7.3 THE DIGITAL STEP EDGE

Robert M. Haralick

Departments of Electrical Engineering and Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

Abstract

We use the facet model to accomplish step edge detection. The essence of the facet model is that any analysis made on the basis of the pixel values in some neighborhood has its final authoritative interpretation relative to the underlying grey tone intensity surface of which the neighborhood pixel values are observed noisy samples.

Pixels which are part of regions have simple grey tone intensity surfaces over their areas. Pixels which have an edge in them have complex grey tone intensity surfaces over their areas. Specifically, an edge moves through a pixel if and only if there is some point in the pixel's area having a zero crossing of the second directional derivative taken in the direction of a non-zero gradient at the pixel's center.

To determine whether or not a pixel should be marked as a step edge pixel, its underlying grey tone intensity surface must be estimated on the basis of the pixels in its neighborhood. For

this, we use a functional form consisting of a linear combination of the tensor products of discrete orthogonal polynomials of up to degree three. The appropriate directional derivatives are easily computed from this kind of a function.

Upon comparing the performance of this zero crossing of second directional derivative operator with Prewitt gradient operator and the Marr-Hildreth zero crossing of Laplacian operator, we find that it is the best performer and is followed by the Prewitt gradient operator. The Marr-Hildreth zero-crossing of Laplacian operator performs the worst.

I. Introduction

What is an edge in a digital image? The first intuitive notion is that a digital edge occurs on the boundary between two pixels when the respective brightness values of the two pixels are significantly different. Significantly different may depend upon the distribution of brightness values around each of the pixels.

We often point to a region on an image and say this region is brighter than its surrounding area, meaning that the mean of the brightness values of pixels inside the region is brighter than the mean of the brightness values outside the region. Having noticed this we would then say that an edge exists between each pair of neighboring pixels where one pixel is inside the brighter region and the other is outside the region. Such edges are referred to as step edges.

Step edges are not the only kind of edge. If we scan through a region in a left right manner observing the brightness values steadily increasing and then after a certain point observe that the brightness values are steadily decreasing we are likely to say that there is an edge at the point of change from increasing to decreasing brightness values. Such edges are called roof edges.

It is, therefore, clear from our use of the word edge that edge refers to places in the image where there appears to be a jump in brightness value or a local extrema in brightness value

derivative. Jumps in brightness values are the kinds of edges originally detected by Roberts (1965). Relative extrema of first derivative in a one dimensional form is used by Ehrich and Schroeder (1981) and in an isotropic two-dimensional suboptimal form by Marr and Hildreth (1980).

In some sense this summary statement about edges is quite revealing since in a discrete array of brightness values there are jumps, in the literal sense, between neighboring brightness values if the brightness values are different, even if only slightly different. Perhaps more to the heart of the matter, there exists no definition of derivative for a discrete array of brightness values. The only way to interpret jumps in value or local extrema of derivatives when referring to a discrete array of values is to assume that the discrete array of values comes about as some kind of sampling of a real-valued function defined on a bounded and connected subset of the real plane R^2. The jumps in value or extrema in derivative really must refer to points of high first derivative of f and to points of relative extrema in the second derivatives of f. Edge detection must then involve fitting a function to the sample values. Prewitt (1970), was the first to suggest the fitting idea. Heuckel (1971, 1973), Brooks (1978), Haralick (1980), Haralick and Watson (1981), Morgenthaler and Rosenfeld (1981), Zucker and Hummel (1979), and Morgenthaler (1981) all use the surface fit concept in determining edges.

Edge finders should then regard the digital picture function as a sampling of the underlying function f, where some kind of random noise has been added to the true function values. To do this, the edge finder must assume some kind of parametric form for the underlying function f, use the sampled brightness values of the digital picture function to estimate the parameters, and finally make decisions regarding the locations of discontinuities and the locations of relative extrema of partial derivatives based on the estimated values of the parameters.

Of course, it is impossible to determine the true locations of discontinuities in value or relative extrema in derivatives directly from a sampling of the functions. The locations are estimated by function approximation. Sharp discontinuities can reveal themselves in high values for estimates of first partial derivatives. Relative extrema in first directional derivative can reveal themselves as zero-crossings of the second directional derivative. Thus, if we assume that the first and second partial derivatives of any possible underlying image function have known bounds, then any estimated first or second order partials which exceed these known bounds must be due to discontinuities in value or in derivative of the underlying function. This is basis for the gradient magnitude and Laplacian magnitude edge detectors. However, edges can be weak but well localized. Such edges, as well as the strong edges just discussed, manifest themselves as local extrema of the deriva ve taken across the edge. This idea

for edges is the basis of the edge detector discussed here.

In this paper, we assume that in each neighborhood of the image the underlying function f takes the parametric form of a polynomial in the row and column coordinates and that the sampling producing the digital picture function is a regular equal interval grid sampling of the square plane which is the domain of f. As just mentioned, we place edges not at locations of high gradient, but at locations of spatial gradient maxima. More precisely, a pixel is marked as an edge pixel if in the pixel's immediate area there is a zero crossing of the second directional derivative taken in the direction of the gradient. Thus this kind of edge detector will respond to weak but spatially peaked gradients.

The underlying functions from which the directional derviatives are computed are easy to represent as linear combinations of the polynomials in any polynomial basis set. That polynomial basis set which permits the independent estimation of each coefficient would be the easiest to use. Such a polynomial basis set is the discrete orthogonal polynomial basis set.

Section II discusses the polynomials. In section II.1 we discuss how to construct the one dimensional family of discrete orthogonal polynomials. In section II.2 we discuss how arbitrary two dimensional polynomials can be computed as linear combinations of the tensor products of one dimensional discrete

orthogonal polynomials. In section II.3, we discuss how the discretely sampled data values are used to estimate the coefficients of the linear combinations: coefficient estimates for exactly fitting or estimates for least square fitting are calculated as linear combinations of the sampled data values.

Having used the pixel values in a neighborhood to estimate the underlying polynomial function we can now determine the value of the partial derivatives at any location in the neighborhood and use those values in edge finding. Having to deal with partials in both the row and column directions makes using these derivatives a little more complicated than using the simple derivatives of one dimensional functions. Section III discusses the directional derivative, how it is related to the row and column partial derivatives, and how the coefficients of the fitted polynomial get used in the edge detector. In section IV we discuss the statistical confidence of the estimate of edge existence and the edge angle. In section V we show results indicating the superiority of the directional derivative zero-crossing edge operator over the Prewitt gradient operator and the related Marr-Hildreth zero-crossing of the Laplacian operator.

II. The Discrete Orthogonal Polynomials

These polynomials are sometimes called the discrete Chebychev polynomials (Beckmann, 1973). In this section we show

how to construct them for one or two variables and how to use them in fitting data.

II.1 Discrete Orthogonal Polynomial Construction Technique

Let the index set R be symmetric in the sense that r∈R implies -r∈R. Let $P_n(r)$ be the n^{th} order polynomial. We define the construction technique for discrete orthogonal polynomials iteratively.

Define $P_0(r) = 1$.

Suppose $P_0(r),\ldots, P_{n-1}(r)$ have been defined. In general, $P_n(r) = r^n + a_{n-1}r^{n-1} + \ldots + a_1 r + a_0$. $P_n(r)$ must be orthogonal to each polynomial $P_0(r),\ldots,P_{n-1}(r)$. Hence, we must have the n equations

$$\sum_{r\in R} P_k(r)(r^n + a_{n-1}r^{n-1} + \ldots + a_1 r + a_0) = 0,\ k=0,\ldots,n-1 \quad (1)$$

These equations are linear equations in the unknown $a_0,\ldots,a_{n-1}$ and are easily solved by standard techniques.

The first five polynomial functions formulas are

$$P_0(r) = 1$$

$$P_1(r) = r$$

$$P_2(r) = r^2 - \mu_2/\mu_0$$

$$P_3(r) = r^3 - (\mu_4/\mu_2)r$$

$$P_4(r) = \frac{r^4 + (\mu_2\mu_4-\mu_6)r^2 + (\mu_2\mu_6-\mu_4^2)}{\mu_0\mu_4-\mu_2}$$

where

$$\mu_k = \sum_{s\in R} s^k$$

II.2 Two Dimensional Discrete Orthogonal Polynomials

Two dimensional discrete orthogonal polynomials can be created from two sets of one dimensional discrete orthogonal polynomials by taking tensor products. Let R and C be index sets satisfying the symmetry condition $r\in R$ implies $-r\in R$ and $c\in C$ implies $-c\in C$. Let $\{P_0(r),\ldots,P_N(r)\}$ be a set of discrete polynomials on R. Let $\{Q_0(c),\ldots,Q_M(c)\}$ be a set of discrete polynomials on C. Then the set $\{P_0(r)Q_0(c),\ldots,P_n(r)Q_m(c),\ldots,P_N(r)Q_M(c)\}$ is a set of discrete polynomials on RXC.

The proof of this fact is easy. Consider whether $P_i(r)Q_j(c)$ is orthogonal to $P_n(r)Q_m(c)$. when $n \neq i$ or $m \neq j$. Then

$$\sum_{r \in R} \sum_{c \in C} P_i(r)Q_j(c)P_n(r)Q_m(c)$$

$$= \sum_{r \in R} P_i(r)P_n(r) \sum_{c \in C} Q_j(c)Q_m(c).$$

Since $n \neq i$ or $m \neq j$ one or other of the sums must be zero.

