BOUNDARY DETECTION IN IMAGES, USING A NONSTATIONARY STATISTICAL MODEL

Robert Haralick and Amrendra Singh
Remote Sensing Laboratory
Center for Research, Inc.
University of Kansas
Lawrence, KS 66045

Abstract

In this paper we suggest two boundary detection techniques for digital images. The techniques are based on decision theoretic and statistical hypothesis testing models using localized characteristics that change from pixel to pixel. The procedures were tested on a section of a LANDSAT image. The paper illustrates detected boundaries for some values of the parameters of the method. The procedures seem to work well and the boundaries detected are thin and consistent.

1.0 Introduction

Boundary detection is an important step in the segmentation of homogeneous regions of an image. Recent literature contains a large amount of work being done in this field. The reader will find extensive references in surveys by Davis[1] and Riseman and Arbib[2].

The problem of edge determination falls into the two broad but partially overlapping classes, each of which contain a variety of techniques. The first class uses purely local information and addresses itself to the initial identification or detection of edge cells. Most methods here create a partial boundary image using thresholding procedures which classify each pixel as a boundary or object pixel based on a homogeneity or gradient criteria.[3,4,5]

However, methods which just identify edge cells using only local criteria are not adequate for detecting complete boundaries. This is because real world images are noisy and the boundaries detected are usually incomplete or erroneous. In order to use these partial boundary images, some sort of enhancement and boundary completion using world model or prior information must be done. Such techniques constitute the second class of techniques in boundary determination. Enhancement includes things like line completion by joining segments or the reduction of noise points. For example, Vanderburg[9] uses template matching to iteratively enhance linear features. Zucker et. al.[6] use a relaxation labeling on probability values associated with pixels to detect lines. Kelly[7] uses a planning method on a compressed image to guide detection in the original picture. Montanari's[8] method uses dynamic programming to detect systems of curves.

The boundary analysis part of the VISIONS[10] incorporates both these classes. The first section contains programs to detect edges while the second improves upon the edges determined.

The two methods discussed in this paper fall into the first class. The aim is to initially classify each pixel of the image as boundary or non-boundary. Other procedures can then be used to improve upon the boundary cells found.

One of the interesting statistical thresholding techniques is the dynamic thresholding method of Chow and Kaneko[11]. The basic assumption they use is that the probability distribution of gray tones for any small region of the picture consisting solely of the object or background is unimodal. Thus, in order to detect boundaries, the image is sectioned into overlapping windows and the distribution in the windows is checked for bimodality. The frequency distributions were assumed to be normal, as it is possible analytically to identify two mixtures of normal distributions.[12] Thus the distribution for the r-th window is represented by

$$f_r(u) = \sum_{k=1}^{2} \frac{\rho_{kr}}{\sigma_{kr}\sqrt{2\pi}^{1/2}} \exp[-(X-\mu_{kr})^2/\sigma_{kr}^2],$$

where ρ_{1r} and ρ_{2r} are the theoretical fractions of areas occupied by the object and the background, and $(\rho_{1r} + \rho_{2r} = 1)$. μ_{1r}, μ_{2r}, σ_{1r}, and $\sigma_{2}r$ are the mean and standard deviation of the object and background respectively. The mean and standard deviation of the mixture are given by

$$\mu_r = \rho_{1r}\mu_{1r} + \rho_{2r}\mu_{2r} \quad \text{and}$$

$$\sigma_r^2 = \rho_{1r}\sigma_{1r}^2 + \rho_{2r}\sigma_{2r}^2 + \rho_{1r}\rho_{2r}(\mu_{1r} - \mu_{2r})^2.$$

Since windows which contain boundaries between object and background are likely to have bimodal histograms with large variances, only those windows with these properties are selected. The five parameters of the above distribution are estimated using a least square fit to the histograms. The bimodality for each of these histograms is tested by the difference in means $\mu = \mu_{2r} - \mu_{1r}$, ratio of variances $\varepsilon = \sigma_{1r}/\sigma_{2r}$ and the valley to peak ratio

$$\delta = \frac{\text{Min. of } f(\mu_r) \text{ in } [\mu_{1r}, \mu_{2r}]}{\text{Min. of } [f(\mu_{1r}), f(\mu_{2r})]}$$

where $f(\mu_r)$ is the obtained distribution for the r-th window and $f(\mu_{1r})$, $f(\mu_{2r})$ are the approximated distributions. A histogram is labelled as bimodal if the values for μ, ε, and δ, lie within certain ranges.

