TR-87-001

Statistical Theory of Edge Detection

Jun S. Huang and Dong H. Tseng

Computer Vision Laboratory

Institute of Information Science

Academia Sinica

Taipei, Taiwan, Rep. of China

November 1985

Revised October, 1986

Abbr. of title: Statistical Theory of Edge Detection

Mailing address: Jun S. Huang

Institute of Information Science

Academia Sinica

Taipei, 11529, Taiwan, Rep. of China

__Abstract__. Conventional way of edge detection is first filtering the image and then using simple techniques to detect edges. However, filtering the noises will also blur the edges since edges correspond to the high frequencies. Our suggestion is that both filtering and edge detection should take place at the same time. The way of doing this is by statistical theory of hypothesis testing. A simple form of decision rule is derived and the generalization of this result to more complicated situations is also discussed in detail. The decision rule can make a decision whether in a given small neighborhood there is an edge, or a line, or a point, or a corner edge, or just a smooth region. During the computation of the decision rule, the by-products are the mean and variance of the neighborhood and these can be used for split and merge analysis. The calculation of mean acts as filtering of the neighborhood pixels. In fact the representation of the neighborhood by its mean and variance can be generalized by Haralick's sloped-facet model, which has better understanding of the local changes of intensities.

List of special symbols: μ, σ^2, λ, α, β, ξ, θ .

1. INTRODUCTION.

Edge detection is critical to the recognition of an object. There are three main reasons. The first: the human eyes glance an unknown object by beginning tracing its outlines, which are composed of edge segments. The second: through experiences if the boundary of an object can be traced out successfully then shape analysis can be simplified and the recognition becomes highly feasible. The boundary is composed of edge segments and thus edge detection plays dominant role. The third: many images do not contain concrete objects and the understanding of these images depends on their texture properties. However the extraction of texture properties is also highly related to the edge detection.

Ther are several difficulties in extracting exact edges. The most serious one is that noises corrupt the edges and this makes the process of edge detection complicated.

The classical gradient or mask methods (see Rosenfeld and Kak [1], also Nevatia and Babu [2]) have suffered a serious drawback that the threshold must be properly selected and is often dynamically changed. Although Frei and Chen [3] use geometric property to avoid tuning threshold, their method however is noise sensitive. Relaxation method [4] has a serious problem of when to stop and is not suitable for real-time processing. In practices, an image input device will contain noises and the conventional methods of reducing noises such as filtering will also blur the edges since deges correspond to the high frequencies as noises do. Thus the consideration of the variations due to noises and edges altogether is critical. Although Yakimovsky [5] suggests using a maximum likelihood principle for detecting step edges, theoretically he does not solve the problem in statistical sense. Thus he pursues a heuristic method of region growing to solve the problem. Haralick [6] uses a sloped facet model for edge analysis. Essentially he is doing three parameter linear regressional analysis. Unfortunately the edge detection for this model corresponds to the regressional change-point problem which is not well solved in statistical literatures. If we analyze different cases of edge structures, the problem is actually in the area of multiple decision theory; a simple F test can not work well. This is why Haralick has to use a non-maximum suppression operator to pick up a local maximum in the F-statitic. Nahi and Jahanshahi [7], Nahi and Assefi [8], Habibi [9], and Nahi [10] all use a statistical model of the object, background, and noise in a Bayesian and/or recursive estimation scheme to improve the image or estimate the boundary. Their approaches are different from statistical change-point theory which is emphasized in this paper. An optimal edge detector based on Laplacian of Gaussian is discussed by Lunscher and Beddoes [11, 12]. Although Laplacian of Gaussian is supported by theoretical evidences from Shanmugam et al. [13], Torre and Poggio [14] and also biological evidences from Marr [15], there are still some problems on tuning parameters as shown in Fig. 6 (C) of [12]. Also it is questionable whether human beings are able to make an image sensor similar to biological ones. We make airplane completely different from birds: we need a powerful engine by physical law not by biological law.

