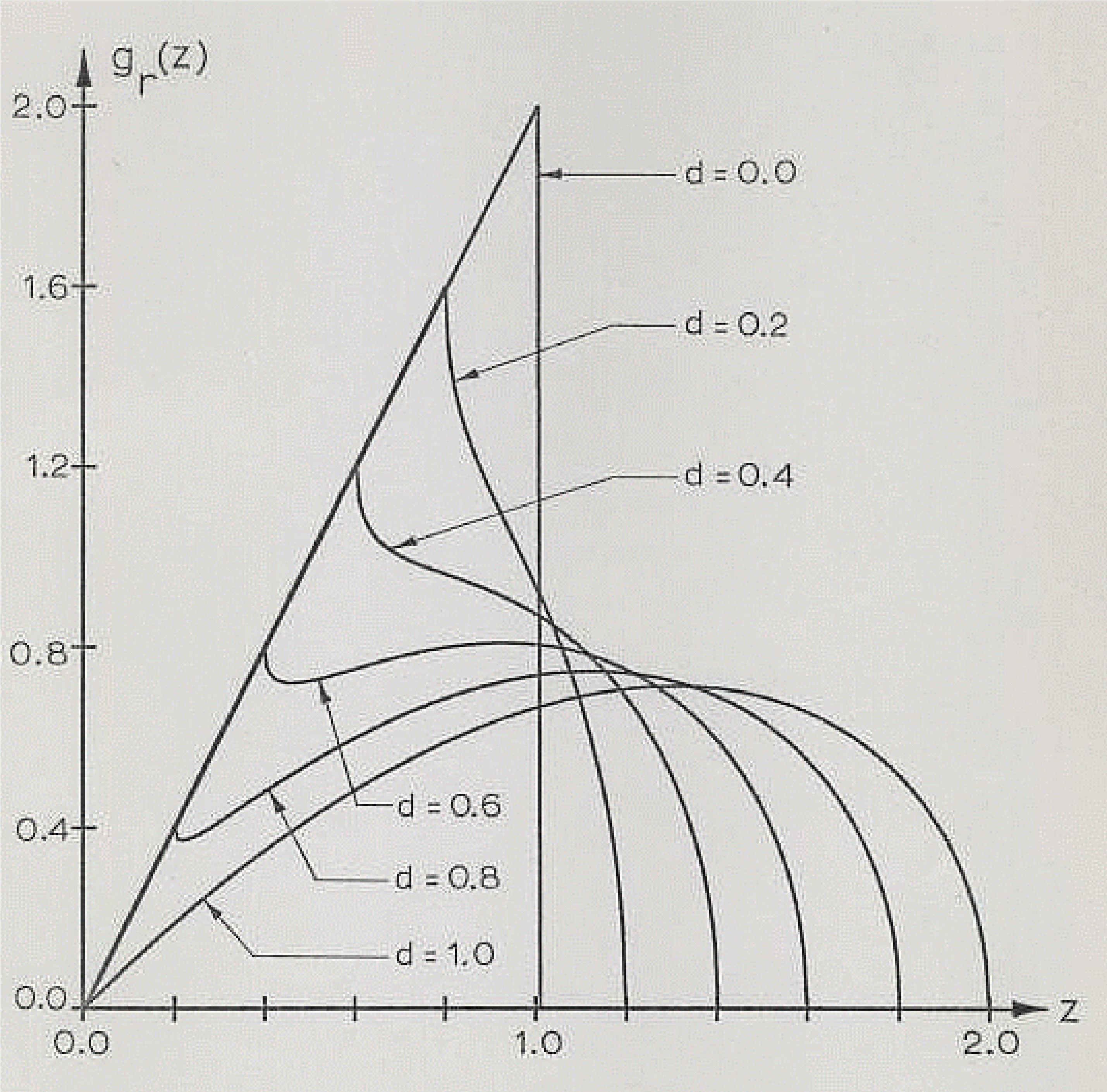

Fig. 3 The density function governing *r* for six values of *d*.

to distinguish high counts then determines the confidence level with which we reject the null hypothesis. With maximum entropy quantization each cell count has the same binomial distribution and, hence, the same threshold applied to different cells reflects the same confidence level. With the second compensation method each cell count, h, can have a different binomial distribution. However, as N gets large, the modified counts $(h - \mu)/\sigma$ all approach the standard normal distribution. Hence, in our second procedure, the same threshold applied to different cells reflects the same confidence level in the limit.

