TR-90-003

Statistical Line Detection and Its Extensions

TR-90-00

Statistical Line Detection and Its Extensions

Jun S. Huang and Yu Fu Chang

Computer Vision Laboratory

Institute of Information Science

Academia Sinica

Taipei, Taiwan, Republic of China

Abstract :

The use of statistical analysis in line detection is necessary because of noises induced by signal amplification, and random variations due to the microtexture of the object surface. In this paper , we first derive two likelihood ratio tests for detecting a line in both gray level and binary images. These tests have some invariance properties. We also present some experimental results of detecting a line in the simulated image and also an edge line in the real image to demonstrate the usefulness of the tests. Then extensions of these tests to detect a parametric curve or a general shape are discussed in detail. Finally, a complete analysis of Bayesian approach to line detection , particular in normal distributed case, is carried out successfully, and practical considerations of the whole theory are discussed with a conclusion that the theory is realistic and can be applied in many practical situations, and in some cases better than Hough transform.

Index terms — Line detection and estimation , Simulation , Likelihood ratio test , Bayesian analysis , normal and binomial distributions .

1. INTRODUCTION

Line detection plays an important role in image analysis. If the image background is fairly uniform with little variations and the signal of a line is significant , then we can filter the image and use some heuristic methods (*e.g.* the Hough transform) to detect a line. But when the image background is pretty complicated (*e.g.* just like a wire passing through a forest) then detecting a line is a difficult problem and it is a challenging research topic. When the image background is not very complicated but corrupted by noises, which can be modelled statistically, then we can develop a statistical line detection theory and this is what concerned in this paper.

The use of statistical analysis for line detection is necessary because of noises induced by signal amplification, and random variations due to the microtexture of the object surface. In a recent paper by Hunt, Nolte and Ruedger [1], the authors try to develop a statistical detection theory on line detection and relate it to Hough transform. However, the model they have developed is clearly incomplete and not adequate in the case of normally distributed observations. Hence the explanation they made for the relationship between the statistical detection theory and the Hough transform is unsatisfactory. We know that the Hough transform has a serious quantization problem when the image is associated with noises and is not capable of detecting short line segments when noises are present. Therefore a rigorous statistical theory of line detection is needed. Here we try to develop a theory based both on the likelihood ratio test and Bayesian analysis.

The likelihood ratio tests are in general consistent [2] and yield sensible good tests. Unfortunately, the likelihood ratio tests we developed here for line detection are clearly not uniformly most powerful and also not consistent , but they are reasonable and have intuitive meanings and some invariance properties. The power of these tests depend heavily on the mean difference between the line and the background, and also the variance of the

noise. These tests are similar to Hough transform in that all possible line segments need to be computed, but there are differences when compared to Hough transform , for example , complicated statistics are computed, other than just counting the number in each cell. A statistical decision rule is given to decide if a line exists or not, whereas the Hough transform can not make a decision like this because of noise corruption. A good survey of Hough transform is given by Illingworth and Kittler [3]. In this paper , we also conduct some experiments for line detection both in simulated images and real images and the results are quite successful. This shows that the likelihood ratio tests are reliable.

Extensions to detecting a general curve or general shape are discussed in detail. These reveal that our method is similar to template matching for images taken with negligible noises and unstable lighting condition , where the correlation coefficient is used as a matching criterion. Subsequently, a detail analysis of Bayesian approach to inference a line is given and a complicated decision rule is derived for the case of normal (Gaussian) observations . Finally, practical considerations and discussions of the whole theory given here are presented and a conclusion is drawn that the theory is realistic and can be applied in many practical situations, and in some cases better than Hough transform .

2. THE LIKELIHOOD RATIO TESTS FOR DETECTING A LINE

Suppose we are given an image containing noises and only a straight line passing through it. We try to develop a statistical model to detect and locate the line. The problem will be discussed in two cases : the gray level case and the binary case.

1. The gray level case. Let the image $\{ X_{ij} \mid i=1,...,m ; j=1,...,n \}$ has an uniform background added by Gaussian white noises (*i.e.* independent identically normally distributed data), and a line (white or black or gray) passing through it. Let ϵ_{ij}'s be independent normally distributed with mean 0 and variances σ^2, which is unknown, i=1,...,m ; j=1,...,n. Let μ_1 be the unknown uniform background intensity and μ_2 be the unknown line intensity. Then the line detection problem can be formulated as the following

hypotheses testing :

$$H_0 : X_{ij} = \mu_1 + \epsilon_{ij} , \qquad \forall\, i,j \text{ and } \mu_1 \text{ unknown},$$

$$H_1 : X_{ij} = \mu_1 + S_{ij} + \epsilon_{ij} , \qquad \forall\, i,j \text{ and}$$

$$S_{ij} = \begin{cases} \mu_2 \text{ (unknown and nonzero)} \quad , \text{ if (i,j) is on the line } L_1 : \\ \gamma = i\cos\theta + j\sin\theta, \text{ where } \gamma \text{ and } \theta \text{ are unknown and quantized,} \\ \\ 0 \quad , \text{otherwise.} \end{cases}$$

Testing H_0 against H_1 using likelihood ratio criterion is developed here. The likelihood function under H_0 is given by

$$L(\mu_1, \sigma_0^2 | H_0) = (2\pi\sigma_0^2)^{-mn/2} \exp\{-1/(2\sigma_0^2)\Sigma_i\Sigma_j(X_{ij}-\mu_1)^2\} ,$$

and the maximum likelihood estimates of μ_1 , σ_0^2 are

$$\hat{\mu_1} = \overline{X} = \Sigma_i\Sigma_j X_{ij} / mn ,$$
$$\hat{\sigma_0}^2 = \Sigma_i\Sigma_j (X_{ij} - \overline{X})^2 / mn.$$

Under H_1, the likelihood function is given by

$$L(\mu_1, \mu_2, \sigma^2, L_1 \mid H_1) = (2\pi\sigma^2)^{-mn/2} \exp\{-1/(2\sigma^2)[\sum_{(i,j)\notin L_1} \Sigma_i \Sigma_j (X_{ij}-\mu_1)^2 + \sum_{(i,j)\in L_1} \Sigma_i \Sigma_j (X_{ij}-\mu_1-\mu_2)^2]\}.$$