The new theory proposed here is by using statistical analysis of change-point problem [16, 17]. A simple form of decision rule is derived and the generalization of this rule to more complicated situations is also discussed in detail. Haralick's sloped-facet model [6] is analyzed and gives a better interpretation of local intensity changes. Also some experimental results are given here and comparison with Sobel operator indicates the superiority of the new decision rule.

2. THE THEORY OF EDGE DETECTION

Consider a small square subimage of properly chosen size, say $n \times n$, n depending on the resolution of image. We would like to choose n such that any square subimage of this size has at most one edge segment passing through. Furthermore, n is not too small, for the $n \times n$ subimage would be noise sensitive. The pattern of this subimage may be classified as the following four types:

(a) uniform type: uniform region,

(b) edge type: two disjoint connected uniform subregions containing at least three pixels, so that the boundary of these subregions corresponds to the edge segment,

(c) line type: line passing through uniform region; and assuming the line is at least two pixels wide and each splitted region contains at least four pixels,

(d) point type: a point in a uniform region and assuming the point contains at least four pixels.

Examples for each of these types are shown in Fig. 1. We assume the resolution is high enough so that a point or a line to be clearly shown up should contain sufficient number of pixels. In general the image sensor is not accurate enough to detect a pixel correctly, rather it detects a group of pixels, with some properties.

12	11	10	12	10
9	10	9	11	10
11	9	11	9	12
10	11	10	12	11
11	11	10	11	10

(a) uniform region

22	101	100	99	102
21	100	100	101	20
23	99	102	100	21
20	102	100	20	23
21	21	20	21	20

(b) two disjoint connected uniform subregions

1	2	91	93	0
1	0	94	92	1
0	91	93	1	2
2	92	90	2	0
1	94	91	0	1

(c) line

0	1	1	2	0
1	2	81	82	1
1	83	80	85	1
2	0	1	0	2
0	1	2	1	1

(d) point

Fig. 1. Examples of four types of pattern of a 5×5 subimage

In Fig. 1, (b), (c), (d) all clearly partition the 5×5 subimage into two disjoint sets (or two uniform regions) and the two sets are connected in (b) and (d), and in (c) one set is splitted into two by another set which corresponds to a line.

If any $n \times n$ subimage contains edge segment, or a line, or a point, then it should be partitioned into two uniform regions (a region here may not be connected) having significant differences in gray levels (as shown in Fig. 1.). If the contrast is low then the difference may not be clear. In practices, the input image sensor and electronic devices have noises (assuming lighting is uniform and focus is precise), and the noises corrupt the edges. Thus sometimes the edges are not clear because of noises. Hence the pixel gray levels X_i's are assumed to be random, $i = 1, 2, \ldots, N (= n^2)$. In general the characteristics of noises are not globally uniform (i.e. not stationary) but are locally dependent. In a small subimage of $n \times n$, we only assume that noises are independent normally distributed because the device noises are in some cases white and Gaussian (normal), and also additive. Whether this is true depends on the manufactures of the camera. Thus $X_i \sim N(\mu_i, \sigma^2)$, (which means X_i has normal distribution with mean μ_i and variance σ^2, both are unknown), and X_i's are independent, $i = 1, 2, \ldots, N$.

Let us relabel the pixel gray levels X_i's as $\{X_i, i = 1, \ldots, m\}$ corresponding to the first partitioned region, and $\{X_i', i = 1, \ldots, m'\}$ corresponding to the second partitioned region where $m+m' = n^2 = N$. Let $X_i' \sim N(\mu_i', \sigma^2)$. Now consider the following hypothesis testing:

$$H_0: \mu_1 = \dots = \mu_m = \mu_1' = \dots = \mu_m' = \mu_0, \text{ (unknown)}$$

against

$$H_a: \mu_1 = \dots = \mu_m \neq \mu_1' = \dots = \mu_m'$$

for some unknown partition of the subimage. Please note that we consider small values of n because we want either H_0 is true or H_a is true. Of course there is possibility that H_0 and H_a both are not true, but the possibility is very small and hence negligible. We would like to find a test statistic based on X_i, X_j', $i = 1, \dots, m$, $j = 1, \dots, m'$, to test whether H_0 is true or H_a is true. If H_a is true we must find the partition that makes H_a true (i.e. find the edge segment or line, or point). The test statistic is derived from the likelihood ratio test.