Examples:

Index Set	Discrete Orthogonal Polynomial Set
$\{-1/2, 1/2\}$	$\{1, r\}$
$\{-1, 0, 1\}$	$\{1, r, r^2 - 2/3\}$
$\{-2/3, -1/2, 1/2, 3/2\}$	$\{1, r, r^2 - 5/4, r^3 - 41/20r\}$
$\{-2, -1, 0, 1, 2\}$	$\{1, r, r^2 - 2, r^3 - 17/5, r^4 + 3r^2 + 72/35\}$
$\{-1, 0, 1\} \times \{-1, 0, 1\}$	$\{1, r, c, r^2 - 2/3, rc, c^2 - 2/3, r(c^2 - 2/3), c(r^2 - 2/3), (r^2 - 2/3)(c^2 - 2/3)\}$

Figure 1 and 2 show some of the window masks used for the 3 x 3 and 4 x 4 cases.

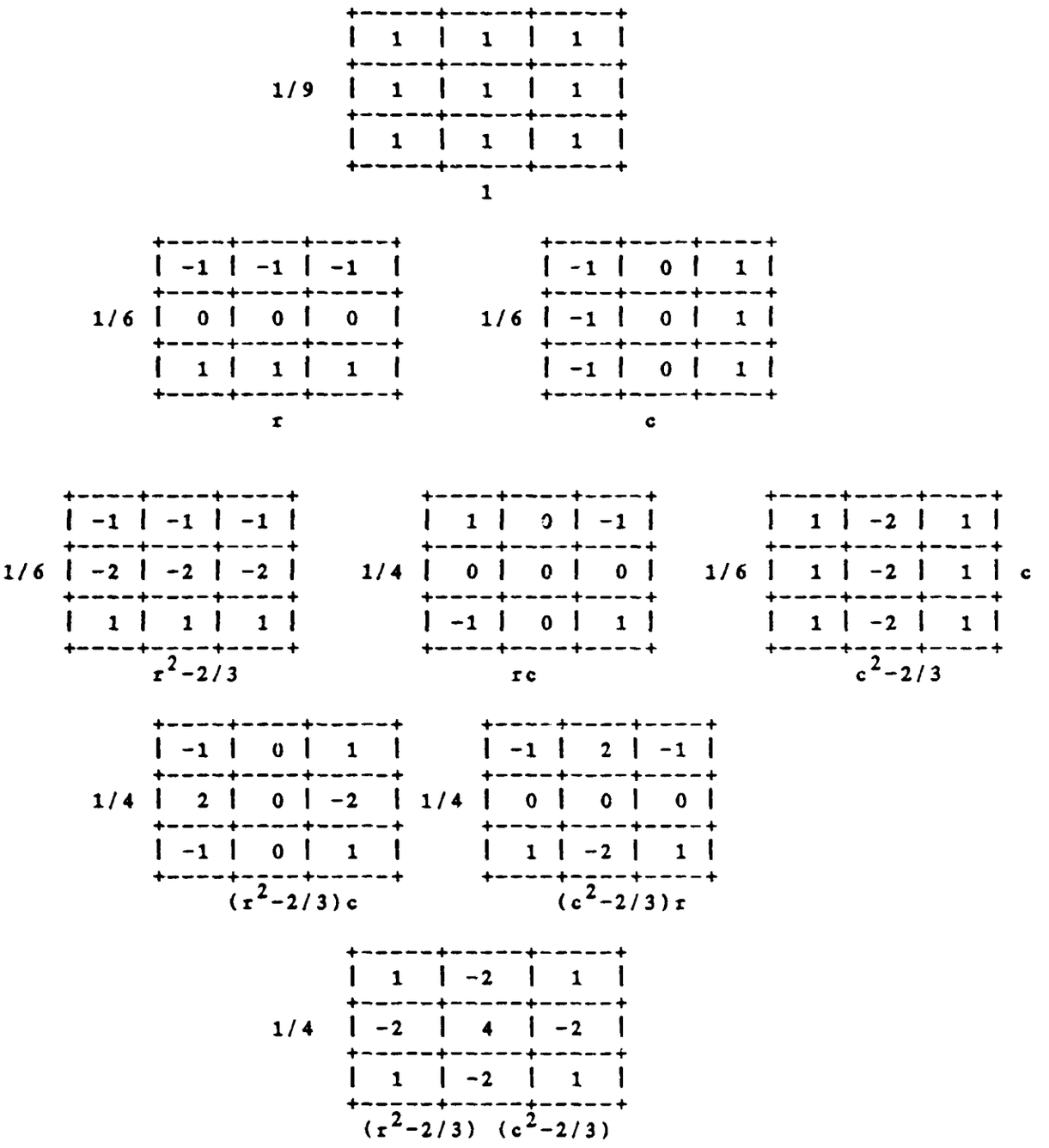

Figure 1 illustrates the 9 masks for the 3x3 window

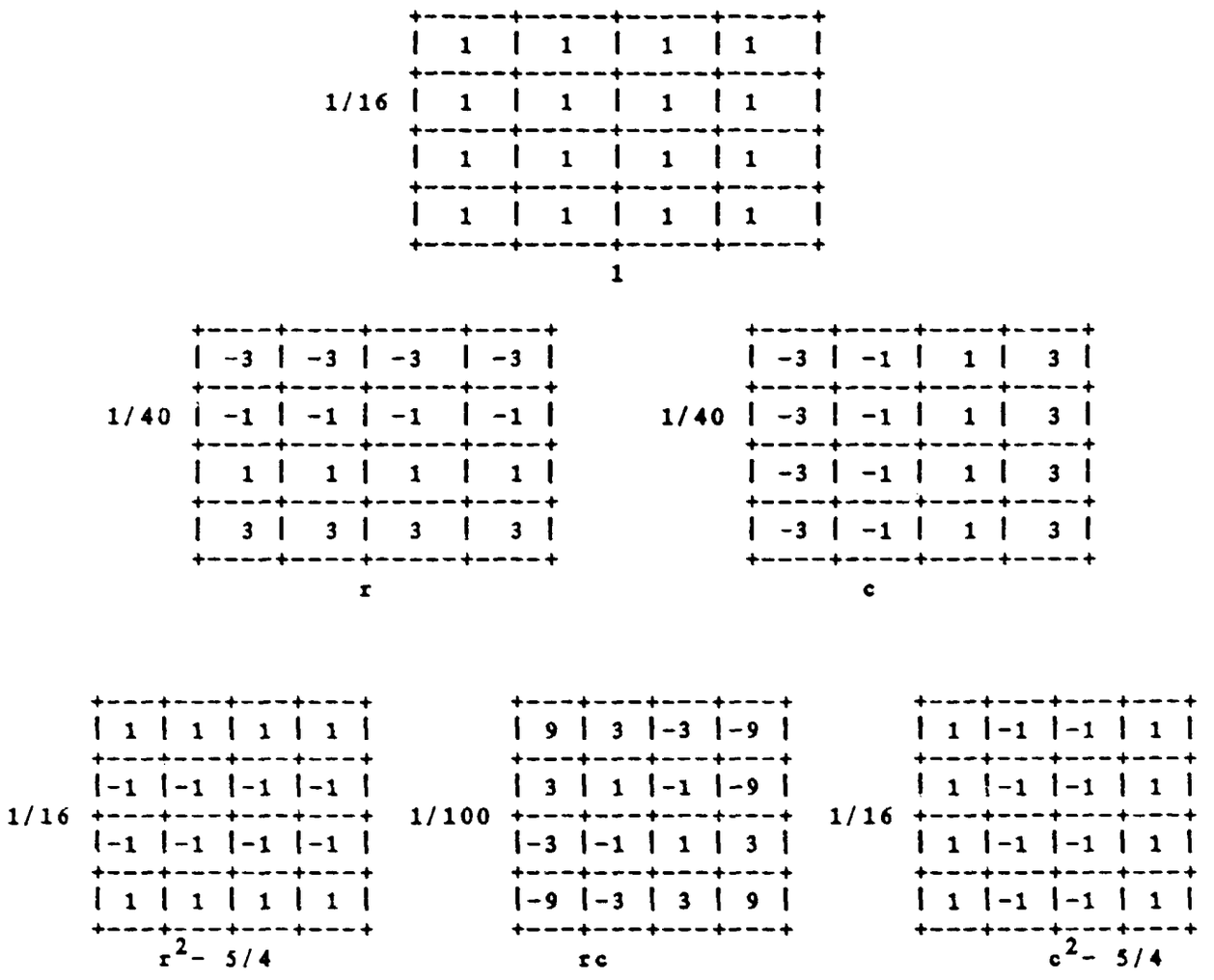

Figure 2 illustrates the masks used to obtain the coefficients of all polynomials up to the quadratic ones for a 4x4 window.

II.3 Fitting Data With Discrete Orthogonal Polynomials

Let an index set R with the symmetry property $r \varepsilon R$ implies $-r \varepsilon R$ be given. Let the number of elements in R be N. Using the construction technique, we may construct the set $\{P_0(r),\ldots,P_{N-1}(r)\}$ of discrete orthogonal polynomials over R.

For each $r \varepsilon R$, let a data value d(r) be observed. The exact fitting problem is to determine coefficients $a_0,\ldots,a_{N-1}$ such that

$$d(r) = \sum_{n=0}^{N-1} a_n P_n(r)$$

The orthogonality property makes the determination of the coefficients particularly easy. To find the value of some coefficient, say a_m, multiply both sides of the equation by $P_m(r)$ and then the sum over all $r \varepsilon R$.

$$\sum_{r \varepsilon R} P_m(r) d(r) = \sum_{n=0}^{N-1} a_n \sum_{r \varepsilon R} P_n$$

Hence,

$$a_m = \sum_{r \in R} P_m(r)d(r) \; / \; \sum_{r \in R} P_m^2(r) \qquad (2)$$

The approximate fitting problem is to determine coefficients $a_0,\ldots,a_K$, $K \leq N-1$ such that

$$e^2 = \sum_{r \in R} [d(r) - \sum_{n=0}^{K} a_n P_n(r)]^2$$

is minimized. To find the value of some coefficient, say a_m, take the partial derivative of both sides of the equation for e^2 with respect to a_m. Set it to zero and use the orthogonality property to find that again

$$a_m = \sum_{r \in R} P_m(r)d(r) / \sum_{r \in R} P_m^2(r) \qquad (3)$$

The exact fitting coefficients and the least squares coefficients are identical for $m = 0,\ldots,K$.