Based on this a threshold, t_r, for each of these windows was computed such that the probability of misclassification was minimized. The thresholds for the remaining windows were computed by interpolating the weighted averages for thresholds of neighboring windows. Using these thresholds for each window, the image was segmented into boundary and non-boundary areas. Because it is insensitive to shading, the

process worked very well in determining the boundary of the heart in a cineangiogram image. However, the procedure is not readily extendable to scenes with many objects each with different mean gray tones as the thresholding would tend to smooth objects into each other, and would make it difficult to determine the mixture for any one window.

The two techniques discussed in this paper also perform dynamic thresholding by using a non-stationary window for each pixel. Where they differ from the Chow and Kaneko approach is that like some other methods[13,14,15], they use spatial differentiation to create a gradient image.

The purpose of a gradient operation is to enhance the edges or boundaries in an image. The gradient at a resolution cell is a value indicating the change in gray tones around that pixel. A high value for a pixel of the gradient image means sharp changes in the gray tones in the neighborhood of that image point, thus indicating a boundary cell. Low or zero gradient pixel values indicate little or no change in the spatial neighborhood of the cell, thereby indicating a non-boundary or homogeneous cell. Thus the homogeneous areas in the original image tend to show up as areas of low pixel values in the gradient image.

There are quite a few gradient operations available for digital images[16,17]. Some of these approximate the two dimensional gradient function '∇', while others, like the Roberts Cross[18] operators have no counter part in the continuous domain.

The simple way to detect the boundaries is to perform a level slicing on the gray levels of the gradient image. Those pixels which fall above the cutoff will be boundary cells, while those below the cutoff will be homogeneous area cells. Unfortunately, the simple approach does not work well because boundaries are often relative. In one area of the image, where there may be much texture, pixels may be considered to be boundaries only if the gradient of the boundary is higher than the gradient of the textured area. In another more quiet and homogeneous part of the image, pixels may be considered to be boundaries only if the gradient of the boundary is higher than the average gradient of the homogeneous areas. In absolute terms, however, a gradient which is higher than the average gradient of the homogeneous area may be much lower than the average gradient in a textured area. Hence, a threshold adequate for a relatively homogeneous area would call texture cells in a busier part of the image boundary cells, and a boundary threshold adequate for a busy textured part of the image would tend to call boundary cells in a relatively homogeneous part of the image a non-boundary cell. Thus to generate a good boundary image by thresholding, the cutoff for each pixel should be sensitive to the pixel gradient and the average gradient of the pixel's neighborhood. This leads to a non-stationary statistical approach for determining thresholds.

2.0 Statistical Boundary Detection

In the boundary detection technique described in the next section, the pixel's neighborhood is a rectangular window centered around the pixel. We use this neighborhood to determine the average spatial variation or gradient in the local rectangular area around each pixel. Boundaries are detected when pixel's gradient is significantly higher than the gradient in its neighborhood.

We use the Roberts Cross operator for the gradient. The gradient value, g, for pixel (i,j) for this operator is defined by

$$g(i,j) = |I(i,j)-I(i+1,j+1)| + |I(i,j+1)-I(i+1,j)|.$$

The image on which we illustrate our results is a 130 rows by 90 column LANDSAT image over a section of California. Figure 1 shows the MSS band 5 of this image and Figure 2 the result of the Roberts Cross operator. The lighter areas in Figure 2 correspond to edges, while the dark areas represent homogeneous regions. In Figure 3 we have the defocused image of the gradient picture of Figure 2, using a rectangular window box filter of size 5 rows by 5 columns. We will take the pixel value of the defocused gradient image to be the mean gradient value in the pixel's neighborhood.