The likelihood function of $\{X_i\}_1^N$ under H_0 (i.e. no partition or no edge exists) is

$$L(\mu_0, \sigma_0^2 | H_0) = \prod_{i=1}^{N} (2\pi\sigma_0^2)^{-1/2} \exp[-\frac{1}{2\sigma^2}(X_i - \mu_0)^2] .$$

The maximum likelihood estimate of μ_0, σ_0^2 are

$$\hat{\mu}_0 = \bar{X}_N = \sum_1^N X_i/N ,$$

$$\hat{\sigma}_0^2 = \sum_1^N (X_i - \hat{\mu}_0)^2/N .$$

The likelihood function of $\{X_i\}_1^N$ under H_a is

$$L(\mu_1, \mu_1', \sigma^2, P|H_a)$$

$$= \prod_1^m \{(2\pi\sigma^2)^{-1/2}\exp[-\frac{1}{2\sigma^2}(X_i-\mu_1)^2]\}$$

$$\times \prod_1^{m'} \{(2\pi\sigma^2)^{-1/2}\exp[-\frac{1}{2\sigma^2}(X_i'-\mu_1')^2],$$

where P is the unknown partition; X_i, μ_1, m corresponds to the first partitioned region and X_i', μ_1', m' corresponds to the second partitioned region. The maximum likelihood estimate of μ_1, μ_1', σ^2 and P are

$$\hat{\mu}_1 = \bar{X},\ \bar{X} = \sum_1^m X_i/m ,$$

$$\hat{\mu}_1' = \bar{X}',\ \bar{X}' = \sum_1^{m'} X_i'/m',$$

$$\hat{\sigma}^2 = \sum_1^m (X_i - \bar{X})^2 + \sum_1^{m'} (X_i' - \bar{X}')^2]/N ,$$

and $\hat{P}$ is the partition that maximizes

$L(\hat{\mu}_1, \hat{\mu}_1', \hat{\sigma}^2, P|H_a)$ over all meaningful P (like those shown in Fig. 1).

The likelihood ratio test is: reject H_0 (accept H_a) if

$$\lambda = \frac{L(\hat{\mu}_1, \hat{\mu}_1', \hat{\sigma}^2, \hat{P}|H_a)}{L(\hat{\mu}_0, \hat{\sigma}_0{}^2|H_0)} \text{ is large.}$$

By simplifying λ further, we obtain an equivalent test statistic, or a decision rule:

$$W^2 = \max_P \frac{mm'}{N} \frac{|\bar{X} - \bar{X}'|^2}{S_p^2} ,$$

and reject H_0 if $W^2 > \xi_\alpha$ where ξ_α depends on the distribution of W^2 and probability of type 1 error;

$$S_p^2 = [\sum_1^m (X_i - \bar{X})^2 + \sum_1^{m'} (X_i' - \bar{X}')^2]/(N - 2),$$

an unbiased estimate of σ^2 under H_a. Please note that variance of $\bar{X} - \bar{X}'$ is $\sigma^2(\frac{1}{m} + \frac{1}{m'}) = \sigma^2(\frac{m + m'}{mm'}) = \sigma^2(\frac{N}{mm'})$ and thus $\sqrt{\frac{mm'}{N}}(\bar{X} - \bar{X}')/S_p$ has t-distribution with $N - 2$ degree of freedom. Thus W^2 is the maximum of t^2-distribution over all meaningful partitions P.