The maximum likelihood estimates of parameters μ_1, μ_2, σ^2, L_1 are :

$$\hat{\mu}_1 = \overline{X}_k^{\,\prime} = \sum_{(i,j)\notin L_1} \sum_i \sum_j X_{ij} \Big/ (mn-k),$$ where k is the number pixels in L_1 ,

$$\hat{\mu}_1+\hat{\mu}_2 = \sum_{(i,j)\in L_1} \sum_i \sum_j X_{ij} \Big/ k = \overline{X}_k ,$$

$$\hat{\mu}_2 = \overline{X}_k - \overline{X}_k^{\,\prime} ,$$

$$\hat{\sigma}^2 = \Big[\sum_{(i,j)\notin L_1} \sum_i \sum_j (X_{ij} - \hat{\mu}_1)^2 + \sum_{(i,j)\in L_1} \sum_i \sum_j (X_{ij} - \hat{\mu}_1 - \hat{\mu}_2)^2 \Big] \Big/ mn ,$$

and $\hat{L}_1$ is the L_1 that maximizes $L(\hat{\mu}_1\ \hat{\mu}_2\ ,\hat{\sigma}^2\ ,L_1 \mid H_1)$. Please note that L_1 corresponds to the unknown parameter (γ,θ). Then the likelihood ratio is

$$\lambda = \frac{L(\ \hat{\mu}_1,\hat{\mu}_2,\hat{\sigma}^2,\ \hat{L}_1\ \mid H_1\ )}{L(\ \hat{\mu}_1,\ \hat{\sigma}_0^2\ \mid H_0\ )} = \max_{L_1} \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right)^{mn/2} .$$

However, $mn\hat{\sigma}_0^2 = \Sigma_i\Sigma_j\ (X_{ij} - \overline{X}\ )^2 = mn\hat{\sigma}^2 + \dfrac{k\ (mn-k)}{mn}\ (\overline{X}_k - \overline{X}_k^{\,\prime})^2$, and hence the likelihood ratio test is

$$\text{reject } H_0 \qquad \text{if } W^2 = \max_{L_1} \frac{|\overline{X}_k - \overline{X}_k^{\,\prime}|^2}{\hat{\sigma}^2.\left(\dfrac{1}{mn-k} + \dfrac{1}{k}\right)} > \xi_\alpha ,$$

where ξ_α is the critical value at level α ,and $\hat{\sigma}^2$ is modified to be

$$\hat{\sigma}^2 = \frac{1}{mn-2}\Big[\sum_{(i,j)\notin L_1} \sum_i \sum_j (X_{ij} - \overline{X}_k^{\,\prime})^2 + \sum_{(i,j)\in L_1} \sum_i \sum_j (X_{ij} - \overline{X}_k)^2 \Big],$$

an unbiased estimate of σ^2 under H_1. The critical value ξ_α can be determined from the

distribution of W^2 statistic under H_0. However, there is no theoretical way of computing the distribution of W^2 statistic except by computer simulation. The simulation process is described as follows . First, generate the random normal data as the image background by Box–Muller transformation [4,p.86] of the uniform data , generated on VAX–11/8530 by RANDU routine . The uniform randomness of the data is tested by Kolmogorov–Smirnov test. The simulated image's resolution is set to be 100×100 . An algorithm is developed to generate all straight lines of two pixel width and of different values of (γ,θ). Having generated all 100×100 normal data with zero means and variances one ,the quantity $|\bar{X}_k - \bar{X}'_k|^2/[\hat{\sigma}^2.(1/(mn-k) + 1/k)]$ is computed for each line L_1 which contains totally k pixels,$k \geq 6$. Then the maximum is found over all (γ,θ)'s. This process is repeated 500 times and the distribution of these 500 values of W^2 is plotted in Fig. 1. The critical values are listed in Table 1.

α	.10	.05	.01
ξ_α	19.95	21.24	23.98

Table 1. Critical values of W^2 for 100×100 image.

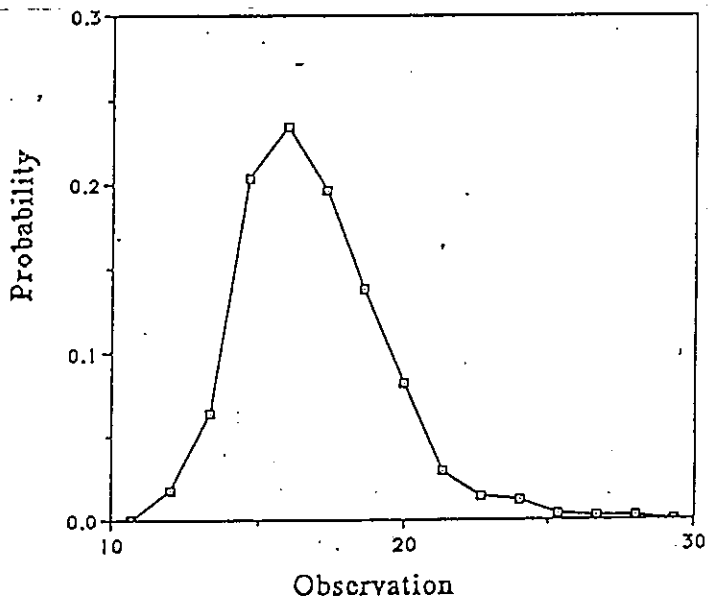


Fig 1. The distribution of W^2 for 100×100 image.

For an input image with Gaussian background, we can make a decision whether there is a line in the image or not ,by checking its W^2 value . If $W^2 > \xi_\alpha$,then a line is present in the image and its location is the value of (γ,θ) that gives this W^2 (or the maximum of $|\bar{X}_k-\bar{X}_k'|^2/[\hat{\sigma}^2.(1/(mn-k)+1/k)]$.

Now question !! Is the decision $W^2 > \xi_\alpha$ invariant under the translation and scaling of the observations? That is, if we have another image with different contrast or brightness will the decision be unchanged ? The answer is yes by looking at $|\bar{X}_k-\bar{X}_k'|^2$ and $\hat{\sigma}^2$. However, this likelihood ratio test is clearly neither uniformly most powerful nor consistent [2]. The power of this test depends on the mean difference between the line and the background ,and also the variance of the noise (*i.e.* $|\mu_2| \ / \ \sigma$). If a line falls on the central region of an image then the test has higher power than that falling near the boundary region. If a line consists of very few pixels , then unless the value of $|\mu_2| \ / \ \sigma$ is large, the detection will probably fail.