2.1 Decision Theory Method

Each cell in the image will be classified as a boundary cell, or a homogeneous cell. We will assume that boundary cells and homogeneous cells belong to two populations with the same probability distribution, but different means and variances. Given a gradient image pixel with magnitude X, we denote the probability that the pixel is a boundary point by $P(b|X)$. Similarly for the homogeneous case the conditional probability is denoted by $P(h|X)$.

$P(b)$ and $P(h)$ are the probabilities of the pixel being a boundary cell and homogeneous cell, respectively. Since a pixel is either a boundary or a homogeneous cell, it follows that $P(b) + P(h) = 1$.

The Bayes criterion would assign a pixel to be a boundary cell if $P(b|X) \geq P(h|X)$. By definition of conditional probabilities,

$$P(b|X) = P(X|b)\ P(b)/P(X) \text{ and } P(h|X) = P(X|h)\ P(h)/P(X)$$

Thus the cell is a boundary cell if

$$P(X|b)\ P(b) \geq P(X|h)\ P(h). \qquad (1)$$

In order to use this statistical model, we need to determine the conditional probabilities $P(X|b)$ and $P(X|h)$. Histograms of the two probability functions $P(X|b)$ and $P(X|h)$ tend to show that each is unimodel. The distribution for homogeneous cells is thin and peaked at low values, while that of the boundary cells is more spread out and peaked at relatively higher values. This can be seen in Figure 4. Hence, the larger the mean gradient, the larger the variance of the gradient.

Becuase the mean and variance of the Poisson distribution are the same, these characteristics can be represented by appropriate Poisson distributions. In addition, the Poisson distribution is a discrete function defined on positive integers, which fits in nicely with the quantized gray tone values of a digital image.

Using the Poisson distribution form, we can expresss $P(X|b)$ and $P(X|h)$ as

$$P(X|b) = \frac{\mu_b^X e^{-\mu_b}}{X!}, \quad X = 0,1,2,3,\ldots$$

and

$$P(X|b) = \frac{\mu_h^X e^{-\mu_h}}{X!}, \quad X = 0,1,2,3,\ldots$$

where μ_b and μ_h represent the means of the boundary and homogeneous populations and $\mu_b > \mu_h$.

Substituting this into the Bayes rule (1), we decide a boundary cell if

$$\frac{\mu_b^X e^{-\mu_b}}{X!} P(b) \geq \frac{\mu_h^X e^{-\mu_h}}{X!} P(h)$$

$$\text{or} \quad \left(\frac{\mu_b}{\mu_h}\right)^X \geq \frac{P(h)}{P(b)} e^{(\mu_b - \mu_h)}$$

$$\text{or} \quad X \geq \frac{(\mu_b - \mu_h)}{\ln \mu_b/\mu_h} + \frac{\ln P(h)/P(b)}{\ln \mu_b/\mu_h} \qquad (2)$$

Denoting μ_b/μ_h by γ, we have a boundary cell if

$$X \geq \frac{\mu_h(\gamma - 1)}{\ln \gamma} + \frac{\ln P(h)/P(b)}{\ln \gamma} \qquad (3)$$

Now consider a rectangular window surrounding the resolution cell (i,j) having magnitude X in the gradient image. Let μ be the mean gradient in the window; μ is then the value of the resolution cell (i,j) in the defocused gradient image. If P(b) and P(h) are the probabilities of boundary and homogeneous cells, respectively for that window, then the mean gray tone μ for the window is given by

$$\mu = \mu_b P(b) + \mu_h P(h). \qquad (4)$$

Using $\gamma = \mu_b/\mu_h$ we have

$$\mu_h = \frac{\mu}{\gamma P(b) + P(h)}.$$

Substituting this in equation (3) a cell with gradient magnitude X is a boundary cell if

$$X \geq \mu \frac{(\gamma - 1)}{(\gamma P(b) + P(h)) \ln \gamma} + \frac{\ln P(h)/P(b)}{\ln \gamma}.$$