The null distribution of W^2 (under H_0) is almost impossible to be derived analytically since W^2 is the maximum of t^2-distributions over all meaningful partitions. Hence we try the other way: computer simulation. Simulation in some sense extracts the knowledge of image signal and we incorporate this knowledge into the decision rule. The simulation result is based on 1000 sample points of W^2 and it takes much effort to find all meaningful partitions that partition the $n \times n$ subimage into two separated regions, each region having at least three points for noise and practical considerations. The VAX-11/780 random number generator is used to generate uniform random numbers and the Kolmogrov-Smirnov test is used to screen out nonuniform random numbers. Transformation to normal data is by log-cosine method (see [18]). The simulated critical values are, for 5×5 window,

$$\alpha = .01 \qquad \xi_\alpha = 38.57$$

$$\alpha = .05 \qquad \xi_\alpha = 24.19$$

$$\alpha = .10 \qquad \xi_\alpha = 22.13$$

and the null distribution of W^2 has approximate shape shown in Fig. 2.

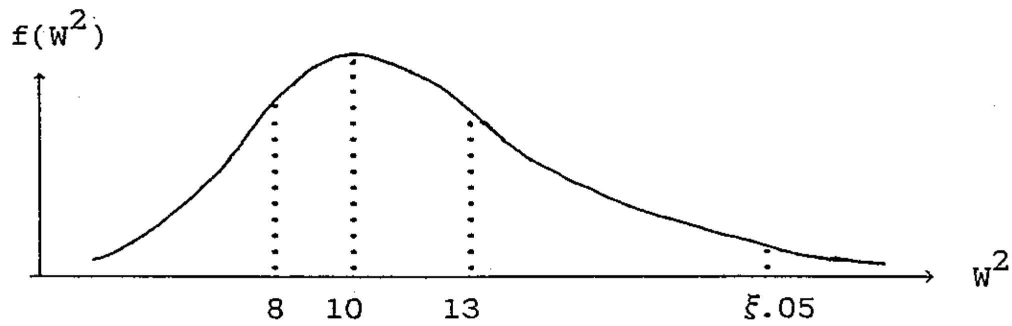


Fig. 2. Density of W^2, n = 5 .

The statistic W^2 has some intuitive meanings: If the edge contrast is low (i.e. $|\bar{X} - \bar{X}'|^2$ is small) but the noises are also small (i.e. S_p^2 is small), W^2 may be large indicating the existence of edges since W^2 is proportional to $|\bar{X} - \bar{X}'|^2/S_p^2$. On the other hand, if the edge contrast is high but the noises are also large then W^2 may be small indicating no edge exists. If $W^2 > \xi_\alpha$ for a prespecified α, then H_0 is rejected and H_a is accepted. Then the edge segment is corresponding to the partition that maximizes W^2 .

Now for a real 5 × 5 subimage, if we choose $\alpha = .05$ and the computed W^2 is greater than $\xi_\alpha = 24.19$ then we reject H_0 and accept H_a: edge, or line, or point pattern exists. The estimate $\hat{P}$ of the partition P when H_a is true is the P that maximizes W^2 . Thus from $\hat{P}$ we can see whether an edge (may be a corner edge), or a line, or a point exists in this subimage.

The advantages of this statistical approach are:

(1) thresholding value is fixed by specifying α for every kind of images,

(2) theorectical justification to get nontrivial decision rule,

(3) the computed values $\hat{\mu}_1$, $\hat{\mu}_2$, $\hat{\sigma}^2$ can be used for further region analysis.

(4) line and point patterns can be detected besides edge.

3. THE NEW METHOD AND SOME EXPERIMENTAL RESULTS.

Here in Academia Sinica we have a Hamamatzu C-1000 T.V. camera as image input device and it contains significant noises. For noise analysis, we take a picture of a uniformly white paper and calculate its mean and variance by splitting it into 441 6×6 subpictures, each having calculated mean and variance. The overall standard deviation mean is 1.064 and the variance of standard deviation is 0.0295. The detail of noise analysis is in another report. We take a picture of resolution 128×128 with 128 gray levels, and choose window size $n = 5$ or $N = 25$ about .15% of resolution. Starting from the left, skip three lines rightward each time to take a new $n \times n$ subimage and after reaching the right end we then return to the left and skip three lines downward and repeat to take subimages. Thus total subimages scanned is 40 × 40 = 1600. The reason of skipping 3 lines is that the adjacent subimages should contain some overlapping area in which edge segment may occur. It takes about one hour to process a picture on VAX-11/780. The most computing time are spent in calculating all different cases of partitions and its corresponding statistics.