Experiments on real and simulated images are conducted to verify our proposed decision procedure. First, three simulated images, resolution 200×200 , each containing a line of two pixels wide (γ=75 , θ=20), are generated and shown in Fig.2 . The background data are normally distributed with mean 128 and variance 25. The mean difference between the background and the line (*i.e.* $|\mu_2|$) is 20 in (a), 40 in (b), 60 in (c). Since each image is of size 200×200, we have to divide each into 4 subimages of 100×100 and run the likelihood ratio test on each subimage. The results are shown in Fig.3, where (a) corresponds to Fig.2(a) ($|\mu_2|$=20), (b) to Fig.2(b) ($|\mu_2|$=40), (c) to Fig.2(c) ($|\mu_2|$=60). The first two maximum W–squares in (a) or (b) have values less than $\xi_{.10}$ = 19.95 indicating that no line in the corresponding subimages. The rest two maximum W^2 in (a) or (b), having significant values greater than $\xi_{.10}$, reveal that there are lines in the corresponding subimages, and the (γ,θ) values of corresponding lines are (45.48,19) and (10.64,20) in (a), and (44.99,21) and (10.64,20) in (b). Note that , because the center of coordinate system is moved to each subimage's center, so the γ values of lines are different in all subimages. The

original θ is 20 and two estimated θ's are correct, which corresponding to subimages with long lines. The other two estimated θ's are 19 and 21, pretty close to 20, which correspond to subimages with short lines, and the little error is due to quantization error and two pixels wide line.

Secondly, two real images of 200×200, each containing an edge line, are taken and shown in Fig. 4(a). Sobel operators are used to find the edge strength of each pixel and the results are shown in Fig. 4(b). Then each edge map is divided into 4 subimages of 100×100 and each subimage is tested to see whether a line exists. The results are shown in Fig. 5 , where 5(a) corresponds to the left image of Fig. 4(b) and 5(b) to the right image. All the subimages containing an edge line segment give very high W^2 values, indicating that existence of a line is highly significant. Those regions containing no line segment have low W^2 values, smaller than critical value $\xi_{.10}$ except one with $W^2 = 26.89$ which is significant. By checking the subimage with this significant W^2 value, we see that the exception is due to the texture of background where significant line segments exist. After all, the estimated value of θ in the left image are 170, 171 ,which are all right, and 13, 11, 11 in the right image . The error is still small so we can make a conclusion that the tests are all successful.

2. *The binary case.* -Let X_{ij}'s be independent Bernoulli distributed B(1,p); p is the probability of success. X_{ij} has value 0 or 1 for all i,j . In general , the image background has sparsely distributed 1's, and the line, if exists , has densely distributed 1's, or vice versa. Thus to detect a line in such a binary image we can set up the following two hypotheses :

$$H_0 : X_{ij} \sim B(1,p_0) , \forall\, i,j , \quad p_0 \text{ unknown},$$

$$H_1 : \begin{cases} X_{ij} \sim B(1,p_0) , & \text{if } (i,j) \notin L_1 : \text{an unknown line}, \\ X_{ij} \sim B(1,p_1) , & \text{if } (i,j) \in L_1 , \; p_1 \text{ unknown and } p_1 \neq p_0 , \end{cases}$$

and try to develop a test statistic to test whether H_0 is true or H_1 is true. If H_1 is true we then locate the line. As before, we compute the two likelihood functions :

$$(1) \qquad L(p_0 \mid H_0) = \prod_{ij} p_0^{X_{ij}}(1-p_0)^{1-X_{ij}} = p_0^{\Sigma_i\Sigma_j X_{ij}}(1-p_0)^{mn-\Sigma_i\Sigma_j X_{ij}} ,$$

and the maximum likelihood estimate of p_0 ,under H_0 , is

$$\hat{p}_0 = \frac{\Sigma_i\Sigma_j X_{ij}}{mn} = \bar{X} .$$

$$(2)\ L(p_0,p_1,L_1 \mid H_1) = \prod_{(i,j)\notin L_1} p_0^{X_{ij}}(1-p_0)^{1-X_{ij}} \prod_{(i,j)\in L_1} p_1^{X_{ij}}(1-p_1)^{1-X_{ij}}$$

$$= p_0^{\Sigma_{(i,j)\notin L_1} X_{ij}} (1-p_0)^{(mn-k)-\Sigma_{(i,j)\notin L_1} X_{ij}} \cdot p_1^{\Sigma_{(i,j)\in L_1} X_{ij}} (1-p_1)^{k-\Sigma_{(i,j)\in L_1} X_{ij}} ,$$

where k is the number of pixels in L_1. The maximum likelihood estimates of p_0 ,p_1 ,L_1 are

$$\hat{p}_0 = \Sigma_{(i,j)\notin L_1} X_{ij} \Big/ (mn-k) = \bar{X}_k' ,$$
$$\hat{p}_1 = \Sigma_{(i,j)\in L_1} X_{ij} \Big/ k = \bar{X}_k ,$$
$\hat{L}_1$ is the L_1 that maximizes $L(\hat{p}_0,\hat{p}_1,L_1 \mid H_1)$.

The log likelihood ratio is

$$\log\lambda = \max_{L_1} \log \frac{L(\hat{p}_0,\hat{p}_1,\hat{L}_1 \mid H_1)}{L(\hat{p}_0 \mid H_0)}$$

$$= \max_{L_1} \{ (mn-k)[\bar{X}_k' \log(\bar{X}_k'/\bar{X}) + (1-\bar{X}_k')\log(\frac{1-\bar{X}_k'}{1-\bar{X}})] + k[\bar{X}_k \log(\bar{X}_k/\bar{X}) + (1-\bar{X}_k)\log((1-\bar{X}_k)/(1-\bar{X}))]\}.$$

We reject H_0 and accept H_1 if $\log \lambda > \xi_\alpha$,a critical value. The problem associated with this testing statistic is " Does the invariant property still hold, under translation and scaling of observations ? ". Because there is no easy way to find an invariant testing statistic, we check this by simulation. Using simulated 100×100 images, under null hypothesis (*i.e.* $p_0 = p_1$), log λ's are calculated 500 times for a fixed p_0 , and the corresponding critical values are extracted. Table 2. summarizes these critical values at different values of p_0. From this table we see that these critical values are different, but not far apart at each fixed type *I* error α. We conclude that there is no significant evidence to prove the invariance of logλ statistic , under different values of p_0. But here we will take the risk and use the average of critical values. Also the simulated density plots of log λ for $p_0 = .2$, .5 , .8 are shown in Fig. 6 where 6(a) has bandwidth 15 and 6(b), 10. From these plots we see that the densities of different p_0's are nearly the same. Note that ,in our simulation process ,we adopt the same random number sequence for all cases.