It is clear that the procedures in this paper can be extended to the detection of any structure or pattern for which a parameterization exists. The effects of noise on the parameters can be determined theoretically, or the probabilities associated with each of the cells can be estimated from data. In either case, once these effects have been characterized the detection procedure can be made to compensate for them.

6. Summary

Hough proposed an algorithm for detecting lines in pictures. His algorithm, based on a slope-intercept parameterization of lines, was improved by Duda and Hart through the use of an angle-radius parameterization. Duda and Hart's procedure works acceptable well with pictures that are relatively noise-free. When pictures contain noise that cannot be removed, however it can yield unsatisfactory results. This paper presents two approaches to compensate for noise. One approach is applicable when the distribution of the noise is known. Knowledge of the distribution allows us to compensate for the noise by producing a histogram through maximum entropy quantization. The second approach can be used when nothing is known about the noise. In this case histogram counts are first estimated with pictures containing only background or noise. These estimates are then subsequently used for standardizing the counts obtained in regular pictures. The proposed approach is illustrated for the case of line and circle detection on a circular retina containing independent, additive, uniformly distributed random noise.

7. Acknowledgment

The second author proposed the Duda-Hart procedure as a test for random number generators. George Marsaglia pointed out that such a test was inadequate because it did not take into account the distribution of ρ, and suggested that the Duda-Hart procedure could be improved. This was the motivation for the work presented here.

8. References

1. A. Rosenfeld, *Picture Processing by Computer*, Academic Press, NY (1969).
2. P. V. C. Hough, Method and means for recognizing complex patterns, U.S. Patent 3,069,654. Dec. (1962).
3. R. O. Duda and P. E. Hart, Use of the Hough transformation to detect lines and curves in pictures, *Comm. ACM* **15**, 11-15 (1972).
4. C. Kimme, Finding circles by an array of accumulators, *Comm. ACM* **18**, 120-122 (1975).