Some experimental results are given below. Fig. 3(a) gives a picture of F-15 fighter in cloudy sky. The edge detection result by our new method is given in Fig. 3(b), in which the outline of the ariplane is clear visible. Fig. 3(c) is the result of using Sobel operator. Please note that the threshold is tuned low so that edges are very thick. Fig. 4(a) is the edges of two Chinese characters indicating clear respresentation of boundaries, and Fig. 4(b) contains a VAX-11/780 disk driver box with many vertical bars. The new edge detection result is given in Fig. 4(c) showing clear vertical bars. We also note that in all cases there are at most 5% of false edges where the critical value is set at $\alpha = .05$.

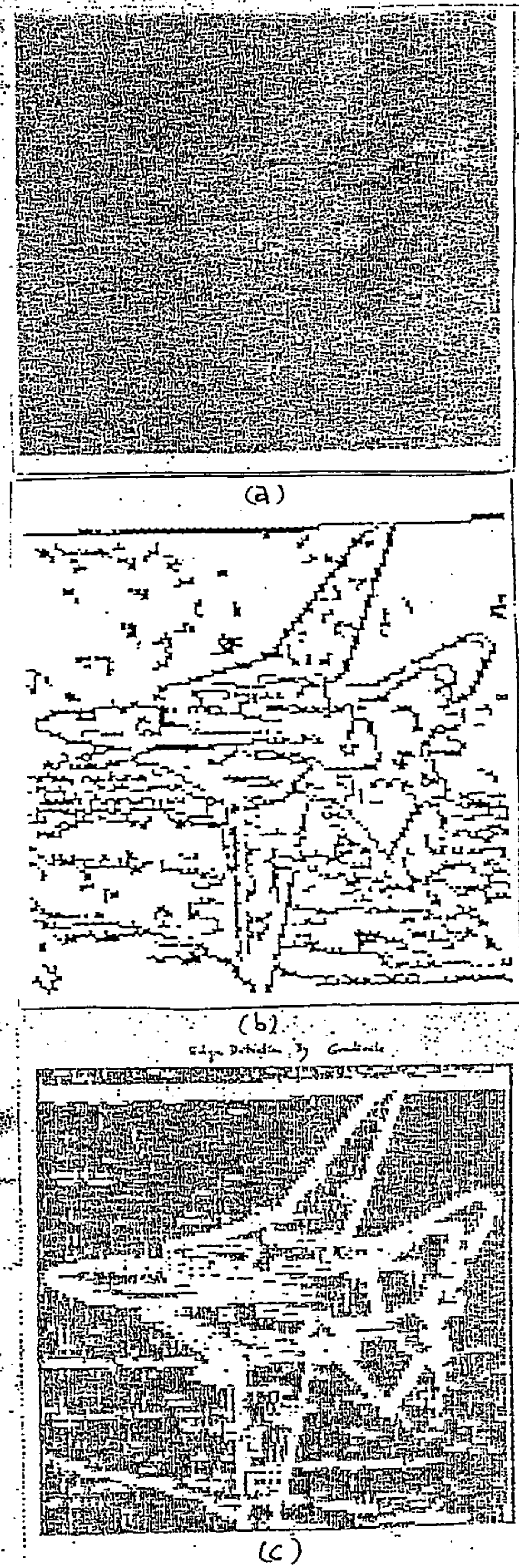
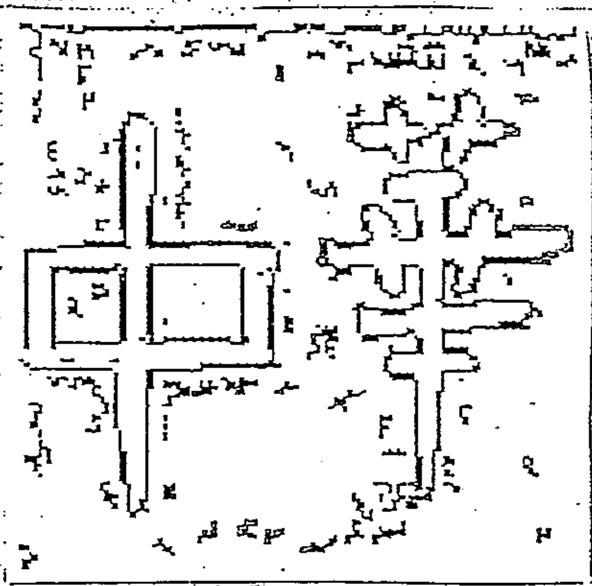
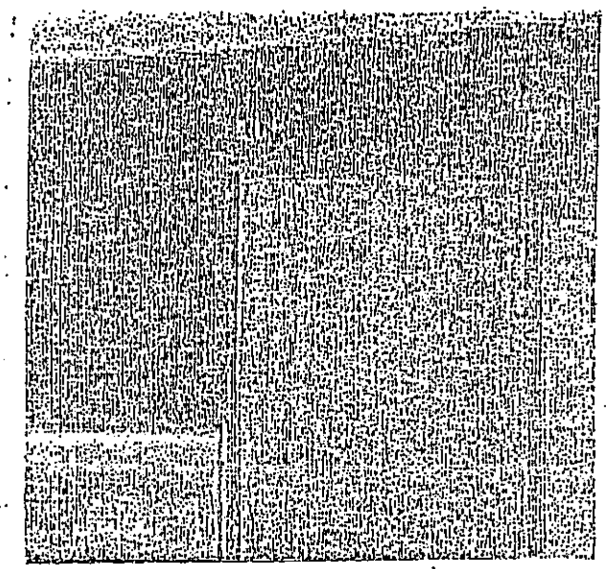