$p_0 \backslash \alpha$	.10	.05	.01
.2	10.25	10.96	12.19
.3	9.84	10.51	12.13
.4	9.77	10.22	12.02
.5	10.07	10.92	12.07
.8	10.24	10.98	12.73
Average	10.03	10.72	12.23

Table 2. Critical values of log λ .

A simulated 200×200 binary image containing a line, two pixels wide, is shown in Fig. 7(a) where $p_0 = .3$, $p_1 = .8$, $\gamma = 75$ and $\theta = 20$. This image is divided into 4 subimages of size 100×100 and the likelihood ratio test is run on each subimage. The test results are shown in Fig. 7(b). The first two subimages have log λ values less than the

average of $\xi_{.10} = 10.03$,indicating that no line is in these subimages . The other two subimages have log λ values far greater than the average of $\xi_{.01} = 12.23$, indicating that a line passes through the corresponding subimage. The estimated (γ,θ)'s are correct and consistent with the normal case shown in Fig. 3 .

3. EXTENSIONS TO GENERAL CURVES AND SHAPES

We have developed two statistics to detect a line in an image with uniform noisy background. Each statistic needs computing the maximum value of certain functions over all possible lines in the image. Clearly, the maximum operation can be extended to all parametric curves or nonparametric curves without changing the formulas. For example , if the curve to be detected is a parabola , then all possible parabolas inside the image are considered, and all possible parameter values of a parabola are searched. If the curve to be detected is a square, then we have to consider all possible squares by translating, rotating, and scaling a standard square. Formulas for these statistics are unchanged but the distribution of each statistic will be changed and so will be the critical values. This is the main shortcoming of statistical approach compared with the Hough transform. In fact, the statistical method presented here is similar to template matching where the correlation coefficient is used as a matching criterion, under negligible noises and unstable lighting.

In the case of slight variations of positions of line pixels, we can consider searching lines with two or three pixel wide in evaluating $\overline{X}_k$. This will improve the detection power. If there are n lines to be detected , then we have to consider all possible configurations of n lines which tends to have exponential complexity. If only one line segment is to be detected then a hybrid method , by looking at a small window and using some heuristic ideas, may be more reliable and efficient.

4. BAYESIAN INFERENCE OF A LINE

Let an image be represented by a set of random variables $\{ X_{ij} \mid -m \leq i \leq m, -n \leq j \leq n\}$, and the parameter space (γ,α) of a line, represented by $\gamma = i\cos\alpha + j\sin\alpha$, be quantized into discrete values $\{ (\gamma_k,\alpha_l) \mid 0 \leq \alpha_l < 2\pi , 0 \leq \gamma_k \leq (m^2+n^2)^{1/2}, l=1,...,b; k=1,...,a\}$. The image is said to have a line $L(\gamma,\alpha)$ located at (γ,α) , if $X_{ij} \sim F_1(x|\theta_1)$ for $(i,j) \notin L(\gamma,\alpha)$ and $X_{ij} \sim F_2(x|\theta_2)$ for $(i,j)\in L(\gamma,\alpha)$, where $F_1(x|\theta_1) \neq F_2(x|\theta_2)$ and θ_1,θ_2 may be multidimensional parameters (known or unknown). We shall consider the situation in which F_1 and F_2 have known functional forms, but the location (γ,α) of the line L is unknown. Given $\{X_{ij}$'s$\}$, the problem concerned is that of making Bayesian inference about (γ,α).

Assuming that the distributions $F_1(x|\theta_1)$ and $F_2(x|\theta_2)$ admit densities $p_1(x|\theta_1)$ and $p_2(x|\theta_2)$ respectively, and X_{ij}'s are independent for all i and j ,the joint distribution of X_{ij}'s conditioned on θ_1,θ_2 and line $L(\gamma,\alpha)$,is given by

$$\begin{aligned} & p(x_{ij} , (i,j)\in I \mid (\gamma,\alpha),\theta_1,\theta_2) \\ & \quad = p_1(x_{ij} ,(i,j)\in I\backslash L \mid \theta_1).p_2(x_{ij} ,(i,j)\in L \mid \theta_2) \\ & \quad = \prod_{(i,j)\in I\backslash L} p_1(x_{ij}|\theta_1) \cdot \prod_{(i,j)\in L} p_2(x_{ij}|\theta_2). \end{aligned} \tag{4.1}$$

where I denotes the set of coordinates of all image points.

Further, we assume a prior distribution, to be specified over the set of possible line parameter values : (γ_k,α_l), k=1,...,a; l=1,...,b ; *i.e.* we specify $p_0(\gamma,\alpha)$ such that $\Sigma_k\Sigma_l\, p_0(\gamma_k,\alpha_l) = 1$. The actual specification must be made in the light of knowledge of the particular application. If the prior knowledge is vague an uniform prior is assumed in general. Also, if θ_i , for some i, is unknown we can assign a prior distribution to it.

Further analysis of the problem depends critically on the assumptions made concerning θ_1 and θ_2.

Case 1 : Both θ_1 and θ_2 are known. From Bayes theorem ,it is easily seen that given θ_1, θ_2 the posterior densities of the possible line, having observed x_{ij}'s , are given by

$$p_{mn}(\ (\gamma,\alpha)|\theta_1,\theta_2) \propto p(x_{ij}'s|(\gamma,\alpha),\theta_1,\theta_2)\ .\ p_0(\gamma,\alpha).$$

Thus by (4.1)

$$\begin{aligned} & p_{mn}((\gamma,\alpha)\mid\theta_1,\theta_2) \Big/ p_0(\gamma,\alpha) \\ & \propto \prod_{(i,j)\in\Lambda\setminus L} p_1(x_{ij}|\theta_1)\ .\ \prod_{(i,j)\in L} p_2(x_{ij}|\theta_2) \\ & \propto \prod_{(i,j)\in L} p_2(x_{ij}|\theta_2) \Big/ \prod_{(i,j)\in L} p_1(x_{ij}|\theta_1) \end{aligned} \qquad (4.2)$$

For a given $p_0(\gamma,\alpha)$, it follows that $p_{mn}((\gamma,\alpha)\mid\theta_1,\theta_2)$ is large when the likelihood ratio of " line " over " no line " , when based on the observations in L , is large. If

$$p_0(\gamma,\alpha) = \frac{1}{ab} \quad \text{,for all } \gamma,\ \alpha$$

(*i.e.* $p_0(\gamma,\alpha)$ has a uniform prior density) , the posterior mode is also the maximum likelihood estimate ,and is given by the value of (γ,α) which maximizes the right–hand side of (4.2).