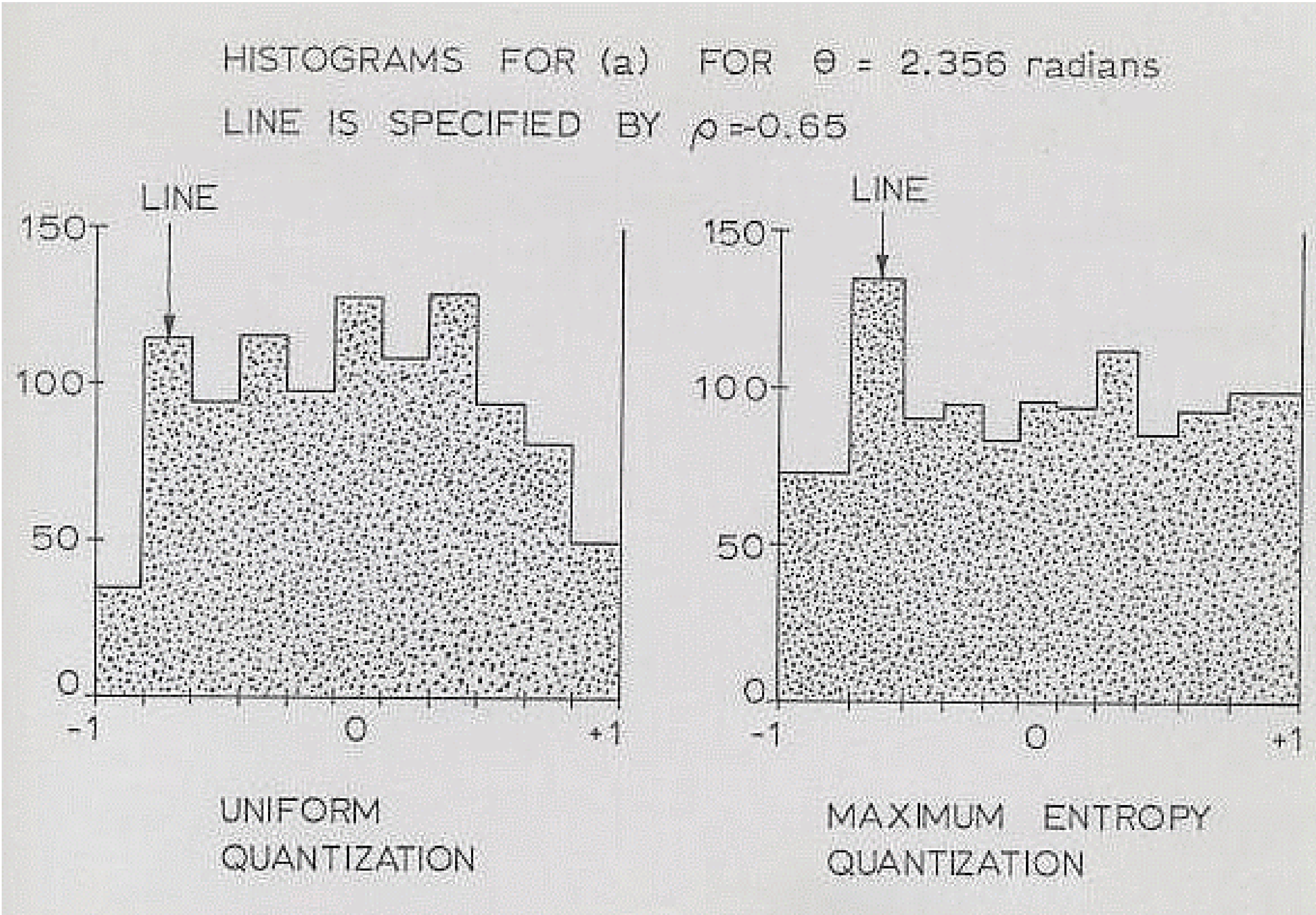

Fig. 2(b) Histograms for picture in (a), when θ = 2.356 radians, for uniform and maximum entropy quantization.