By introducing the parameters Θ and κ where

$$\Theta = \frac{\gamma - 1}{(\gamma P(b) + P(h)) \ln \gamma} \quad \text{and} \quad \kappa = \frac{\ln P(h)/P(b)}{\ln \gamma}$$

the condition reduces to the form

$$X \geq \Theta\kappa + \mu \qquad (5)$$

This form is intuitively appealing. It states that the threshold criterion is a linear function of the mean of the window and a scaling factor. The boundary non-boundary decision regions may be illustrated on a scatterogram as shown in Figure 5. The horizontal axis represents the gray tones of the defocused gradient image, while the vertical axis represents the values of the gradient image. Resolution cells having (defocused gradient, gradient) values lying on or above the line $\Theta\mu + \kappa$ will correspond to boundary points, while those below correspond to homogeneous region cells.

Figure 6 shows some examples of thresholding using this model for a few values of P(b) and γ. The corresponding values for Θ and κ are also shown. As expected, increasing the P(b) estimates increases the number of boundary cells. In general, the best values for P(b) and γ would vary a little for different images and might have to be determined by trial and error.

In our discussion we made the assumption that P(b) and P(h) remain the same over all windows. One way to generalize this assumption is to use the procedure iteratively. The boundary image from the first iteration could be used to determine the P(b) and P(h) values for each window. The thresholding could be carried out again with these varying probability values. Of course, this would be at the cost of increasing the complexity of the procedure.

2.2 Hypothesis Testing Method

The second procedure is based on a different kind of statistical model which also uses the gradient value and the average gradient value in a pixel's neighborhood. It perhaps is slightly more intuitively pleasing.

The box car filter or rectangular convolution operation is a low pass digital filtering function and the effect is to smooth out rapid changes in gray tones. This is best seen in Figure 7 which shows a cross-sectional view along one direction in a gradient image. The peaks represent the high gray tone values as we cross the edges. The convolution function spreads these higher gray tone values over a larger spatial area as shown by the dotted curve. Over the homogeneous areas both these curves will have about the same values as there is little spatial change to average out. Thus to determine edge cells, we will look for cells in the defocused image for which the corresponding cells in the gradient image have large enough values. Because of the noise and variation in the image we only want to label a pixel a boundary cell if this difference is relatively large.

To see what relatively large means we put the problem in a statistical hypothesis testing context. Let the neighborhood around a resolution cell have a mean gradient of μ and standard deviation of σ. Let X be the observed noisy gradient value. We wish to test the hypothesis: "the pixel has a gradient value coming from the population having mean at most μ and standard deviation σ." It is natural to reject the hypothesis if $\frac{X - \mu}{\sigma} \geq k$ for some constant k.

$\frac{X - \mu}{\sigma}$ is too large when the probability of rejecting the hypothesis when it is true is too large. We denote this probability by

$$P\left(\frac{X - \mu}{\sigma} \geq k \mid \text{mean} = \mu, \text{ standard deviation} = \sigma\right)$$

Fixing the probability to β we can solve the equation

$$P\left(\frac{X-\mu}{\sigma} \geq k \mid \text{mean} = \mu, \text{standard deviation} = \sigma\right) = \beta$$

for the threshold k.

The technique can be implemented as a two-pass operation. In the first pass a temporary image can be created in which each pixel value is the difference between the corresponding gradient and defocused or spatially averaged gradient. At the same time the standard deviation (σ) of the differences can also be computed. In the second pass a pixel can be set to a boundary cell if the difference value between the gradient and convolution cells exceeded the variable, $k' = \alpha.\sigma$, where α is a user supplied parameter which controls the level of thresholding. By increasing or lowering α, the user can, to some degree, increase or decrease the number of cells classified as boundary points. Figure 8 shows results of this procedure for α values of 0.1, 0.3, and 0.5.

3.0 Evaluation of Edge Detection Techniques

Pratt[19] has suggested some criteria for what a good edge detector should do and defined a performance measure based on these criteria. However, in order to check those the edge detector has to be applied to an image for which the ideal edge locations are known.

Pratt classifies three major types of errors associated with determination of an edge location:

(1) missing valid edge points,
(2) misclassification of noise as edge points and
(3) non-localized edge points.