Case 2. : θ_1 is known and θ_2 is unknown. We assign a prior density $p_0(\theta_2|\theta_1)$, independent of $p_0(\gamma,\alpha)$, over Θ_2 : the range of possible values of θ_2, which could depend on θ_1. It follows that for given θ_1 ,

$$p_{mn}((\gamma,\alpha)|\ \theta_1) \propto \int_{\Theta_2} p(x_{ij}'s|(\gamma,\alpha),\theta_1,\theta_2).p_0(\theta_2|\theta_1)\ d\theta_2.p_0(\gamma,\alpha).$$

From (4.2) we have

$$p_{mn}((\gamma,\alpha) \mid \theta_1) \Big/ p_0(\gamma,\alpha)$$

$$\propto \frac{\int_{\Theta_2} \prod_{(i,j)\in L} p_2(x_{ij}|\theta_2).p_0(\theta_2|\theta_1)d\theta_2}{\prod_{(i,j)\in L} p_1(x_{ij}|\theta_1)} .$$

Thus $p_{mn}((\gamma,\alpha) \mid \theta_1) \,/\, p_0(\gamma,\alpha)$ is now determined by the expected likelihood ratio with respect to the prior for θ_2 based on the observations in L. The marginal posterior density for θ_2 is given by

$$p_{mn}(\theta_2|\theta_1) = \sum_{\gamma}\sum_{\alpha} p_{mn}(\theta_2|(\gamma,\alpha),\theta_1).p_{mn}((\gamma,\alpha) \mid \theta_1) ,$$

where

$$p_{mn}(\theta_2|(\gamma,\alpha),\theta_1) \propto p(x_{ij}'s|(\gamma,\alpha),\theta_1,\theta_2).p_0(\theta_2|\theta_1) .$$

The maximum likelihood estimates of θ_2 and (γ,α) correspond to the joint posterior mode resulting from the choice $p_0(\gamma,\alpha) = 1 \,/\, ab$, with a possibly improper, uniform prior for θ_2. Defining $\hat{\theta}_{2,(\gamma,\alpha)}$ to be the value of θ_2 maximizing $\prod_{(i,j)\in L} p_2(x_{ij}|\theta_2)$, it is easy to see that the maximum likelihood estimates are given by $(\hat{\gamma},\hat{\alpha})$ and $\hat{\theta}_{2,(\hat{\gamma},\hat{\alpha})}$, where $(\hat{\gamma},\hat{\alpha})$ maximizes

$$\prod_{(i,j)\in L} p_2(x_{ij}|\hat{\theta}_{2,(\gamma,\alpha)}) \Big/ \prod_{(i,j)\in L} p_1(x_{ij}|\theta_1) .$$

Case 3. : Both θ_1 and θ_2 are unknown. Assigning a prior density $p_0(\theta_1,\theta_2)$ over Θ_{12} , the range of possible values of (θ_1,θ_2), independent of $p_0(\gamma,\alpha)$, we obtain

$$p_{mn}(\gamma,\alpha) \propto p(x_{ij}'s|(\gamma,\alpha)).p_0(\gamma,\alpha) \qquad (4.3)$$

where

$$p(x_{ij}'s|(\gamma,\alpha)) = \int_{\Theta_{12}} p(x_{ij}'s|(\gamma,\alpha),\theta_1,\theta_2).p_0(\theta_1,\theta_2)\ d\theta_1\ d\theta_2. \tag{4.4}$$

Inference about θ_1 and θ_2 can be based upon

$$p_{mn}(\theta_1,\theta_2) = \sum_{\gamma}\sum_{\alpha} p_{mn}(\theta_1,\theta_2|(\gamma,\alpha)).p_{mn}(\gamma,\alpha)\ ,$$

where

$$p_{mn}(\theta_1,\theta_2|(\gamma,\alpha)) \propto p(x_{ij}'s|(\gamma,\alpha),\theta_1,\theta_2).p_0(\theta_1,\theta_2). \tag{4.5}$$

When we assume uniform priors, the joint posterior mode, which gives the maximum likelihood estimates, is at $(\hat{\gamma},\hat{\alpha})$, $\hat{\theta}_{1,(\hat{\gamma},\hat{\alpha})}$,and $\hat{\theta}_{2,(\hat{\gamma},\hat{\alpha})}$ where $(\hat{\gamma},\hat{\alpha})$ maximizes

$$\prod_{(i,j)\in\Lambda\setminus L} p_1(x_{ij}|\hat{\theta}_{1,(\gamma,\alpha)}) \cdot \prod_{(i,j)\in L} p_2(x_{ij}|\hat{\theta}_{2,(\gamma,\alpha)})\ ,$$

and $\hat{\theta}_{1,(\gamma,\alpha)}$ maximizes $\prod_{(i,j)\in\Lambda\setminus L} p_1(x_{ij}|\theta_1)$, and $\hat{\theta}_{2,(\gamma,\alpha)}$ is defined similarly.