Fitting the data values $\{d(r) | r \in R\}$ to the polynomial

$$Q(r) = \sum_{n=0}^{K} a_n P_n(r)$$

now permits us to interpret Q(r) as a well behaved real-valued function defined on the real line. To determine

$$\frac{dQ}{dr}(r_0)$$

we need only to evaluate

$$\sum_{n=0}^{N} a_n \frac{dP_n}{dr}(r_0)$$

In this manner, any derivative at any point may be obtained. Similarly for any definite integrals. Beaudet (1978) uses this technique for estimating derivatives employed in rotationally invariant image operators.

It should be noted that the kernel used to estimate a derivative depends on the neighborhood size, the order of the fit, and the basis functions used for the fit. Figure 3 illustrates one example of the difference the assumed model makes. This difference means that the model used must be justified, the justification being that it is a good fit to the data. In particular, a not sufficiently good justification for

using first order models is that first order partial derivatives are being estimated.

Assumed Model	Kernel Mask for Row Derivative
$g(r,c)=a_{00}+a_{10}r+a_{01}c$	$1/6 \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$
$g(r,c)= a_{00}+a_{10}r+a_{01}c$ $a + a_{20}(r^2-2/3)+a_{11}rc +a_{02}(c^2-2/3)$ $+ a_{21}(r^2-2/3)c + a_{12}(c^2- 2/3)r$	$1/2 \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

Figure 3 illustrates that the assumed model does make a difference in the kernel mask used to estimate a quantity such as row derivative.

III. The Directional Derivative Edge Finder

We denote the directional derivative of f at the point (r,c) in the direction α by $f_\alpha'(r,c)$. It is defined as

$$f_\alpha'(r,c) = \lim_{h\to 0} \frac{f(r+h\sin\alpha, c+h\cos\alpha) - f(r,c)}{h} \qquad (4)$$

The direction angle α is the clockwise angle from the column axis. It follows directly from this definition that

$$f_\alpha'(r,c) = \frac{\partial f(r,c)}{\partial r} \sin\alpha + \frac{\partial f(r,c)}{\partial c} \cos\alpha \qquad (5)$$

We denote the second directional derivative of f at the point (r,c) in the direction α by $f_\alpha''(r,c)$ and it quickly follows that

$$f_\alpha'' = \frac{\partial^2 f \sin^2\alpha}{\partial r^2} + \frac{2\partial^2 f \sin\alpha \cos\alpha}{\partial r c} + \frac{\partial^2 f \cos^2\alpha}{\partial c^2} \qquad (6)$$

Taking f to be a cubic polynomial in r and c which can be estimated by the discrete orthogonal polynomial fitting procedure, we can compute the gradient of f and the gradient

direction angle at the center of the neighborhood used to estimate f. Letting f be estimated as a two dimensional cubic

$$\begin{aligned} f(r,c) = {} & k_1 + k_2 r + k_3 c \\ & + k_4 r^2 + k_5 rc + k_6 c^2 \\ & + k_7 r^3 + k_8 r^2 c + k_9 r c^2 + k_{10} c^3 \end{aligned} \tag{7}$$

we obtain α by

$$\begin{aligned} \sin\alpha &= k_2/(k_2^2 + k_3^2)^{.5} \\ \cos\alpha &= k_3/(k_2^2 + k_3^2)^{.5} \end{aligned} \tag{8}$$

At any point (r,c), the second directional derivative in the direction α is given by

$$\begin{aligned} f_\alpha''(r,c) = {} & (6k_7 \sin^2\alpha + 4k_8 \sin\alpha\cos\alpha + 2k_9 \cos^2\alpha)r \\ & + (6k_{10} \cos^2\alpha + 4k_9 \sin\alpha \cos\alpha + 2k_8 \sin^2\alpha)c \\ & + (2k_4 \sin^2\alpha + 2k_5 \sin\alpha \cos\alpha + 2k_6 \cos^2\alpha) \end{aligned} \tag{9}$$

We wish to only consider points (r,c) on the line in direction α. Hence, $r=\rho\sin\alpha$ and $c=\rho\cos\alpha$. Then

$$f_\alpha''(\rho) = 6[k_7\sin^3\alpha + k_8\sin^2\alpha\cos\alpha + k_9\sin\alpha\cos^2\alpha + k_{10}\cos^3\alpha]\rho + 2[k_4\sin^2\alpha + k_5\sin\alpha\cos\alpha + k_6\cos^2\alpha] = A\rho + B \tag{10}$$

If for some ρ, $|\rho| < \rho_0$, $f_\alpha''(\rho) = 0$ and $f_\alpha'(\rho) \neq 0$ we have discovered a zero-crossing of the second directional derivative taken in the direction of the gradient and we mark the center pixel of the neighborhood as an edge pixel.

IV. Statistical Analysis

In this section we show how the randomness of the noise induces a randomness in the least squares coefficients and then how the randomness of the least squares coefficients induces a randomness in the estimated gradient value, the estimated angle of the gradient, and the estimated location of the zero-crossing.

IV.1 General Model

We let p_n, $n=1,\ldots,N$ denote the names of the discrete orthonormal basis functions, η denote the independent and identically distributed noise, and g denote the gray tone intensity function. Under this model, the observed image can be

written as

$$g(r,c) = \sum_{n=1}^{N} a_n \, p_n(r,c) + \eta(r,c) \tag{11}$$

where $$\sum_{r,c} p_n(r,c) \, p_m(r,c) = \begin{cases} 0, & n \neq m \\ 1 & n = m \end{cases}$$

and the least squares estimates $a_1', \ldots, a_N'$ for the unknown coefficients $a_1, \ldots, a_n$ are given by

$$a_n' = \sum_{r,c} g(r,c) \, p_n(r,c) \tag{12}$$

Substituting the formula for $g(r,c)$ into the equation for a_n' and simplifying results in

$$a_n' = a_n + \sum_{r,c} p_n(r,c) \, \eta(r,c) \tag{13}$$

clearly showing that a_n' has a deterministic part and a random part, the randomness being due to the noise. We assume that the noise is independent normal having mean 0 and variance σ^2. Therefore, the estimated coefficient a_n' has mean a_n, variance σ^2 and is uncorrelated with every other coefficient:

$$E\,[a_n'] = a_n$$

$$E\,[a_m'\,a_n'] = a_m a_n,\ m \neq n$$

$$E\,[a_n'^2] = a_n^2 + \sigma^2$$

$$V\,[a_n'] = \sigma^2$$

The residual error e is defined as the difference between the observed values and fitted values. It too is a random variable.

$$e(r,c) = g(r,c) - \sum_{n=1}^{N} a_n'\,p_n(r,c) \tag{14}$$

$$= \sum_{n=1}^{N} (a_n - a_n')\,p_n(r,c) + \eta(r,c)$$

It is not difficult to see that at each (r,c), the residual error has mean zero and is uncorrelated with each estimated coefficient a_n' since

$$E\,[a_n'\,e(r,c)] = 0$$

After some algebraic substitutions and manipulation, the total residual error, S^2, can be written as

$$S^2 = \sum_{r,c} e^2(r,c) = \sum_{r,c} \eta^2(r,c) - \sum_{n=1}^{N} (a_n - a_n')^2 \qquad (15)$$

Thus, if the noise is assumed normal and there are K pixels in a window

$$\sum_{r,c} \eta^2(r,c)/\ \sigma^2 \text{ has } X_k^2,$$

a chi-squared variate with K degree of freedom,

$$\sum_{n=1}^{N} (a_n - a_n')^2/\sigma^2 \quad \text{has } X_N^2$$

which makes $\sum_{r,c} e^2(r,c)$ have X_{k-N}^2

IV.2 Estimating the First Partials

If the discrete orthogonal basis functions are polynomials then each first partial derivative at (0,0) in the row and column directions is given as some linear combination of the estimated coefficients. Furthermore, the linear combination for the row partial will be orthogonal to the linear combination in the

column partial. Letting the coefficients of the linear combination for the row partial be $s_1,\ldots,s_N$ and the coefficients of the linear combination for the column partial be $t_1,\ldots,t_N$, where

$$\sum_{n=1}^{N} s_n^2 = \sum_{n=1}^{N} t_n^2 = k,$$

we have,

$$\mu_r = \sum_{n=1}^{N} s_n a_n$$

$$\mu_c = \sum_{n=1}^{N} t_n a_n$$

as the true but unknown values of the row and column partials. The estimates are

$$\mu_r' = \sum_{n=1}^{N} s_n \, a_n'$$

$$\mu_c' = \sum_{n=1}^{N} t_n \, a_n'$$

and they have mean and variance given by

$$E[\mu_r'] = \mu_r$$

$$E[\mu_c'] = \mu_c$$

$$V[\mu_r'] = \sigma^2 k$$

$$V[\mu_c'] = \sigma^2 k$$

$$E[\mu_r' \mu_c'] = \mu_r \mu_c$$

Hence, the estimates for the row and column partial derivatives are uncorrelated.

IV.3 Hypothesis Testing For Zero Gradient

To see the effect of the randomness on the estimate of the gradient magnitude, consider testing the hypothesis that $\mu_r = \mu_c = 0$. This hypothesis must be rejected if there is to be a zero-crossing of second directional derivative. Under this hypothesis,

$$\frac{\mu'^2_r + \mu'^2_c}{k\,\sigma^2}$$

has a X^2_2 distribution.

