Zero crossing of second directional derivative edge operator

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

In this paper, we assume that in each neighborhood of the image the underlying grey tone intensity 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. 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 which is discussed in Haralick (1981).

In section II, we discuss how the discretely sampled data values are used to estimate the coefficients of the linear combinations of the tensor product polynomials: coefficient estimates for exactly fitting or estimates for least square fitting are calculated as linear combinations of the sampled data values. The masks used are shown in figures 1 and 2. Figure 3 illustrates that the order of the fit does make a difference in the mask used to estimate a low order coefficient.

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 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.

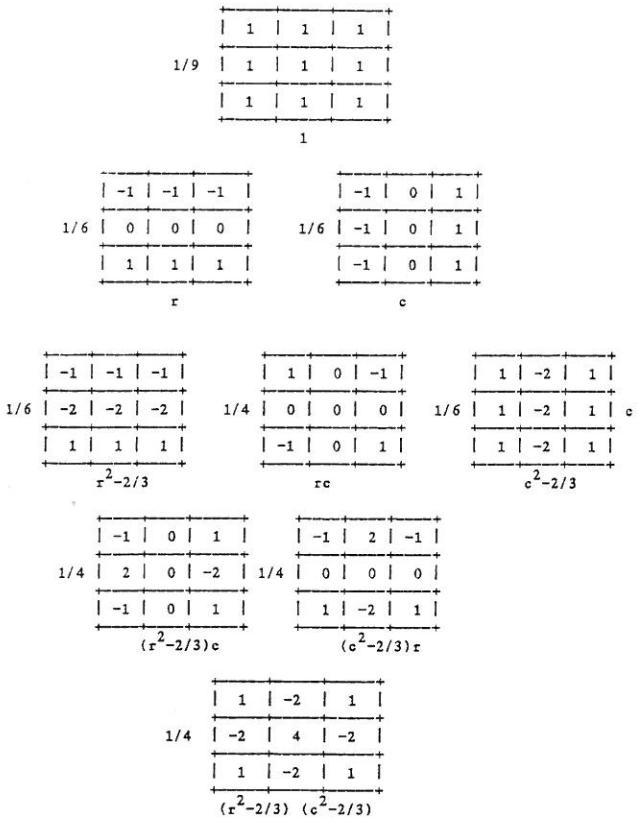

Figure 1 illustrates the 9 masks for the 3x3 window.

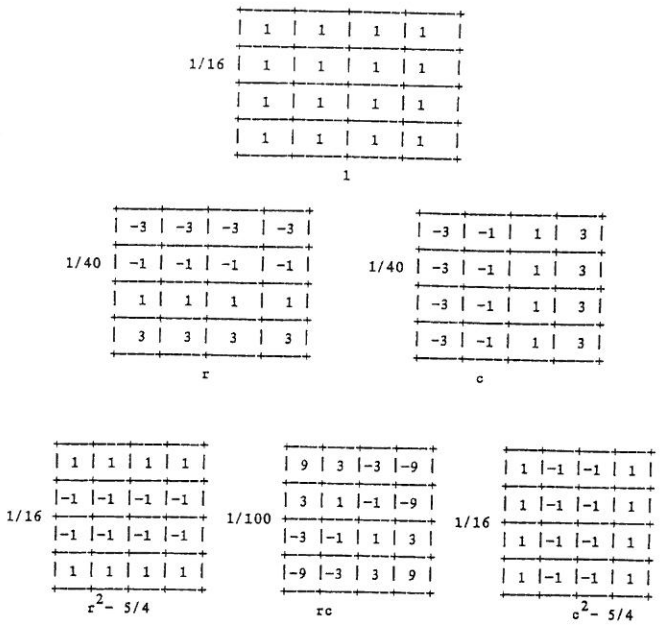

Figure 2 illustrates the masks used to obtain the coefficients of all polynomials up to the quadratic ones for a 4x4 window.

II. 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

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

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,

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

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

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

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

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

The exact fitting coefficients and the least squares coefficients are identical form = 0,...,K.

Fitting the data values $\{d(r) | r\varepsilon R\}$ 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.

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

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

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

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

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

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

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,

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

we obtain α by

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

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

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

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

$$\begin{aligned} f''_\alpha(\rho) &= 6[k_7\sin^3\alpha + k_8\sin^2\alpha\cos\alpha \\ &\quad + k_9\sin\alpha\cos^2\alpha + k_{10}\cos^3\alpha]\rho \\ &\quad + 2[k_4 \sin^2\alpha + k_5 \sin\alpha \cos\alpha + k_6 \cos^2\alpha] \\ &= A\rho + B \end{aligned} \qquad (10)$$

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. 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 4.