Estimates for (γ,α), θ_1 and θ_2 are easily obtained by minimizing the posterior expected loss for some appropriate choice of loss function, which is quadratic in most applications. In the case of quadratic loss function the estimate of each unknown parameter is the mean of the posterior distribution of that parameter , given the observations. As a Bayesian solution to the hypothesis testing problem , we see that posterior odds ratio is easily calculated [5,p.146] and in general the Bayes factor B = (posterior odds ratio)/(prior odds ratio) is the " weighted " likelihood ratio. For example, if S is a subset of the set of all possible (γ,α)'s, testing $H_0 : (\gamma,\alpha)\in S$ versus $H_1 : (\gamma,\alpha)\notin S$ is provided by calculating $p_{mn}(S).(1-\pi_0) \ / \ \{\pi_0.[1-p_{mn}(S)]\}$, where

$$p_{mn}(S) = \sum_{(\gamma,\alpha)\in S} \sum p_{mn}(\gamma,\alpha) \text{ , and } \pi_0 = \sum_{(\gamma,\alpha)\in S} \sum p_0(\gamma,\alpha).$$

5. BAYESIAN ANALYSIS WITH NORMAL OBSERVATIONS

Let both θ_1 and θ_2 be unknown and both F_1 and F_2 be normal (Gaussian). Taking $\theta_i = (\mu_i, \tau_i)$, $i = 1,2$, the mean and the precision (reciprocal of the variance), respectively , then (4.1) can be written as

$$\begin{aligned} & p(x_{ij}\text{'s} \mid (\gamma,\alpha),\theta_1,\theta_2) \\ & = (\tau_1/2\pi)^{(mn-a)/2}\,(\tau_2/2\pi)^{a/2}\exp\{-(\tau_1/2)\sum\sum_{(i,j)\in I\setminus L}(x_{ij}-\mu_1)^2\} \\ & \qquad .\exp\{-(\tau_2/2)\sum\sum_{(i,j)\in L}(x_{ij}-\mu_2)^2\}, \end{aligned} \tag{5.1}$$

where a is the number of observations in $L(\gamma,\alpha)$. In considering the prior distributions of each (μ_i,τ_i) , it is better to select the conjugate exponential family [2,p.77].
It is not hard to see that the conjugate prior distribution of (μ_i,τ_i) is given by

$$\begin{aligned} \pi(\mu_i,\tau_i \mid t_{i1},t_{i2},t_{i3}) \propto & (t_{i3}\,\tau_i/2\pi)^{1/2}\exp\{-(t_{i3}\,\tau_i/2)\,.\,(\mu_i-(t_{i2}/t_{i3}))^2\} \\ & .\,\tau_i^{(t_{i3}-1)/2}\exp\{\,[\,(t_{i2}^2/t_{i3})-t_{i1}\,]\,.\,(\tau_i/2)\,\}, \end{aligned}$$

where , $t_{i3} > 0$, $t_{i1} > (t_{i2}^2/t_{i3}) \geq 0$.

Conditioned on τ_i, μ_i is assumed to have normal prior with mean t_{i2}/t_{i3} and precision $t_{i3}\tau_i$, and τ_i is assumed to have marginal gamma distribution $\Gamma((t_{i3}+1)/2,(t_{i1}-(t_{i2}^2/t_{i3}))/2\,)$. By properly selecting t_{i1},t_{i2},t_{i3} we can obtain the desired prior distribution with realistic meanings. Thus

$$p_0(\theta_1,\theta_2) = \pi(\mu_1,\tau_1 \mid t_{11},t_{12},t_{13})\,.\,\pi(\mu_2,\tau_2 \mid t_{21},t_{22},t_{23}).$$

Fixing (γ,α) , *i.e.* fixing L, the posterior distribution of μ_1,τ_1 is also normal–gamma

distribution with parameters

$$t_{11}^{*} = t_{11} + \sum\sum_{(i,j)\in I\setminus L} x_{ij}^{2} ,$$
$$t_{12}^{*} = t_{12} + \sum\sum_{(i,j)\in I\setminus L} x_{ij} ,$$
$$t_{13}^{*} = t_{13} + mn - a \quad . \tag{5.2}$$

Similar results also hold for (μ_2,τ_2). From these results and (4.4),(4.5),(4.6) we see that,

$$p_{mn}(\gamma,\alpha) \;/\; p_0(\gamma,\alpha) \propto \Gamma((1+t_{13}^{*})/2).\Gamma((1+t_{23}^{*})/2) \;/\; (t_{13}^{*}t_{23}^{*})^{1/2}$$
$$. (t_{11}^{*} - (t_{12}^{*2}/t_{13}^{*}))^{-(1+t_{13}^{*})/2}.(t_{21}^{*} - (t_{22}^{*2}/t_{23}^{*}))^{-(1+t_{23}^{*})/2}, \tag{5.3}$$

where t_{11}^{*}, t_{12}^{*}, t_{13}^{*}, are given by (5.2) and

$$t_{21}^{*} = t_{21} + \sum\sum_{(i,j)\in L} x_{ij}^{2} ,$$
$$t_{22}^{*} = t_{22} + \sum\sum_{(i,j)\in L} x_{ij} ,$$
$$t_{23}^{*} = t_{23} + a \quad . \tag{5.4}$$

The right hand side of (5.3) is the multiplication of two multivariate elliptical t–distributions [6,p.48]. In particular, for noninformative prior, letting $t_{i2} = 0$, $t_{i3}\to 0$, $t_{i1} \to 0$ we find that the right hand side of (5.3) becomes

$$\prod_{i=1}^{2} \Gamma((1+a_i)/2).a_i^{-1/2}\{\sum\sum_{(k,l)\in I_i} (x_{kl} - \bar{x}_i)^2\}^{-(1+a_i)/2},$$

where $a_1 = mn - a$, $a_2 = a$, $I_1 = I\setminus L$, $I_2 = L$ and $\bar{x}_i$ is the mean of observations in I_i , i=1,2. This $p_{mn}(\gamma,\alpha)$ is clearly shift and scale invariant. If $p_{mn}(\gamma,\alpha)$ is fairly even , we can say there is no line in the image.

6. PRACTICAL CONSIDERATIONS AND DISCUSSIONS

Is the whole theory developed here compatible with the real situations in image processing? The serious constraint in the theory is that the observations are independent, whereas the normal assumption is an approximation and in general tolerable , since an image is generally taken by integration of several frames. We know that many CCD sensors in the market are of 300 or 400 sensing elements, not exactly 512 per row and this gives the correlated observations [7] . However, a newly genius designed sensor from GE is CID 512 (TN 2250) which has 512×512 square sensing elements [8] and we believe that observations taken from this sensor are quite nearly independent although we do not have such one (Please note, TN 2250 has interlace 1:1 and hence not compatible with most commercial image processors which need interlace 2:1). Any reader who has such device may try the test of independence. A serious problem is that many commercial sensors are unstable in sensing the light, *i.e.* there are temporal variations when the lighting looks stable [9, p.40]. This will affect the heuristic procedures but not ours.