The first two can be estimated by fixing the edge detection threshold at a level such that the probability of false detection resulting from noise alone does not exceed some value. The probability of the edge detection can then be computed by counting the known and detected edge points.

The third type of error is more complex as it incorporates different kinds of errors. For example, the edge detected may be ideal but fragmented, or it may be continuous but offset, or the edges may be correct but too thick. The figure of merit proposed for non-localized edge errors is defined by

$$R = \frac{1}{I_n} \sum_{i=1}^{I_a} \frac{1}{1 + \alpha d_i^2} ,$$

where $I_n = \text{Max}(I_i, I_a)$ and I_i and I_a are numbers of ideal and actual edge points. α is a scaling constant and d_i is the distance between the i-th actual edge point and a line of ideal edge points to which it is supposed to belong. The rating factor is normalized so that R = 1.0 for a perfectly detected edge. α allows for adjustment to penalize edges which are localized but offset from true position. The normalization ensures a deterioration of the merit figure for thick edges.

This figure of merit was tested for various edge techniques for an image which contained a single vertical edge of a fixed height and Gaussian noise with known variance. The values for R correspond well with the visual analysis of the boundary images.

Finally, we would add to Pratt's criteria one of edge simplicity. All other things being equal, edges which zig-zag about a line of resolution cells are not as good as edges which are simple lines or smooth curves. The micro busyness b of the edges in an image can be measured by the entropy of the distribution of edge directions:

b =-[P(north) log P(north) + P(north-east) log P(north-east) + P(east) log P(east) + P(south-east) log P(south-east)].

4.0 Results

As can be seen from Figures 6 and 8, the edges detected by the two procedures are fairly thin and consistent. Comparison between the two procedures is not as straight forward as might be expected. The advantage (or disadvantage, depending upon the point of view) that the first procedure has, is that there are two parameters P(b) and γ that the user can vary, as opposed to the single variable α of the second procedure. It should be noted that in this implementation the σ that we use for the thresholding in the second procedure, is computed over the entire image. This could cause problems if the boundaries and objects were not evenly distributed over the image as they are in these figures. If, for example, we had a large homogeneous category such as water, the small variations for σ over a large section of the image could deteriorate the thresholding over the more textured part of the image.

5.0 Future Research

We plan to evaluate the edge detection techniques discussed in Section 2 by the evaluation technique discussed in Section 3. We also plan to extend each of these techniques making them sensitive to directions of edges instead of just magnitude of edges. The results of these evaluations and extensions will be discussed in a future paper.

References

1. Davis, L.S., "A Survey of Edge Detection Techniques," Computer Graphics and Image Processing, 4, 248-270, (1975).

2. Riseman, E.M. and Arbib, M.A., "Computational Techniques in the Visual Segmentation of Static Scenes," Computer Graphics and Image Processing, 6, 221-276, (1977).

3. Haralick, R.M. and Dinstein, I., "A Spatial Clustering Procedure for Multi-Image Data," IEEE Transactions on Circuits and Systems, Vol. CAS-22, No. 5, (1975).

4. Rosenfeld, A., Thurston, M. and Lee, Y., "Edge and Curve Detection: Further Experiments," IEEE Transactions on Computers, Vol. C-21, No. 7, (1972).

5. Kasvand, T., "Iterative Edge Detection," Computer Graphics and Image Processing, 4, 279-286, (1975).

6. Zucker, S.W., Hummel, R.A. and Rosenfeld, A., "An Application of Relaxation Labeling to Line and Curve Enhancement," IEEE Transactions on Computers, Vol. C-26, No. 4, (1977).

7. Kelly, M., "Edge Detection by Computer Using Planning," Machine Intelligence VI, Edinburgh University Press, Edinburgh, 397-409, (1971).

8. Montanari, U., "On the Optimal Detection of Curves in Noisy Pictures," CACM, 14, 335-345, (1971).

9. Vanderburg, G.J., "Experiments in Iterative Enhancement of Linear Features," Technical Report TR-425, Computer Science, University of Maryland, (1975).