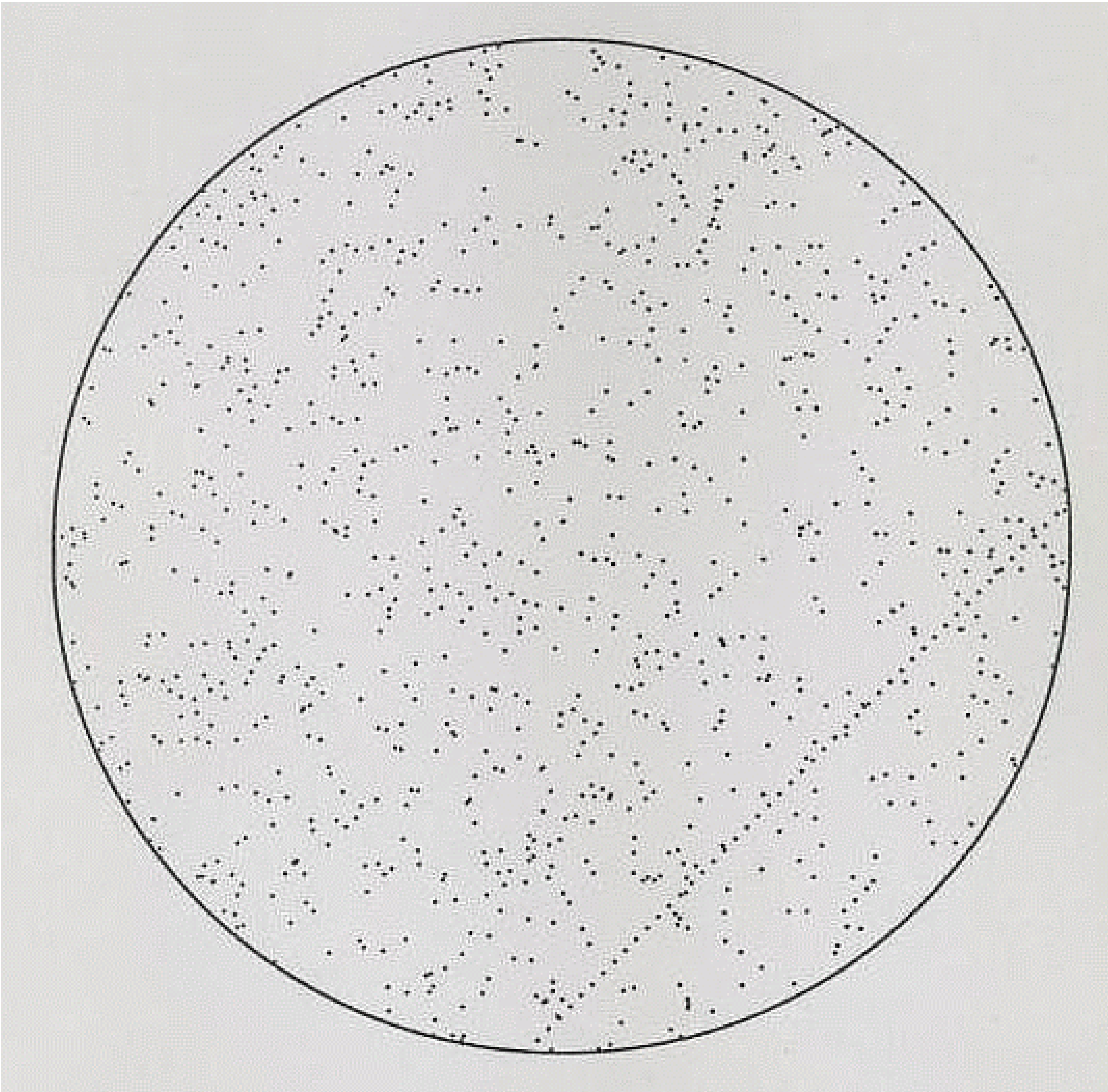

Fig. 2(a) 1000 noise points and a line consisting of 50 points at $\rho = -0.65$ and $\theta = 2.356$ radians.

It is knowledge of this density that enables us to compensate for the noise, by changing the quantization of the ρ-axis from equal-length sub-intervals to equal probability sub-intervals. This is maximum entropy quantization.

To partition the interval (-1, +1) into m equal-probability sub-intervals, we should obtain the sub-interval endpoints $-1 = x_0 < x_1 < x_2 <...< x_{m-1} < x_m = 1$ for which the probability $P(x_{i-1} < \rho < x_i) = 1/m$, or equivalently, for which $F_\rho(x_i) = i/m$ $i = 0, 1,..., m$, where $F_\rho(x)$ is the cumulative distribution function of ρ. To actually determine in which sub-interval a particular ρ_j occurs, we need not search through the above-mentioned x_i's. Instead, the sub-interval 0, 1,..., m-1 can be directly obtained by computing the greatest integer less than $mF_\rho(y)$ for $y - \rho_j$. With this technique the computation involved does not grow as the number of sub-intervals used increases and it can be faster than searching through a large number of sub-interval end-points.

Now consider the problem of detecting circles, also against a background of uniform noise. For each point (a,b) the noise induces a distribution for r that depends on the distance $d = \sqrt{a^2 + b^2}$ of (a,b) from the origin. The density function of r is given by

$$g_r(z) = \begin{cases} 2z & 0 < z < 1 - d \\ z + \dfrac{2z}{\pi d}\left\{\sqrt{1-q^2} - \sqrt{z^2 - W^2} + d\left(\csc\left(\dfrac{W}{z}\right)\right)\right\} & 1 - d < z < 1 + d \end{cases}$$

where $W = (1 - z^2 - d^2)/2d$ and $q = (1 - z^2 + d^2)/2d$. This density function is illustrated in Fig. 3 and is quite peaked for some values of d. These peaks might well overshadow the presence of some circles, and can be compensated for by using a maximum entropy quantization of r.

4. The Modification When Noise is of Unknown Distribution

For a given quantization, the count, h, observed in any cell can be standardized by replacing h with $(h - \mu)/\sigma$ where μ and σ are the mean and standard deviation, respectively, of h. By doing this we give each cell a more equal opportunity to "show off" any high count it may have.

Given θ or (a,b) in the case of the line or circle detection, respectively, estimate for each quantization interval the probability, p, that a picture point is projected to that interval. This estimate, $\bar{p}$, can be obtained from pictures containing only background or noise and no lines nor circles. Then for that cell, h has $\bar{\mu} = N\bar{p}$ and $\sigma = [N\bar{p}(1-\bar{p})]^{1/2}$ where N is the number of points in the picture.

It should be noted that this second modification can also be used when the distribution of the noise is known, in which case p is obtained from the distribution induced on ρ or r.