The total residual error normalized by the noise variance, S^2/σ^2, has a X^2_{K-N} distribution. Hence

$$\frac{(\mu'^2_r + \mu'^2_c)/2}{k\,S^2/(K-N)}$$

has a $F_{2,K-N}$ distribution and the hypothesis of $\mu_r = \mu_c = 0$ would be rejected for suitably large values.

IV.4 Confidence Interval For Gradient Direction

To see the effect of the randomness on the estimate of the direction of the gradient, consider the relationships portrayed in figure '. The axes are the row and column partials μ_r and μ_c. The direction angle θ of the gradient is given by

$$\cos\theta = \mu_r/(\mu_r^2 + \mu_c^2)^{1/2}$$
$$\sin\theta = \mu_c/(\mu_r^2 + \mu_c^2)^{1/2} \qquad (16)$$

The center of the circle is at the estimate (μ_r', μ_c'). Upon substituting the estimates μ_r' and μ_c' for μ_r and μ_c, we obtain the estimated direction angle θ' by

$$\cos\theta' = \mu_r'/(\mu'^2_r + \mu'^2_c)^{1/2}$$
$$\sin\theta' = \mu_r'/(\mu'^2_r + \mu'^2_c)^{1/2} \qquad (17)$$

From a Bayesian point of view, the area of the circle represents the conditional probability that the unknown (μ_r, μ_c) lies within a distance R from the observed (μ_r', μ_c') given that the variance of μ_r' and μ_c' is known and equal to $k\sigma^2$. Assuming a normal distribution for the noise, this conditional probability is $q = 1-e^{-R^2/2k\sigma^2}$ Hence, if probability q is given, the corresponding radius R is

$$R = k\,\sigma\,[-2\,\log(1-q)]^{1/2} \qquad (18)$$

To determine a confidence interval for θ of the form $\theta' - \Delta \leq \theta \leq \theta' + \Delta$, we have from figure 4 that

$$\sin^2\Delta = \frac{k\,\sigma^2(-2\,\log(1-q))}{\mu'^2_r + \mu'^2_c} \qquad (19)$$

Note that the 2Δ confidence interval length depends on the probability q of the circle confidence region for (μ_r, μ_c) and the unknown noise variance σ^2. Although σ^2 is not known, we do know S^2 which has a $\sigma^2 X^2_{K-N}$ distribution. We can handle the problem of ths unknown σ^2 by determining a joint confidence region for (μ_r, μ_c) and σ^2 (Foutz,1981). Taking p to be the probability that a chi-squared random variable with $K-N$ degrees of freedom has an observed value greater than $X^2_{K-N,p}$ we have the confidence interval $[0, S^2/X^2_{K-N,p}]$ for σ^2 having at least probability p. Replacing σ^2 in equation (19) by $S^2/X^2_{N-K,p}$ we obtain

$$\sin^2\Delta = \frac{k\, S^2(-2\log(1-q))}{X^2_{K-N,p}\,(\mu'^2_r + \mu'^2_c)} \qquad (20)$$

A confidence interval for θ having at least probability pq is then $(\theta' - \Delta, \theta' + \Delta)$.

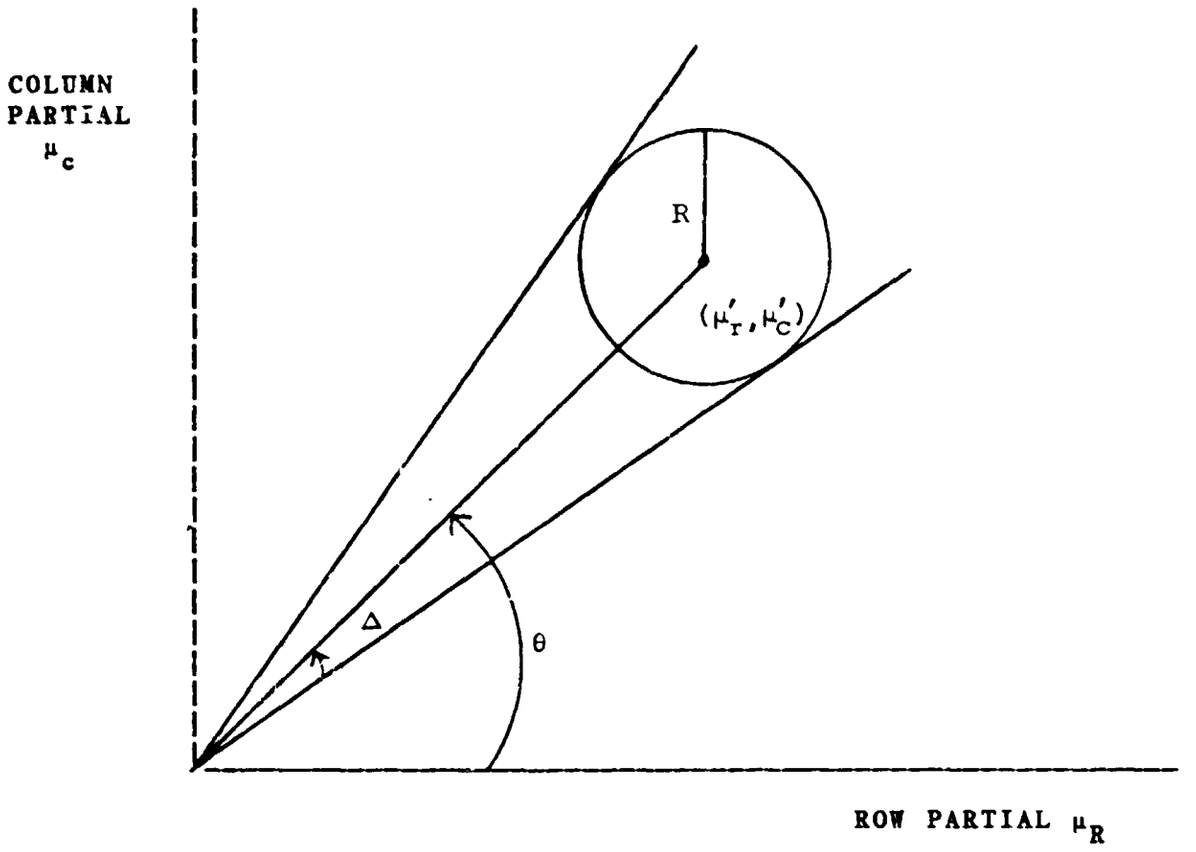

Figure 4 illustrates the geometry of the confidence interval estimation for the edge angle.

IV.5 Edge Hypothesis Testing

In this section we first take the edge direction α to be a fixed constant. We let μ_A and μ_B be the expected values of the random variables A and B appearing in equation (10). The null hypothesis is that an edge exists. The null hypothesis is satisfied if for some ρ, $o \leq \rho \leq d$, $\mu_A\rho + \mu_B = 0$.

The observed random variables are A, B, and the residual fitting error S^2. The bivariate random variable $\begin{pmatrix} A \\ B \end{pmatrix}$ is normal having mean $\begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix}$ and covariance $\sigma^2 \begin{bmatrix} k_A & 0 \\ 0 & k_B \end{bmatrix}$ where k_A and k_B are known constants. For a window of K pixels and a cubic fit, S^2/σ^2 has a X^2_{K-10}.

From this it follows that

$$Z(\mu_A,\mu_B) = \frac{[(A-\mu_A)/k_A]^2 + [(B-\mu_B/k_B{}^2)]/2}{S^2/(K-10)}$$

has an $F_{2\ K-10}$ distribution.

We define $R = \{(x,y) \mid$ for some ρ, $o \leq \rho \leq d$, $x\rho + y = 0\}$ Then the null hypothesis is rejected at the p significance level if

$$\min_{(\mu_A,\mu_B) \in R} Z(\mu_A,\mu_B)$$

is larger than $F_{2\ K-10\ 1-p}$.

An edge strength probability can be defined by q where q satisfies

$$\min_{(\mu_A,\mu_B)\in R} Z(\mu_A,\mu_B) = F_{2\ k-10\ q}$$

Of course the edge direction α is not fixed. But we do have a confidence interval for it. And for each value of α in the confidence interval, the random variable $A(\alpha)$ and $B(\alpha)$ can be computed and the null hypothesis tested. If for all α in the confidence interval the null hypothesis is rejected, then the existence of an edge is also rejected.

In practice, we can perform a non-exact hypothesis test selecting only the left end, middle, and right end values of α from its confidence interval. If for each of these three values of α the null hypothesis is rejected, then the existence of an edge is also rejected.

V. Experimental Results

To understand the performance of the second directional derivative zero-crossing digital step edge operator we examine its behavior on a well structured simulated data set and on a real aerial image. For the simulated data set, we use a 100x100 pixel image of a checkerboard, the checks being 20x20 pixels. The dark checks have gray tone intensity 75 and the light checks

have gray tone intensity 175. To this perfect checkerboard we add independent Gaussian noise having mean zero and standard deviation 50. Defining the signal to noise ratio as 10 times the logarithm of the range of signal divided by RMS of the noise, the simulated image has a 3 db signal to noise ratio. The perfect and noisy checkerboards are shown in figure 5.

Section V.1 illustrates the performance of the classic 3x3 edge operators with and without preaveraging compared against the generalized Prewitt operator. Section V.2 illustrates the performance of the Marr-Hildreth zero-crossing of Laplacian operator, the 11x11 Prewitt operator, and the 11x11 zero-crossing of second directional derivative operator. The zero-crossing of second directional derivative surpasses the performance of the other two on the twofold basis of probability of correct assignment and error distance which is defined as the average distance to closest true edge pixel of pixels which are assigned non-edge but which are true edge pixels.