10. Prager, J.M., Hanson, A.R. and Riseman, E.M., "Extracting and Labelling Boundary Segments in Natural Scenes," COINS Technical Report 77-7, University of Massachusetts, Amherst, (1977).

11. Chow, C.K. and Kaneko, T., "Boundary Detection of Radiographic Images by a Threshold Method," Frontiers of Pattern Recognition," Academic Press, (1972).

12. Teicher, H., "Identifiability of Finite Mixtures," Annals of Mathematical Statistics, Vol. 34, (1963).

13. Bullock, B., "The Performances of Edge Operators on Images with Texture," Technical Report, Hughes Research Labs, Mailbu, (1974).

14. Marr, D., "Early Processing of Visual Information, Al Memo 340, Artificial Intelligence Laboratory," MIT, Cambridge, (1975).

15. Hueckel, M., "A Local Visual Operator which Recognizes Edges and Lines," JACM, 20, 636-647, (1973).

16. Hueckel, M., "An Operator which Locates Edges in Digitized Pictures," JACM, 20, 113-125, (1971).

17. Kirsch, R.A., "Computer Determiniation of the Constituent Structure of Biological Images," Computers and Biomedical Research, 4, 315-328, (1971).

18. Duda, R.O. and Hart, P.F., "Pattern Classification and Scene Analysis," John Wiley & Sons, New York, (1973).

19. Pratt, W.K., "Figure of Merit for Edge Detection," USCIPI Report 60, University of Southern California, 85-93, (1976).

Figure 1.
Test Image MSS band 5.

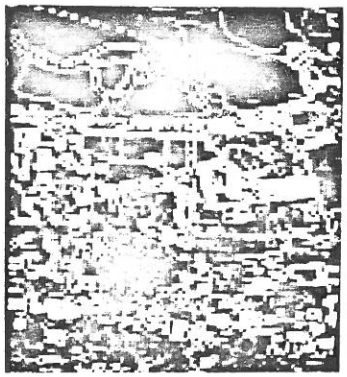


Figure 2.
Gradient of Image shown in Figure 1.

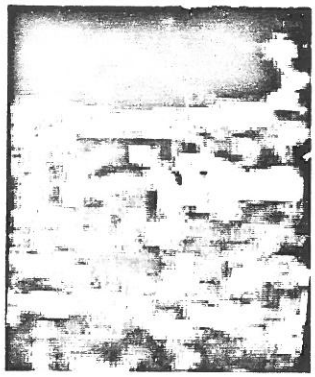


Figure 3.
5 x 5 Window Rectangular Convolution of the Gradient Image.