5. Discussion and Conclusions

From a statistical standpoint the count for each histogram cell can be viewed as a binomial random variable with parameters N and p. Detecting a structure then becomes a matter of rejecting the null hypothesis that a particular accumulator count is indeed binomial (N,p). The threshold used

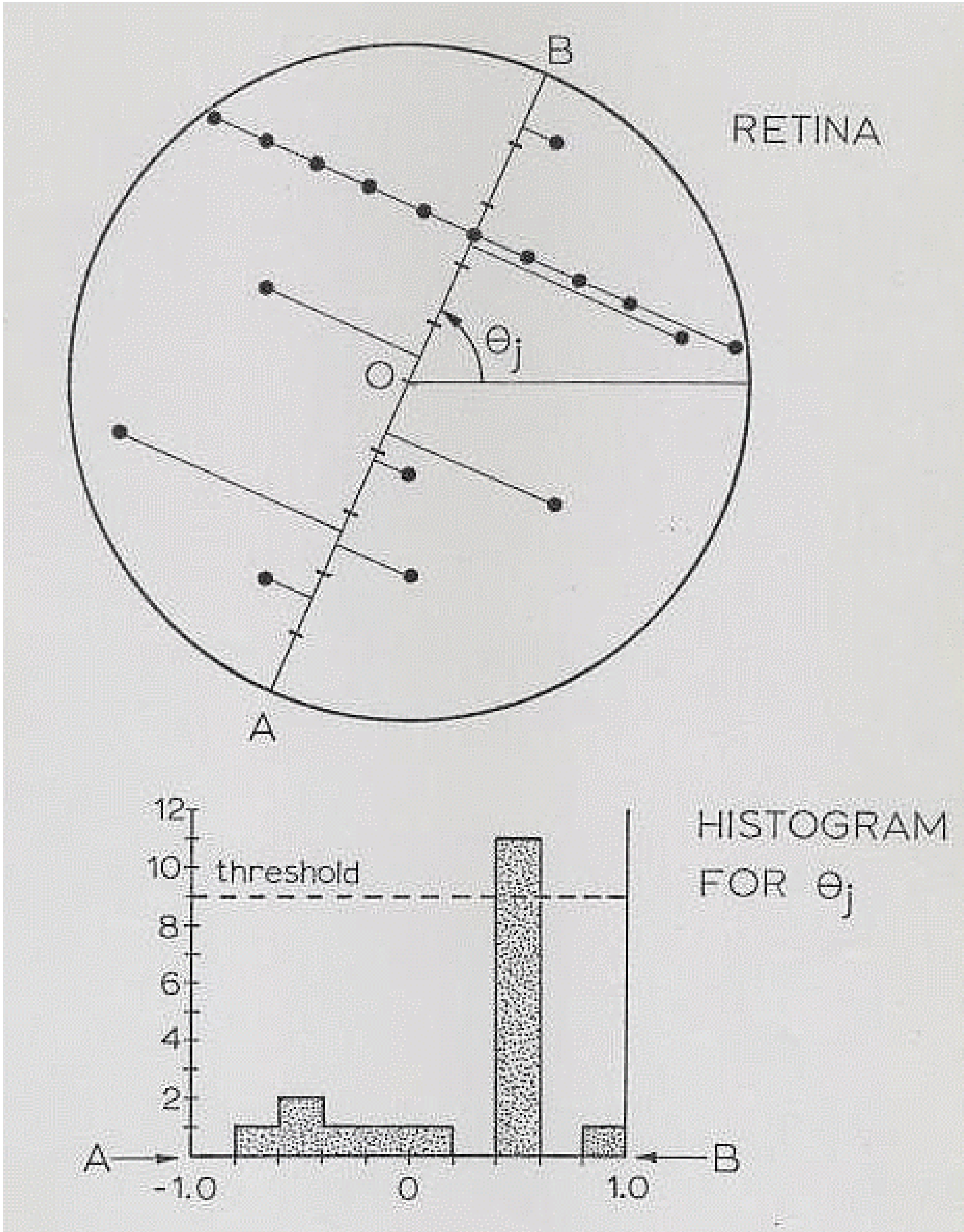

Fig. 1 continued)