V.1 The Classic Edge Operators

The classic 3x3 gradient operators all perform badly as shown in figure 6. Note that the usual definition of the Roberts operator has been modified in the natural way so that it uses a 3x3 mask.

Averaging before the application of the gradient operator is considered to be the cure for such bad performance on noisy

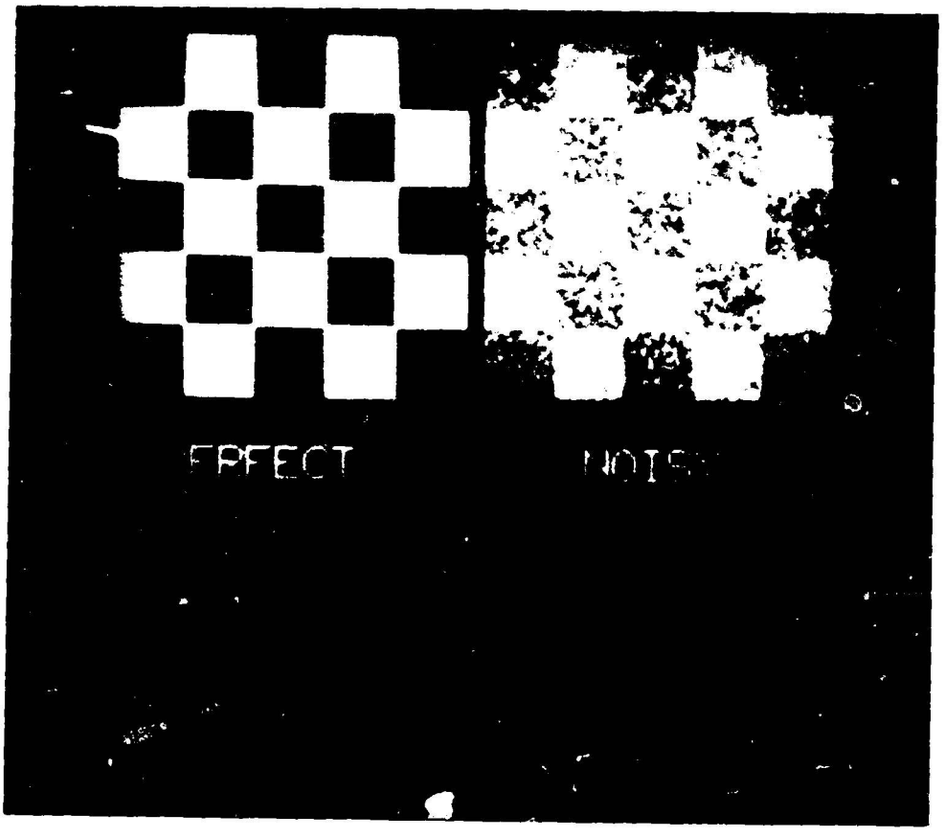

Figure 5 illustrates the noisy checkerboard used in the experiments. Low intensity is 75 high intensity is 175. Standard deviation of noise is 50.

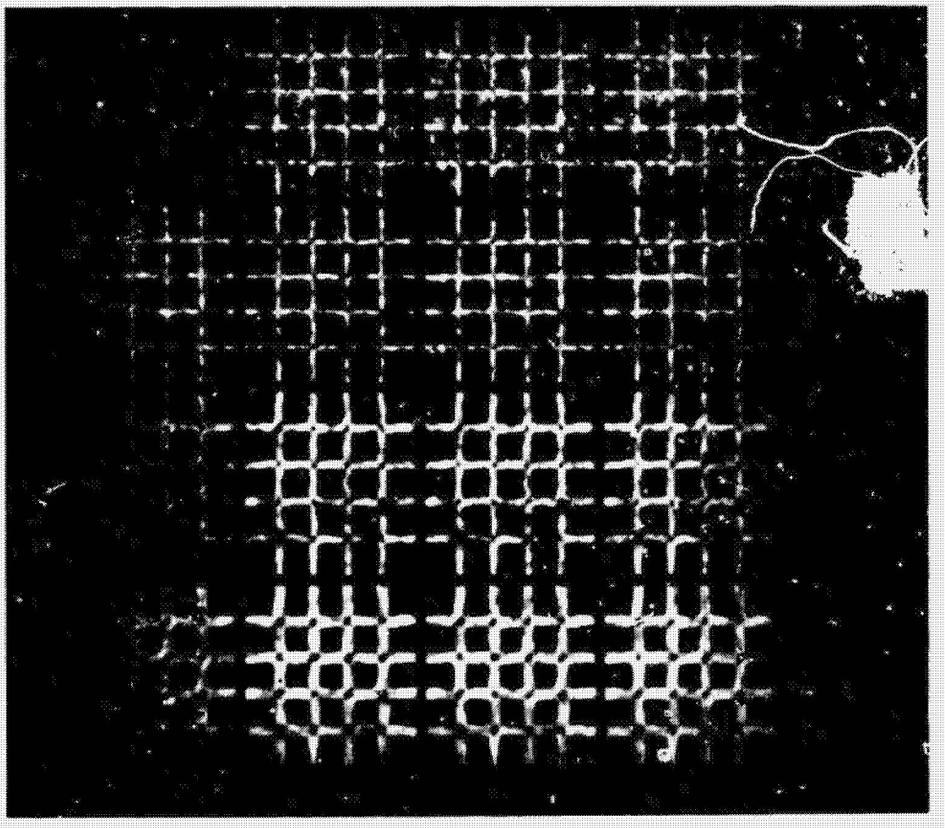

Figure 6 illustrates the 3x3 Roberts, Sobel, Prewitt, and Kirsch edge operators with a box filter preaveraging of 1x1, 3x3, 5 5 and 7x7.

images (Rosenfeld and Kak, 1976). Figure 6 also shows the same operators applied after a box filtering with a 3x3, 5x5, and 7x7 neighborhood sizes.

An alternative to the preaveraging is to define the gradient operator with a larger window. This is easily done with the Prewitt operator (Prewitt,1970) which fits a quadratic surface in every window and uses the square root of the sum of the squares of the coefficients of the linear terms to estimate the gradient. (A linear fit actually yields the same result for the polynomial basic function. A cubic fit is the first higher order fit which would yield a different result.) This is illustrated in figure 7. A 3x3 pre-average followed by a 3x3 gradient operator yields a resulting neighborhood size of 5x5. Thus in figure 7 we also show the 3x3 preaverage followed by a 3x3 gradient under the 5x5 Prewitt and we show the 5x5 pre-average followed by the 3x3 gradient under the 7x7 Prewitt. The noise is higher in the pre-average edge-detector. For comparison purposes the 5x5 Nevatia and Babu (1979) compass operator is shown alongside the 5x5 Prewitt in figure 8. They give virtually the same result. The Prewitt operator has the advantage of requiring half the computation.

It is obvious from these results that good gradient operators must have larger neighborhood sizes than 3x3. Unfortunately, the larger neighborhood sizes also yield thicker edges.

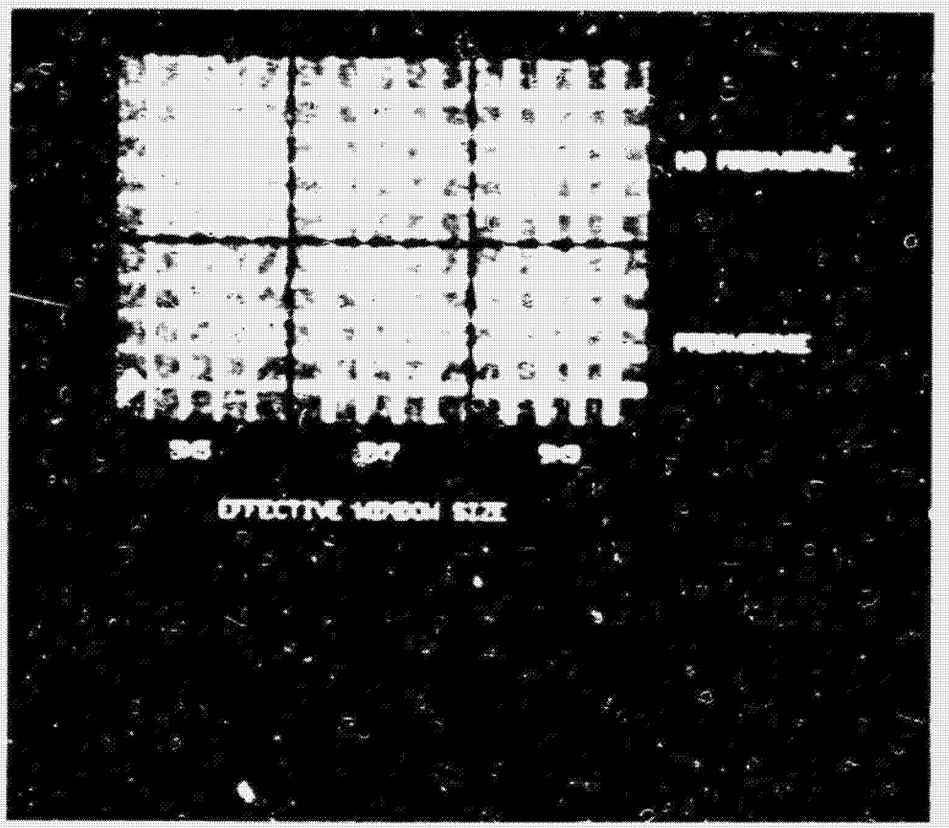

Figure 7 illustrates the Prewitt Operator done by using a least squares quadratic fit in the neighborhood versus doing preaveraging and using a smaller fitting neighborhood size. The no preaveraging results show slightly higher contrast.