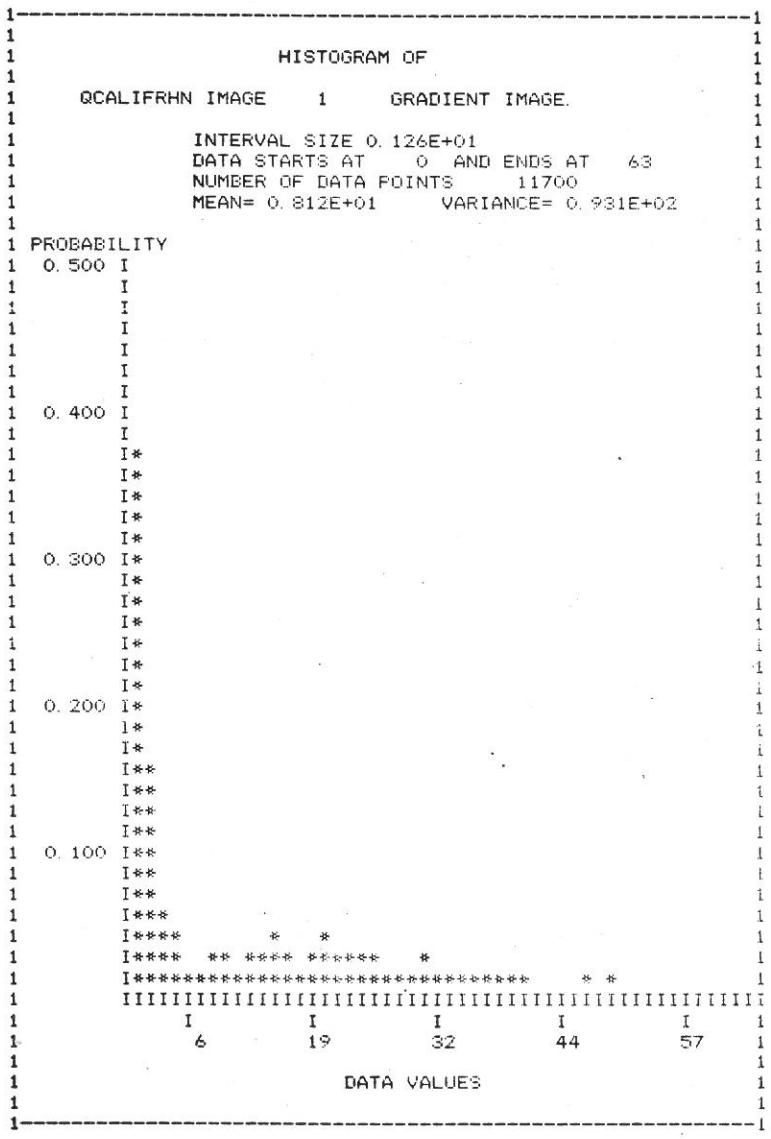


Figure 4.
Histogram of the Gradient Image. The peak at low **values** corresponds to homogeneous regions. The **boundary** cells correspond to the spread out peak **at** the higher values.

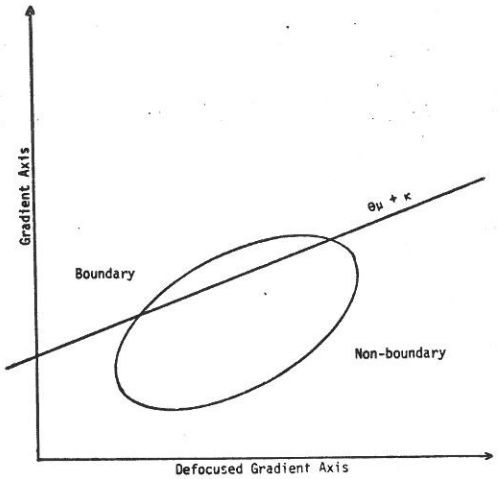
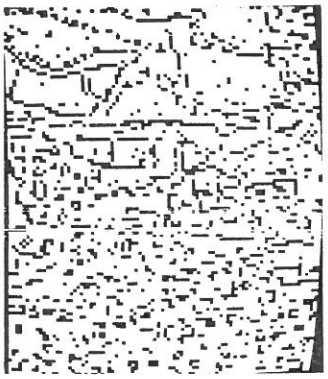
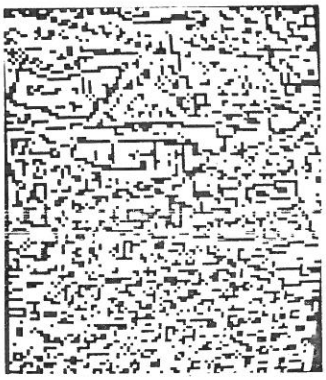


Figure 5
Schematic Scatterogram of the Gradient and the Defocused Gradient Image, showing how boundary pixels can be discriminated from non-boundary pixels.

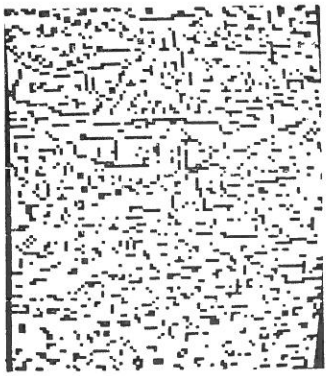
P(b) = 0.15, γ = 2.50
Θ = 1.34, κ = 1.89



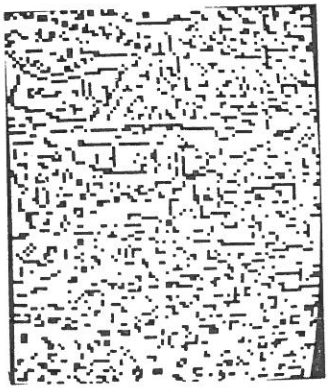
P(b) = 0.20, γ = 2.50
Θ = 1.23, κ = 1.51