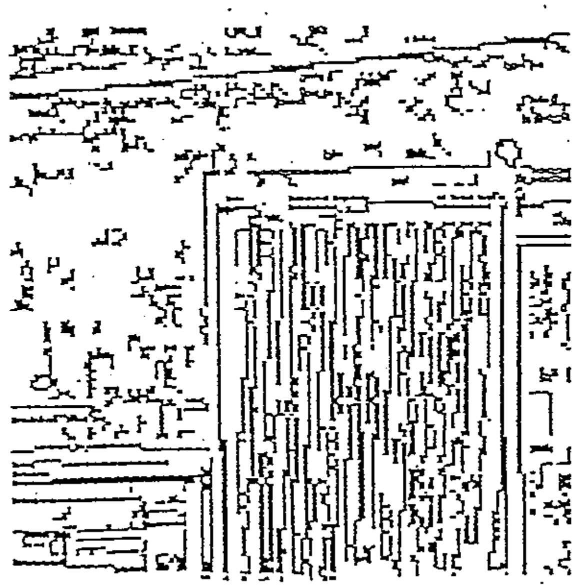
FIG. 3. (a) An F-15 jet fighter in cloudy sky. (b) The result of new edge detection. (c) The result of edge detection by gradient method with low threshold.



(a)



(b)



(c)

FIG. 4. (a) Edges of Chinese characters by new method. (b) The scene contains VAX 11/780 disk driver box. (c) Edges of (b) by new method.

4. GENERLIZATION OF W^2 STATISTICS

In formulating a statistical decision model, we assume that each pixel X_i is independent and distritruted as $N(\mu_i, \sigma^2)$. Now we are going to relax the assumptions and give some further analyses and discussions.