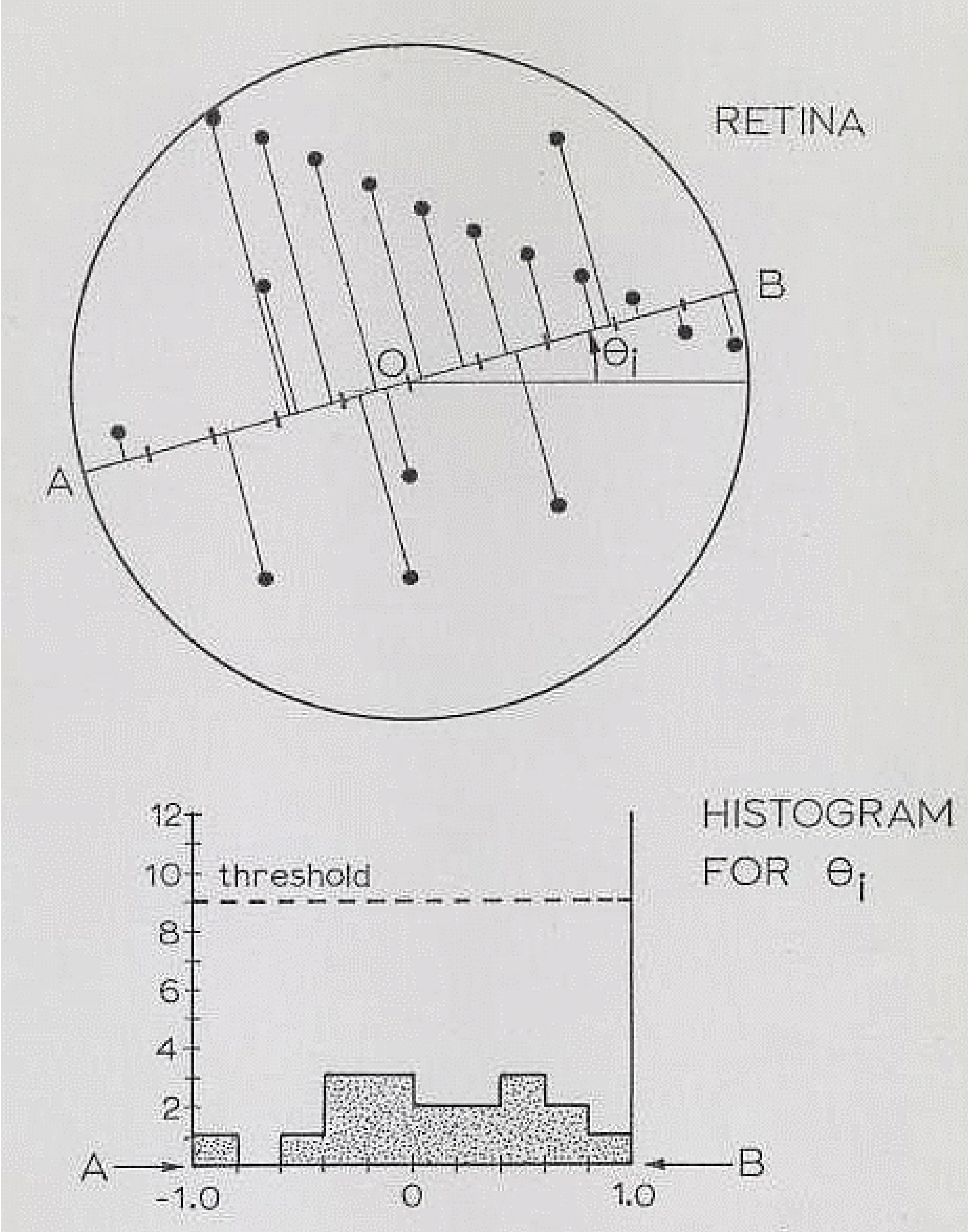

Fig. 1 A picture abd corresponding histograms for two angles illustrating the Duda-Hart line detection procedure. (cont. next page)

2. The Duda-Hart Procedure

Before describing our modifications it is instructive to review the procedure of Duda and Hart[(3)]. Their procedure, as well as ours, is based on the normal parameterization of a line. Such a parametrization specifies a straight line by its distance ρ from the origin and the angle θ of its normal.

Computational efficiency is realized by specifying an acceptable tolerance in ρ and θ and quantizing the ρ-θ plane into a quadruled grid. If we quantize ρ to take on *m* values and θ to take *n* values, then the ρ-θ plane has the capacity to specify *n* x *m* different lines. The line detection procedure is implemented by considering the quadruled grid as a two-dimensional histogram or array of accumulators. The procedure is illustrated in Fig. 1 where the original picture, in this case a dotted line and some noise, is displayed on a circular retina of radius one, centered at the origin.