P(b) = 0.30, γ = 2.50
Θ = 1.30, κ = 0.92



P(b) = 0.20, γ = 10.00
Θ = 1.40, κ = 0.60



P(b) = 0.20, γ = 15.00
Θ = 1.36, κ = 0.51

Figure 6.
Examples of Boundary Detection using the Decision Theoretic model. The black area represents boundaries and the white non-boundary.

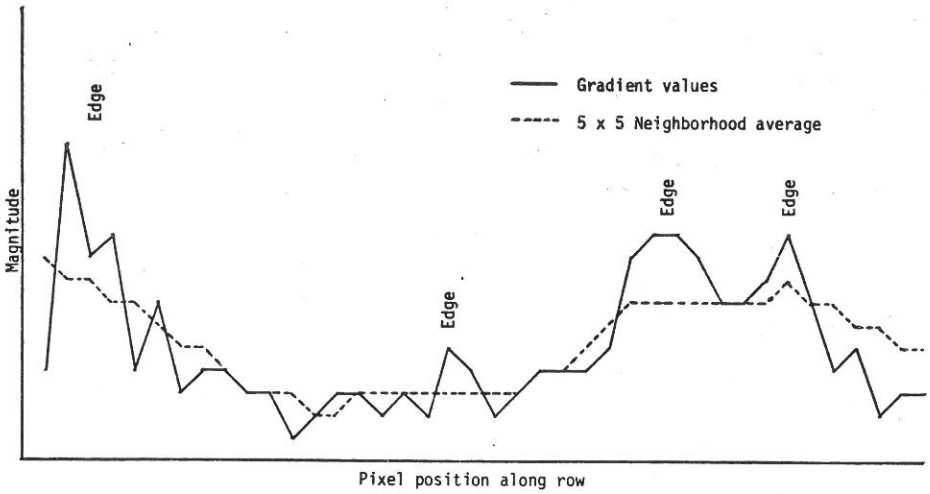
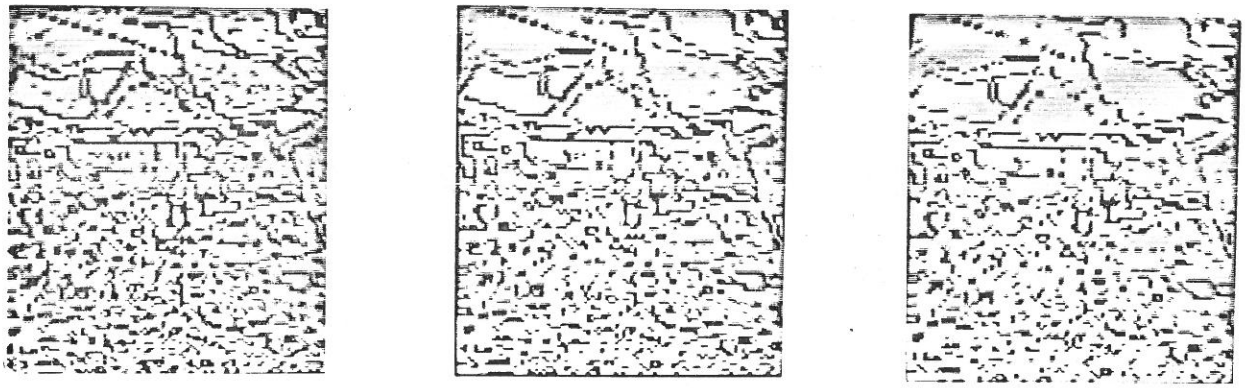


Figure 7
Cross-section view of the Gradient values along an image row. The dotted line represents the 5 x 5 neighborhood average for the corresponding pixels.



$\alpha = 0.1$ $\alpha = 0.3$ $\alpha = 0.5$

Figure 8.
Examples of Boundary Detection using the Statistical Hypothesis Testing model.