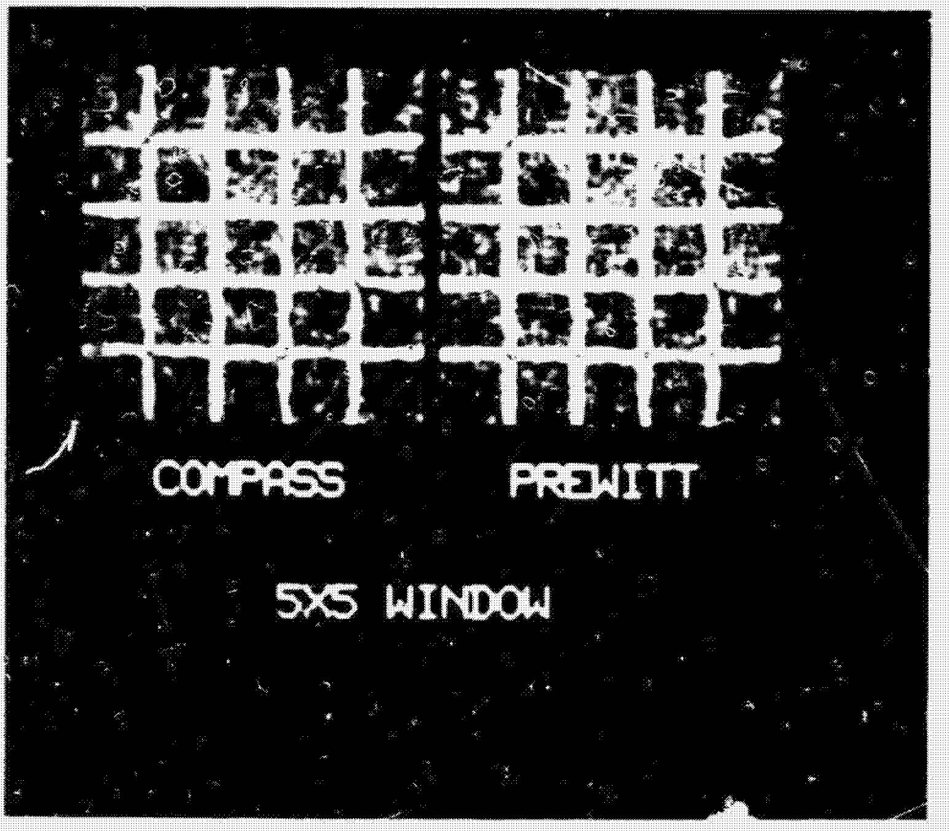

Figure 8 compares the Nevatia and Babu compass operator with the Prewitt operator in a 5x5 neighborhood.

To detect edges, the gradient value must be thresholded. In each case, we chose a threshold value which makes the conditional probability of assigning an edge given that there is an edge equal to the conditional probability of there being a true edge given that an edge is assigned. True edges are established by defining them to be the two pixel wide region in which each pixel neighbors some pixel having a value different from it on the perfect checkerboard. Figure 9 shows the thresholded Prewitt operator (quadratic fit) for a variety of neighborhood sizes. Notice that because the gradient is zero at the saddle points (the corner where four checks meet), any operator depending on the gradient to detect an edge will have trouble there.

V.2 The Second Derivative Zero Crossing Edge Operators

Marr and Hildreth (1980) suggest an edge operator based on the zero crossing of a generalized Laplacian. In effect, this is non-directional or isotropic second derivative zero crossing operator. The mask for this generalized Laplacian operator is given by sampling the kernel

$$1 - k\frac{r^2+c^2}{\sigma^2} e^{-1/2\frac{r^2+c^2}{\sigma^2}}$$

at row colum coordinates (r,c) designating the center of each pixel position in the neighborhood and then setting the value k so that the sum of the resulting weights is zero. Edges are

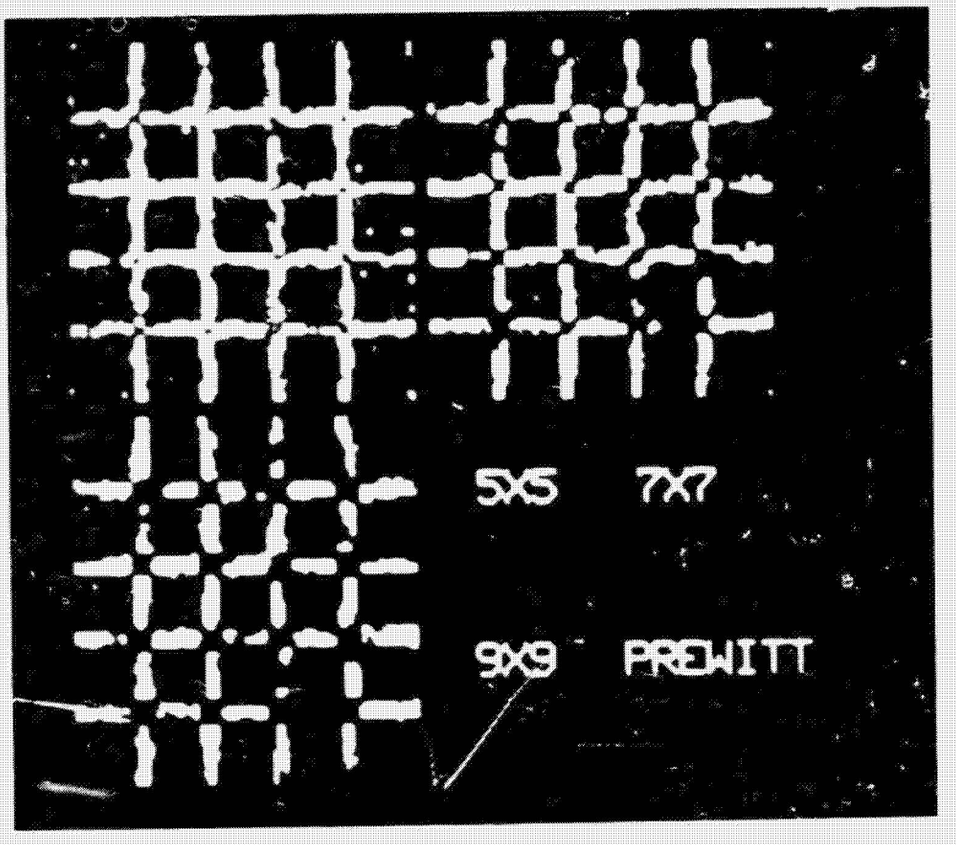

Figure 9 illustrates the edges obtained by thresholding the results of the Prewitt operator.

detected at all pixels whose generalized Laplacian value is of one sign and one of whose neighbors has a generalized Laplacian value of the opposite sign. A zero-crossing threshold strength can be introduced here by insisting that the difference between the positive value and the negative value must exceed the threshold value before the pixel is declared to be an edge pixel. Figure 10 illustrates the edge images produced by this technique for a variety of threshold values and a variety of values for σ for an 11 by 11 window. It is apparent that if all edge pixels are to be detected, there will be many pixels declared to be edge pixels which are really not edge pixels. And if there are to be no pixels which are to be declared edge pixels which are not edge pixels, then there will be many edge pixels which are not detected. Its performance is poorer than the Prewitt operator.

The directional second derivative zero crossing edge operator introduced in this paper is shown in figure 11 for a variety of gradient threshold values. If the gradient exceeds the threshold value and a zero-crossing occurs in a direction of $\pm$ 14.9 degrees of the gradient direction within a circle of one pixel length centered in the pixel, then the pixel is declared to be an edge pixel. This technique performs the worst at the saddle points, the corner where four checks meet because of these being a zero gradient there.

Table 1 shows the comparison among the Prewitt operator and the directional and the Marr-Hildreth non-directional second

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

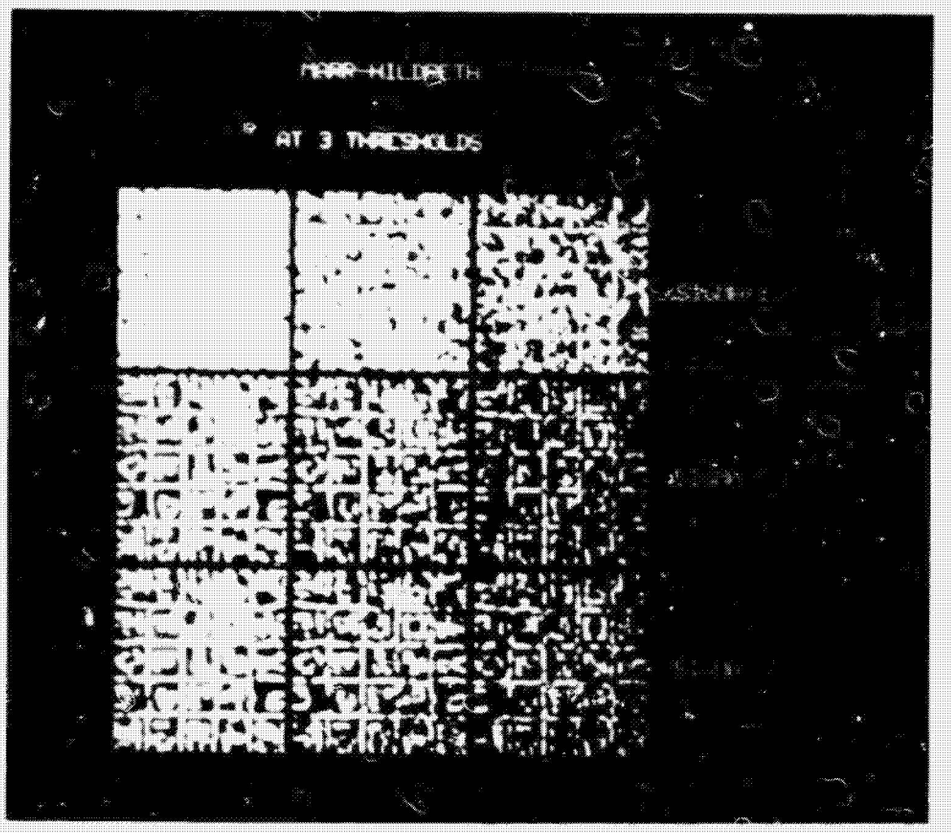

Figure 10 illustrates the edges obtained by the 11x11 Marr-Hildreth zero-crossing of Laplacian operator set for three different zero-crossing thresholds and three different standard deviations for the associated Mexican hat filter.