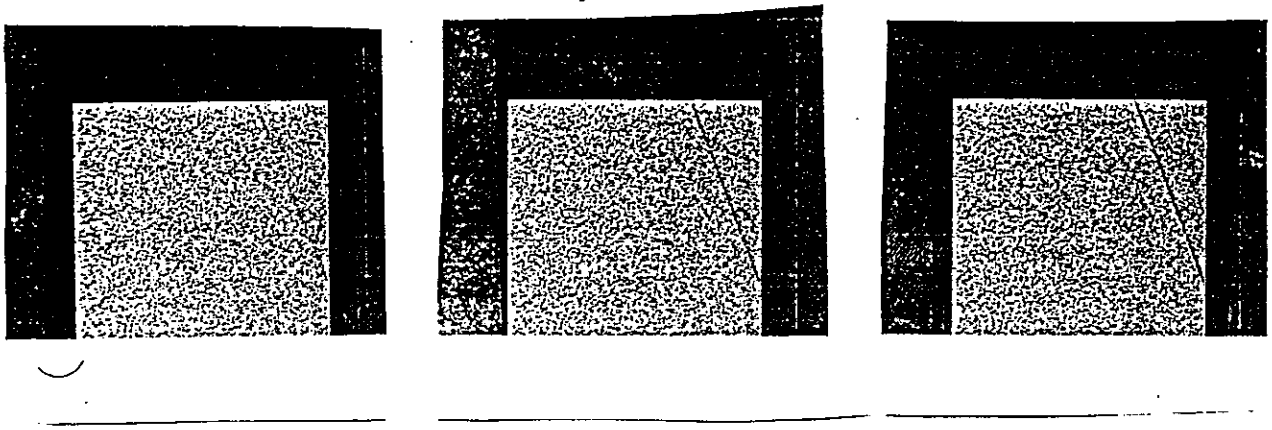
Geometric distortions, due to optics, present a problem to our methods as well as to Hough transform. As mentioned before we may use lines of two pixels wide to overcome the effects of distortions. In this case ,we loose one pixel precision. The Hough transform will have better tolerance to the geometric distortions if the algorithm of finding peaks in accumulator array is quite reliable.

The Hough transform needs multidimensional accumulator array, which is generally formidable large, whereas our methods do not. However, our methods take too much computing time and fortunately this can be implemented in parallel. Thus for a SIMD or MIMD machine with a large number of processing elements , our methods will be able to achieve real time performance and reach practical usage. Also, Hough transform has a shortcoming of inability to detect no line in an image when noises are present, whereas our methods haven't. In summary our methods are realistic and can be applied in many

practical situations, and in some cases better than Hough transform.

REFERENCES

1. D. J. Hunt, L. W. Nolte and W. H. Ruedger, " Performance of the Hough transform and its relationship to statistical signal detection theory " , Computer Vision, Graphics, and Image Processing 43, 221 – 238 , 1988.
2. P. J. Bickel and K. A. Doksum, " Mathematical Statistics " , Holden – Day Inc. , San Francisco , California, 1977.
3. J. Illingworth and J. Kittler, " A survey of the Hough transform " , Computer Vision, Graphics, and Image Processing 44, 87 – 116 , 1988.
4. R. Y. Rubinstein , " Simulation and the Monte Carlo Method " , John Wiley and Sons , New York , 1981.
5. J. O. Berger , " Statistical Decision Theory and Bayesian Analysis " , Second Edition , Springer – Verlag , New York ,1985.
6. R. J. Muirhead , " Aspects of Multivariate Statistical Theory " , Wiley , New York , 1982.
7. G. D. Boreman , " Fourier spectrum techniques for characterization of spatial noise in imaging arrays " , Optical Engineering 26 , No. 10 , 985 – 991 ,1987.
8. P. Amorese and J. Bloomfield , " A slew of standards for camera systems " , Datacube World Review , Vol. 2 , No. 4 , 2 – 6 ,1988.
9. E. P. Krotkov , " Exploratory visual sensing for determining spatial layout with an agile stereo camera system " , Technical Report MS–CIS–87–29 , Department of Computer and Information Science , University of Pennsylvania , 1987.



2(a) 2(b) 2(c)

Fig. 2. Three simulated images (200×200) , each containing a line of two pixel wide (γ=75 , θ=20). The mean difference $|\mu_2|$ is 20 in (a), 40 in (b), and 60 in (c).

```
GEN20.DAT

I1 , I2 :              0          0

 MAXIMUM W - SQUARE IS:      12.28862
             THETA  IS:      95.00000
             RADIUS IS:     -44.16806

I1 , I2 :              0          0

 MAXIMUM W - SQUARE IS:      15.13044
             THETA  IS:      158.0000
             RADIUS IS:      43.14138

I1 , I2 :              0          0

 MAXIMUM W - SQUARE IS:      52.16462
             THETA  IS:      19.00000
             RADIUS IS:      45.47770

I1 , I2 :              0          0

 MAXIMUM W - SQUARE IS:      92.01628
             THETA  IS:      20.00000
             RADIUS IS:      10.64178
FORTRAN STOP
$ SH TIME
   4-APR-1989 15:21:05
$ DEASS SYS$OUTPUT
```

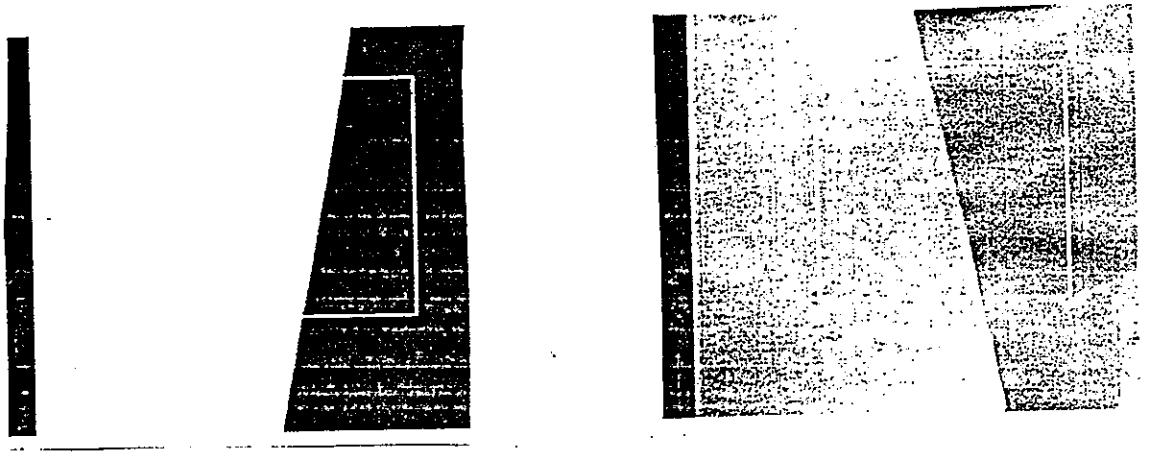
3(a)

```
[.UF]GEN40.DAT

I1 , I2 :              0          0

 MAXIMUM W - SQUARE IS:      12.28862
             THETA  IS:      95.00000
             RADIUS IS:     -44.16806

I1 , I2 :              0          0

 MAXIMUM W - SQUARE IS:      15.13044
             THETA  IS:      158.0000
             RADIUS IS:      43.14138

 MAXIMUM W - SQUARE IS:      176.0871
             THETA  IS:      21.00000
             RADIUS IS:      44.98809

 MAXIMUM W - SQUARE IS:      361.8475
             THETA  IS:      20.00000
             RADIUS IS:      10.64178
FORTRAN STOP
$ SH TIME
```