In Fig. 1 ρ is quantized uniformly into 10 levels which are marked off on the ρ-axis. For each quantized value of θ a one-dimensional histogram is obtained by projecting all the picture points to the ρ-axis of the corresponding θ value and counting the number of points that fall in each cell. Fig. 1 illustrates this process for two values of θ: θ_i in Fig. 1(a) and θ_j in Fig. 1(b). To complete the line detection procedure we need only specify a suitable threshold for the histogram counts. If the count in a particular cell in the ρ-θ histogram exceeds the threshold, a line is detected and specified by the ρ and θ values of the corresponding cell.

Duda and Hart suggested that their line detection procedure could be extended to detect other curves. As an example they consider detecting circles by choosing a parametric representation, *a-b-r*, for the family of all circles having centers within the retina, where *a* and *b* are the Cartesian coordinates of the center point of a circle, and *r* is the radius. The process can be implemented efficiently by using a three-dimensional array of accumulators representing the three-dimensional parameter space. A circle can be detected by looking for a high count in the accumulator array, where the counts are obtained by going to each quantized value of (*a,b*), computing all the distances from the picture points to (*a,b*), and recording these distances on a quantized r axis. Kimme et al.[(4)] applied a similar procedure to the detection of chest tumors.

3. The Modification When Noise is of Known Distribution

Consider first the problem of detecting lines on a circular retina with independent, additive, uniformly distributed random noise, and suppose we were to apply Duda and Hart's procedure to a picture containing only noise and no lines. For any given angle θ, the counts for the small-magnitude ρ values will be considerably higher than for those further away from the center because larger regions of the retina project to the middle portions of the ρ-axis. This is undesirable, for if high histogram counts are used to indicate the presence of lines, we would hope that a retina containing nothing but noise would yield a histogram with roughly equal counts for each sub-interval. In fact, the high counts near the center can actually overshadow the increased counts that result from a line near the outskirts of the retina, as illustrated by the histogram for uniform quantization in Fig. 2.

This tendency of ρ to assume values close to zero is explicitly described by its probability density function: $f_\rho(y) = \frac{2}{\pi}\sqrt{1-y^2}$ $\quad -1 < y < 1$.

On the Detection of Structures in Noisy Pictures*

Melvin Cohen
and
Godfried Toussaint
School of Computer Science
McGill University
Montreal, Quebec, Canada

Published in: ***Pattern Recognition***, Vol. 9, 1977, pp. 95-98.

ABSTRACT

Hough proposed an algorithm for detecting lines in pictures. His algorithm, based on a slope-intercept parameterization of lines, was improved by Duda and Hart through the use of angle-radius parameterization. When pictures contain random noise that cannot be removed, the Duda-Hart procedure can yield unsatisfactory results. This paper presents two modifications of that procedure which compensate for noise. One method is applicable when the distribution of the noise is known and the other can be used when it is not. The proposed modification is also illustrated for circle detection.

Picture processing, Pattern recognition, Detection in noise, Geometrical probability.

1. Introduction

An important and frequently occurring problem in digital image processing is the detection of straight lines. In a simple version of this problem the digitized image contains a set of black points, representing one or more lines, lying on a white background. The problem is then one of detecting the presence of collinear points.

Rosenfeld(1) described an interesting method proposed earlier by Hough(2) in which the black points are transformed into lines in a slope-intercept parameter space. Duda and Hart(3) showed that the unboundedness of the slope and intercept complicates the application of Hough's technique and suggested the angle-radius parameterization as a solution to this problem.

Duda and Hart's procedure works acceptably well with pictures that are relatively noise-free. Furthermore, with pictures that contain essentially unbroken lines, spurious noise can be removed by standard techniques before the procedure is applied. However, when pictures contain discontinuous lines or lines composed of non-adjacent black points (as in the case of bubble-chamber photographs of particle tracks) the noise removal techniques fail because they remove the dotted lines as well.

---

* This work was supported by the National Research Council of Canada under grant NRC-A9293. An earlier version of this paper was presented at the CSCSI Workshop on Computational Studies of Intelligence, University of Ottawa, 28-29 May, 1975.