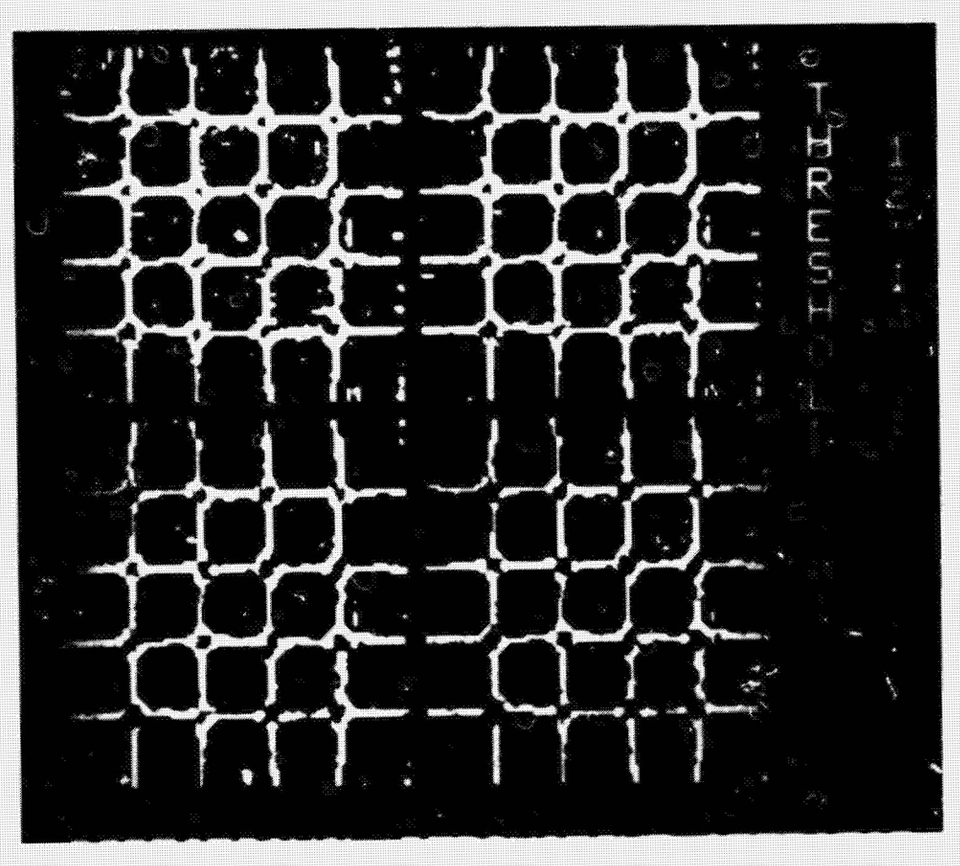

Figure 11 illustrates the directional derivative edge operator for 4 different thresholds.

derivative zero crossing edge operators. The threshold used is, as before, the one equalizing the conditional probability of assigned edge given true edge and the conditional probability of true edge given assigned edge. It is clear that the performance of the directional derivative operator is better than the Prewitt operator and the Marr-Hildreth operator, both on the basis of the correct assignment probability and the error distance which is the average distance to closest true edge pixels of pixels which are assigned non-edge labels but which are true edge pixels.

Figure 12 shows the corresponding edge images of the 11x11 Prewitt operator using a cubic fit rather than a quadratic fit, the 11x11 Marr-Hildreth operator, and the 11x11 directional derivative zero-crossing operator. The thresholds used are the ones to equalize the conditional probabilities as given in Table 1. A visual evaluation also leaves the impression that the directional derivative operator produces better edge continuity and has less noise than the other two.

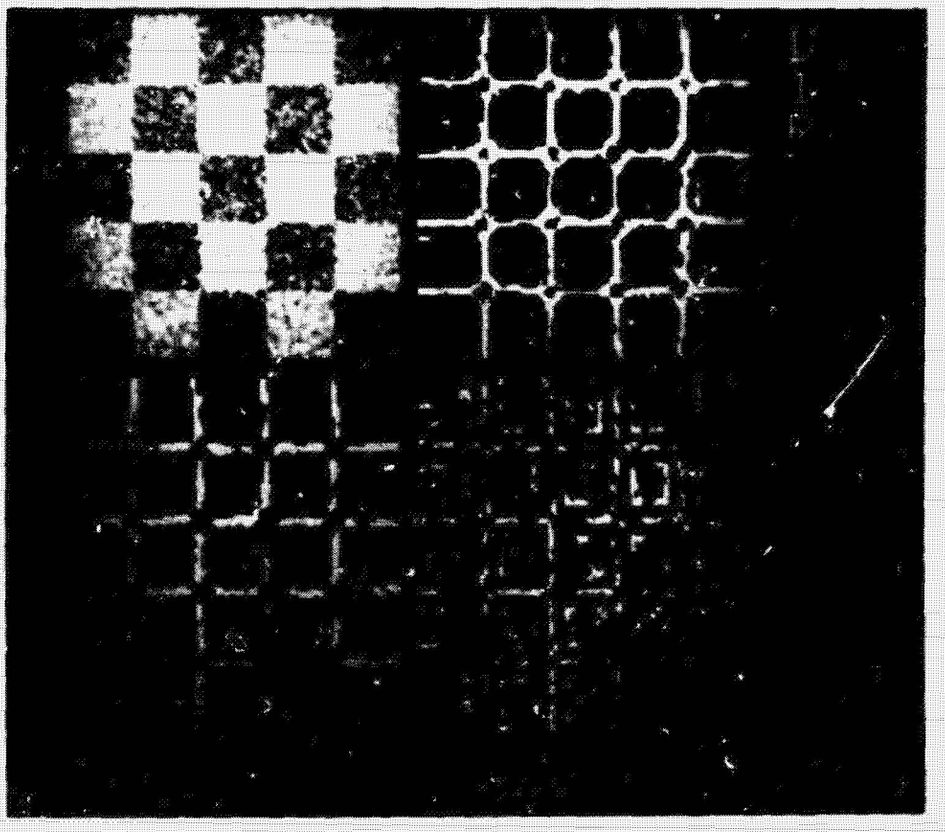

Figure 12 Compares the directional derivative edge operator with the Marr-Hildreth edge operator and the Prewitt edge operator. The thresholds chosen were the best possible ones.

For the case of constant variance additive noise, thresholding on the basis of the hypothesis test of section IV.3 yields essentially the same results as simply thresholding the gradient value.

Figure 13 illustrates the second directional derivative zero crossing operator on an aerial image which has been median filtered and then enhanced by replacing each pixel with the closer of its 3x3 neighborhood minimum or maximum. The technique is so good that it is possible to determine region boundaries essentially by doing a connected components on non-edge pixels. Figure 13b shows the cleaned edge image which is obtained by doing a connected components on the non edge pixels, then removing all pixels whose region has fewer than 20 pixels. The resulting boundaries are given as pixels which have a neighbor with a different label than its own.

Initial raw edges which leave gaps in a region boundary will in effect make the regions merge in the connected components step. Thus the small number of missing boundaries is surprising. To be sure, we are not advocating connected components as an image segmentation technique. The fact that it works as well as it does is an indication of the strength of the edge detector.

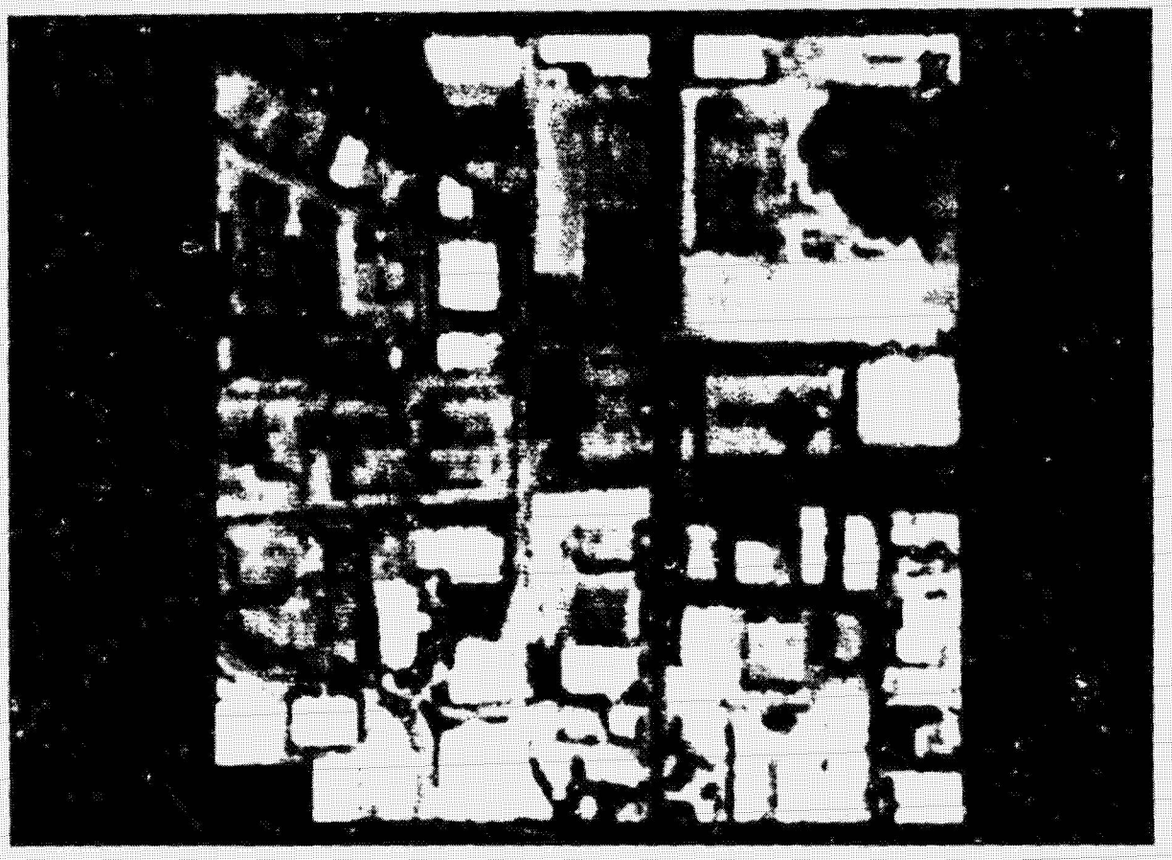

Figure 13a illustrates an aerial photograph.