(A). Unequal variances: The variance σ^2 is equal to σ_1^2 for each pixel in region 1 , and equal to σ_2^2 in region 2 , where σ_1^2 and σ_2^2 may be equal. The derived statistics is

$$W_1^2 = \max_P \frac{|\bar{X} - \bar{X}'|^2}{(\frac{S_1^2}{m} + \frac{S_2^2}{N-m})} ,$$

where S_1^2 and S_2^2 are sample variances in region 1 and 2 respectively. Unfortunately, the distribution of W_1^2 depends on the ratio σ_1^2/σ_2^2 which is generally continuous, not discrete. But in most real cases $\frac{1}{3} \leq \sigma_1^2/\sigma_2^2 \leq 3$ and by symmetry we can only quantize the ratio in the interval $[\frac{1}{3}, 1]$ into some k ; say 10, levels that will give good approximation. Thus from this quantization we should be able to build a table of critical values of W_1^2 that depends on both N and σ_1^2/σ_2^2 .

(B). Haralick's sloped-facet model: In an $n \times n$

subimage let us move the origin of coordinate system to the center point or the one nearest to it. This arrangement can simplify the calculation of inverse of design matrix to be mentioned later. Let each pixel at (i, j) has gray level X_{ij} and Haralick's sloped-facet model is $X_{ij} = a + bi + cj + \epsilon_{ij}$ where ϵ_{ij} is random variation and $\epsilon_{ij} \sim N(0, \sigma^2)$ and all ϵ_{ij}'s are independent. Putting Y = column vector of X_{ij}'s and A = design matrix containing rows of $1, i, j$ at position corresponding to X_{ij} in Y, then the estimates of a, b, c and σ^2 are

$$\hat{\theta} = \begin{bmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{bmatrix} = (A'A)^{-1}A'Y \quad \text{and} \quad \hat{\sigma}^2 = (Y'Y - \hat{\theta}'A'Y)/(m-3) ,$$

where the prime sign denotes the transpose of a matrix and m is the dimension of Y or the number of X_{ij}'s .

Now the subimage is partitioned into two regions, and we use Haralick's sloped-facet model to get

$$\hat{\theta}_1, A_1, \hat{\sigma}_1^2 \text{ in region } 1 \text{ , and}$$

$$\hat{\theta}_2, A_2, \hat{\sigma}_2^2 \text{ in region } 2 \text{ .}$$

Then the derived decision rule is

$$T^2 = \max_p(\hat{\theta}_1 - \hat{\theta}_2)'[\hat{\sigma}_1{}^2(A_1'A_1)^{-1} + \hat{\sigma}_2{}^2(A_2'A_2)^{-1}]^{-1}(\hat{\theta}_1 - \hat{\theta}_2).$$

Please note that we just test whether $\theta_1 = \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}$ is equal to $\theta_2 = \begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix}$ and this is a further generalization of (A) where $b_2 = c_2 = b_1 = c_1 = 0$. Clearly during the evaluation of T^2 more informations about edges lines, points and regions are computed.

The distribution of T^2 is too complicated and it seems that the T^2 just depends on N and ratio $\sigma_1{}^2/\sigma_2{}^2$. The only one way to find out the detail is by simulation and this will take too much time.

(C). Nonparametric approximation: Now we drop the assumption of normality and assume the distributional form be unknown. Nonparametric edge detection has been studied by Bovik, Huang and Munson, Jr [19] without using techniques of change-point analysis. We can discuss this in two case.

(1) Shift in location: To test H_0: each pixel's distribution F is equal to some unknown function F_0, against H_a: $F = F_1(x)$ for pixels in region 1 and $F = F_1(x$

$-\theta)$, $\theta \neq 0$ and θ unknown, for pixels in region 2, F_1 unknown. The test statistics is

$$d_1 = \max_p \{|U_{K,N-K} - \frac{K(N-K)}{2}|/[K(N-K)(N+1)/12]^{1/2}\}$$

where K is the number of pixels in region 1 and $U_{K,N-K} = \sum_{(i,j)_2} \sum_{(i,j)_1} \psi(X_{(i,j)_1} - X_{(i,j)_2})$, the Mann-Whitney-Wilcoxon statistic, where $\psi(a) = 1, 0$ as $a \geq, < 0$, and $(i, j)_1$, is a point in region 1 and $(i, j)_2$ in region 2.

(2) Change in distribution: It is easy to see the test statistic is the maximum of two sample Kolmogorov-Smirnov statistic over all meaningful partitions (see [20]). However, this statistic may be quite conservative.