3(b)

Fig 3. The test results of Fig. 2, where 3(a) corresponds to Fig. 2(a), and 3(b) to Fig. 2(b).

4(a)



4(b)

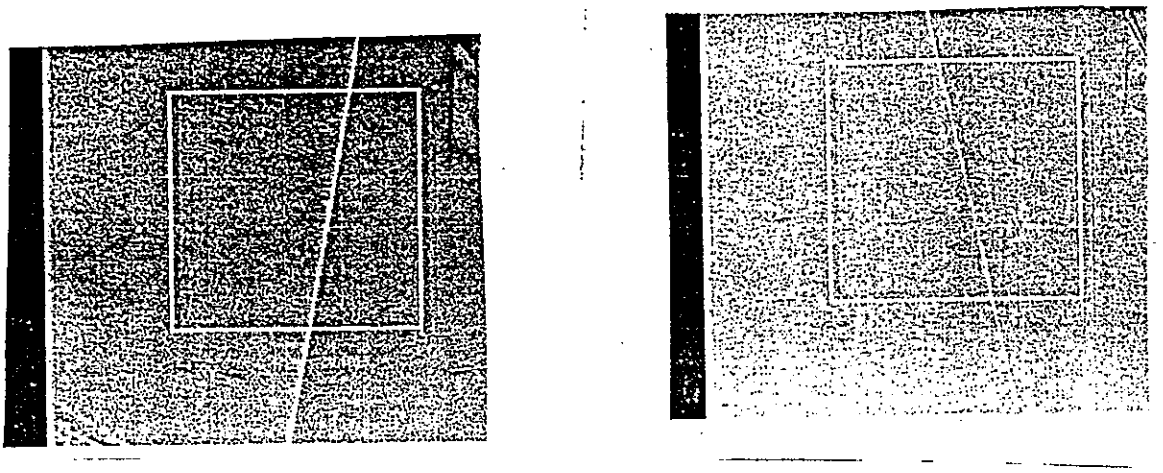


Fig 4. 4(a) Two true images of 200×200 , each containing an edge line. 4(b) Results by using Sobel operators on 4(a).

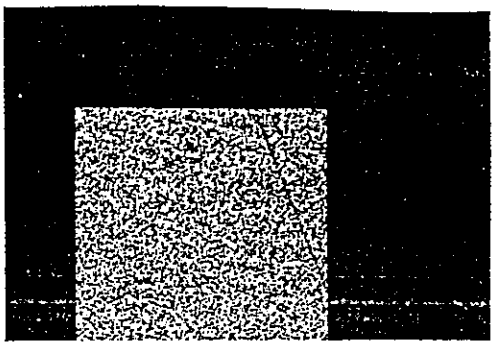
```
                    0          0

 MAXIMUM W - SQUARE IS:     26.89224
             THETA  IS:     86.00000
             RADIUS IS:    -42.10255

 I1 , I2 :          0          0

 MAXIMUM W - SQUARE IS:     17.12052
             THETA  IS:     79.00000
             RADIUS IS:    -43.80480

 MAXIMUM W - SQUARE IS:     2759.249
             THETA  IS:     170.0000
             RADIUS IS:     29.44737

 I1 , I2 :          0          0

 MAXIMUM W - SQUARE IS:     3781.766
             THETA  IS:     171.0000
             RADIUS IS:     12.14958

 FORTRAN STOP
```

5(a)

```
                    0          0

 MAXIMUM W - SQUARE IS:     15.17580
             THETA  IS:     10.00000
             RADIUS IS:     29.44737

 I1 , I2 :          0          0

 MAXIMUM W - SQUARE IS:     417.2392
             THETA  IS:     13.00000
             RADIUS IS:     51.31517

 I1 , I2 :          0          0

 MAXIMUM W - SQUARE IS:     2377.474
             THETA  IS:     11.00000
             RADIUS IS:    -27.50535

 I1 , I2 :          0          0

 MAXIMUM W - SQUARE IS:     1126.398
             THETA  IS:     11.00000
             RADIUS IS:    -46.86094

 FORTRAN STOP
```

5(b)

Fig 5. The test results of Fig 4. , where 5(a) corresponds to the left image and 5(b) to the right image.



7(a)

```
   4 APR-1989 09:53:32
$ RUN USR$DISK2:[UF]FINDB
KEY IN INPUT FILENAME
GENB.DAT

I1 , I2 :            0            0

 MAXIMUM W - SQUARE IS:     7.897822
             THETA  IS:     78.00000
             RADIUS IS:     33.73724

I1 , I2 :            0            0

 MAXIMUM W - SQUARE IS:     9.848311
             THETA  IS:     16.00000
             RADIUS IS:     37.45079

I1 , I2 :            0            0

 MAXIMUM W - SQUARE IS:     75.34846
             THETA  IS:     21.00000
             RADIUS IS:     44.98809

I1 , I2 :            0            0

 MAXIMUM W - SQUARE IS:     137.3337
             THETA  IS:     20.00000
             RADIUS IS:     10.64178
FORTRAN STOP
$ SH TIME
```

7(b)

Fig 7. 7(a) A simulated 200×200 binary image with a line ($p_0 = .3$, $p_1 = .8$,). 7(b) The test results on four subimages of 7(a).