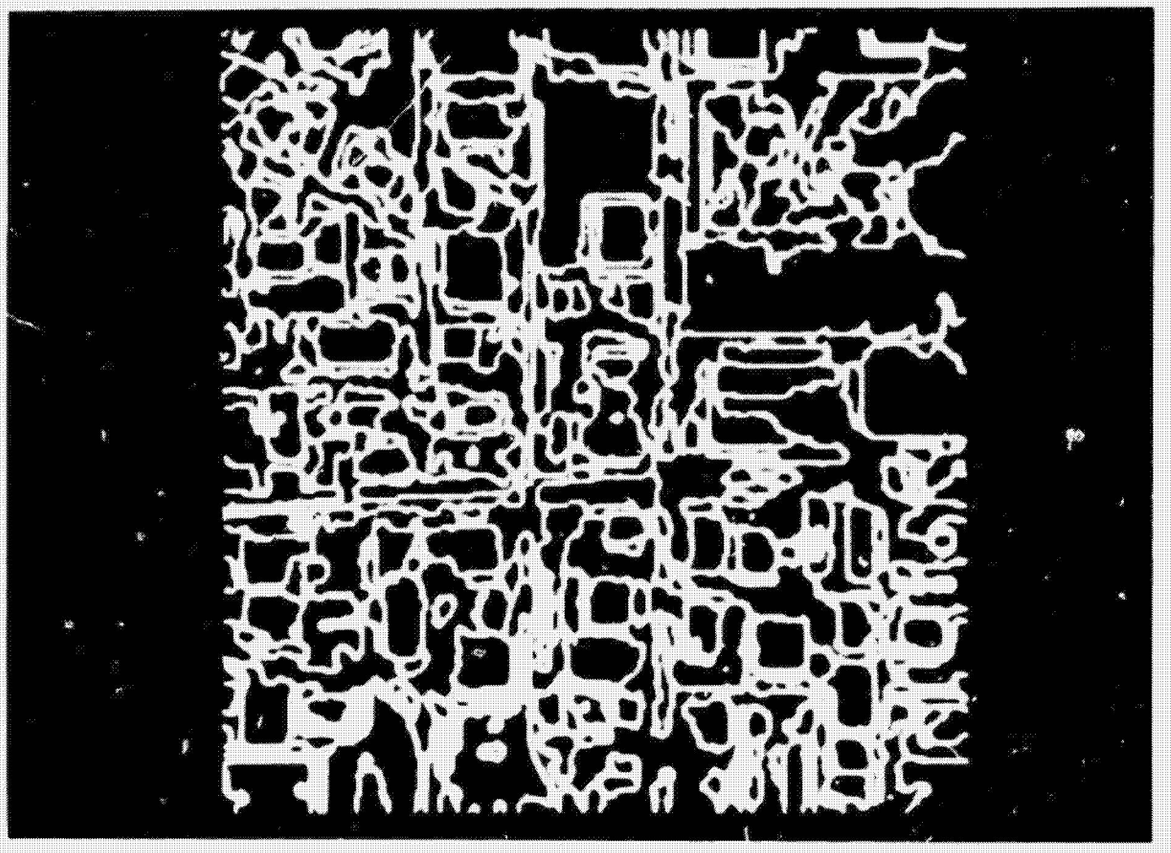

Figure 13b illustrates the directional derivative edges obtained from the aerial photograph by first 3x3 median filtering, then replacing each pixel by the closer of its 3x3 neighborhood minimum or maximum, then taking the directional derivative edges using a 7x7 window, then doing a connected components on the non-edge pixels, and removing all regions having fewer than 20 pixels, and then displaying any pixel neighboring a pixel different than it as an edge pixel.

	Prewitt	Marr-Hildreth	Directional Derivative
Parameters	Gradient Threshold = 18.5	Zero-crossing Strength = 4.0 σ = 5.0	Gradient Threshold=14.0 ρ = .5
P(AE|TE)	.6738	.3977	.7207
P(TE|AE)	.6872	.4159	.7197
Error Distance	1.79	1.76	1.16

Table 1 compares the performance of three edge operators using an 11x11 window on the noisy checkerboard image. Thresholds are chosen to equalize, as best as possible, P(AE|TE), the conditional probability of assigned edge given true edge and the conditional probability, P(TE|AE) of true edge given assigned edge. The error distance is the average distance to closest true edge pixels of pixels which are assigned non-edge but which are true edge.

VI. Conclusions

We have argued that numeric digital image operations should be explained in terms of their actions on the underlying gray tone intensity surface of which the digital image is an observed noisy sample. We called this model, the facet model for digital image processing and showed how the facet model can be used to estimate in each neighborhood the underlying gray tone intensity surface.

We described a digital step edge operator which detects edges at all pixels whose estimated second directional derivative taken in the direction of the gradient has a zero crossing within the pixel's area. We discussed the statistical analysis of this technique, illustrating how to determine confidence intervals for the direction of the gradient and how this [illegible] determines a confidence interval for the placement of the zero-crossing

We have compared the performance of the directional derivative zero crossing edge operator with that of the classic edge operators, the generalized Prewitt gradient operator, and the Marr-Hildreth zero crossing edge operator. We found that in both the simulated and real image data sets the directional derivative zero crossing edge operator had superior performance.

We have illustrated that for good performance it is important to use larger neighborhood sizes than 3x3 and have shown that better results are achieved by defining the edge operator naturally in the large neighborhood rather than pre-

averaging and then using a smaller neighborhood edge operator on the averaged image.

There is much work yet to be done. We need to explore the relationship of basis function kind, (polynomial, trignometric polynomial etc.), order of fit, and neighborhood size to the goodness of fit. Evaluation must be made of the confidence intervals produced by the technique. The technique needs to be generalized so that it works on saddle points created by two edges crossing. A suitable edge linking method needs to be developed which uses these confidence intervals. Ways of incorporating semantic information and ways of using variable resolution need to be developed. An analogous technique for roof edges needs to be developed. We hope to explore these issues in future papers.

REFERENCES

Peter Beckmann, Orthogonal polynomials for Engineers and Physicists The Golem Press, Boulder, Colorado 1973.

Paul Beaudet "Rotationally Invariant Image Operators" 4th International Joint Conference on Pattern Recognition, Tokyo, Japan, November 1978, p579-583.

M.J. Brooks "Rationalizing Edge Detectors" Computer Graphics and Image Processing, Vol 8, 1978 p277-285.

Roger Ehrich and Fred Schroeder, "Contextual Boundary Formation by One-Dimensional Edge Detection and Scan Line Matching" Computer Graphics and Image Processing, Vol 16, 1981, p116-149.

Robert Foutz, Personal communication, 1981.

Robert Haralick, "Edge and Region Analysis For Digital Image Data" Computer Graphics and Image Processing, Vol. 12, 1980 p60-73.

Robert Haralick and Layne Watson, "A Facet Model for Image Data" Computer Graphics and Image Processing, Vol 15, 1981, p113-129.

M. Hueckel, "A Local Visual Operator Which Recognizes Edges and Lines" J. Assoc. Computing Machinery Vol 20, 1973, p634-647.

M. Hueckel, "An Operator Which Locates Edges in Digitized Pictures" J. Assoc. Computing Machinery, Vol 18, 1971, p113-125.

David Marr and Ellen Hildreth, "Theory of Edge Detection" Proc. Royal Society of London, B, vol 207, 1980, p187-217.

David Morgenthaler, "A New Hybrid Edge Detector" Computer Graphics and Image Processing, vol 16, 1981, p166-176.

David Morgenthaler and Azriel Rosenfeld, "Multidimentional Edge Section by Hypersurface Fitting" IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol PAMI-3, no 4, July 1981, p482-486.

Ramokaut Nevatia and Ramesh Babu, "Linear Feature Extraction and Description" Computer Graphics and Image Processing, Vol 13, 1980, p257-269.

Judith Prewitt, "Object Enhancement and Extraction" Picture Processing, and Psychopictories (B. Lipkin and A. Rosenfeld Ed.), Academic Press, New York, 1970, p75-149.

L. G. Roberts " Machine Perception of Three-Dimensional Solids" in Optical and Electrooptical Information Processing, J.T. Trippett et. al., Eds, MIT Press, Cambridge, Mass, 1965, p159-197

Azriel Rosenfeld and A. Kak Digital Picture Processing, Academic Press, New York, 1976.

Steve Zucker and Robert Hummel, "An Optimal Three-Dimensional Edge Operator" Pattern Recognition and Image Processing Conference, Chicago, August 1979, p162-168.