(D). Dependent observations: In a small $n \times n$ subimage it is likely that pixel gray levels X_{ij}'s are dependent. This makes the problem very complicated. Some people use spatial time series model to analyze the data. We suggest that (1) buy a high quality camera that the noises are less dependent on different lighting conditions, and (2) do extensive experiments and analyses of the noises by using the high pass filter and the low pass filter and also proper objects (e.g. a fairly uniform paper). If we can model the noises by some multivariate normal distribution with special form of covariance matrix, then we can incorporate this matrix into Haralick's sloped-facet model to derive a better decision rule.

5. CONCLUSION.

We have developed a statistical theory of edge detection based on likelihood ratio test. The new method based on this theory has been implemented to run on VAX-780 and some experimental results are given here indicating the high feasibility of the method. The generalization to more general cases is also discussed and it seems that the Haralick's sloped-facet model is the most suitable one. The decision rules shown in this paper are computationally expensive, but besides detecting edges they also give line and point detection, and also give estimates of region mean and variance for further analysis.

REFERENCES

1. A. Rosenfeld and A. C. Kak, Digital Picture Processing, Vol. 1. II, Academic Press, New York, 1982.

2. R. Nevatia and K. R. Babu, Linear feature extraction and description, computr Graphics and Image Processing, 13, 257-269, 1980.

3. W. Frei and C. C. Chen, Fast boundary detection: a generalization and a new algorithm, IEEE Trans. on Computer, Vol. C-26, 1977, 988-997.

4. J. Prager, Extracting and labeling boundary segments in natural scenes, IEEE Trans. on Pattern Analy. Mach. Intell., PAMI-2, 1980, 16-27.

5. Y. Yakimovsky, Boundary and object detection in real world images, JACM, Vol. 23, 599-618, 1976.

6. R. M. Haralick, Edge and region analysis for digital image data, Computer Graphic and Image Processing, 12, 60-73, 1980.

7. N. E. Nahi and T. Assefi, Bayesian recursive image estimation, IEEE Trans. Computers, C-21, July 1972.

8. A. Habibi, Two-dimensional Bayesian estimate of images, Proc. IEEE 60, July 1972.

9. N. E. Nahi, Role of recursive estimation in statistical image enhancement, Proc. IEEE 60, July 1972

10. N. E. Nahi and M. H. Jahanshahi, Image boundary estimation, IEEE Trans. Computers C-26, Augest 1977, 772-781.

11. W. H. Lunscher and M. P. Beddoes, Optimal edge detector design I: Parameter selection and noise effects, IEEE Trans. on Pattern Anal. Machine Intelligence, PAMI-8, 1986, 164-177.

12. ————————————, Optimal edge detector design II: coefficient quantization, IEEE Trans. on Pattern Anal. Machine Intelligence, PAMI-8, 1986, 178-187.

13. K. S. Shanmugan, F. Dickey, and J. Green, An optimal frequency domain filter for edge detection in digital images, IEEE Trans. Pattern Anal. Machine Intelligence PAMI-1, January 1979, 39-47.

14. V. Torre and T. A. Poggio, On edge detection, IEEE Trans. on Pattern Anal. Machine Intelligence, PAMI-8, 1986, 147-163.

15. David Marr, Vision, Springer-Verlag, N. Y., 1982.

16. K. J. Worley, On the likelihood ratio test for a shift in location of normal populations, J. Amer. Statist. Asso., Vol. 74, 1979, 365-367.

17. D. A. Wolfe and E. Schechtman, Nonparametric statistical procedures for the change-point problem, J. Statist. Planning and Inference, 1984, 388-396.

18. R. Y. Rubinstein, Simulation and the Monte Carlo Method, John wiley Co., New York, 1981.

19. A. C. Bovik, T. S. Huang and D. C. Munson, Jr, Nonparametric tests for edge detection in noise, Pattern Recognition, Vol. 19, 1986, 209-219.

20. W. J. Conover, Practical Nonparametric Statistics, Academic Press, N. Y., 1977.