Section IV.1 illustrates the performance of the classic 3x3 edge operators with and without preaveraging compared against the generalized Prewitt operator. Section IV.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 5. Note that the usual definition of the Roberts operator has been modified in the natural way so that it uses a 3x3 mask.

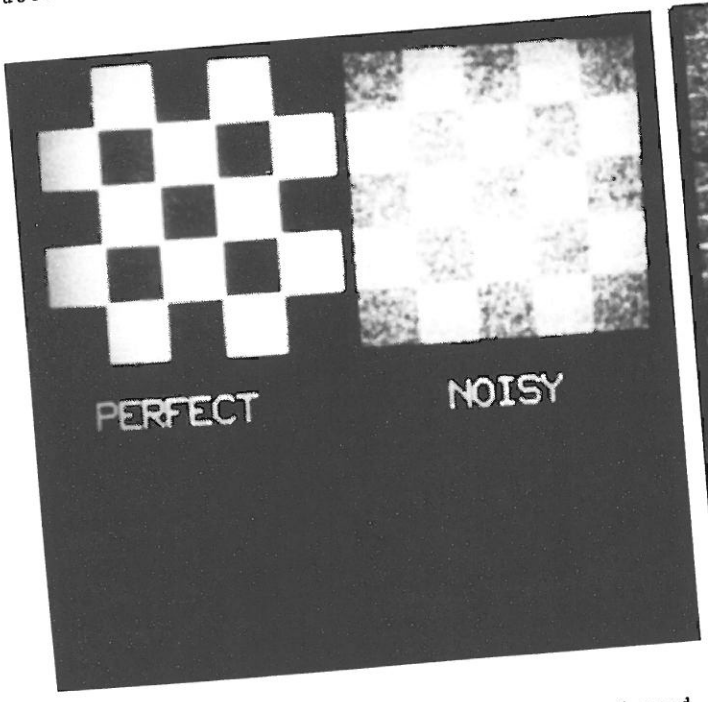

Figure 4 illustrates the noisy checkerboard used in the experiments. Low intensity is 75 high intensity is 175. Standard deviation of noise is 50.

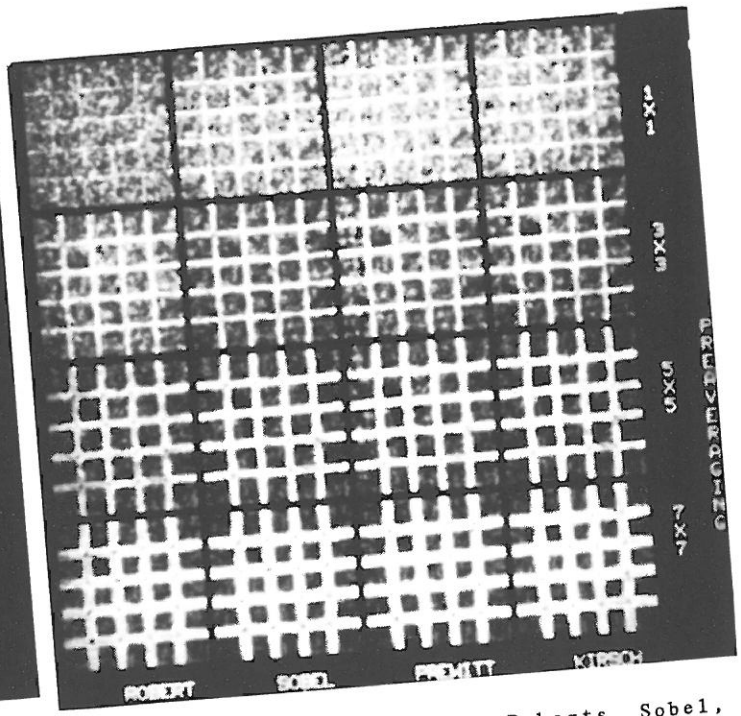

Figure 5 illustrates the 3x3 Roberts, Sobel, Prewitt, and Kirsch edge operators with a box filter preaveraging of 1x1, 3x3, 5x5, and 7x7.

Averaging before the application of the gradient operator is considered to be the cure for such bad performance on noisy images (Rosenfeld and Kak, 1976). Figure 5 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 6. A 3x3 pre-average followed by a 3x3 gradient operator yields a resulting neighborhood size of 5x5. Thus in figure 6 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 (1980) compass operator was compared with the 5x5 Prewitt. 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.

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 7 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.

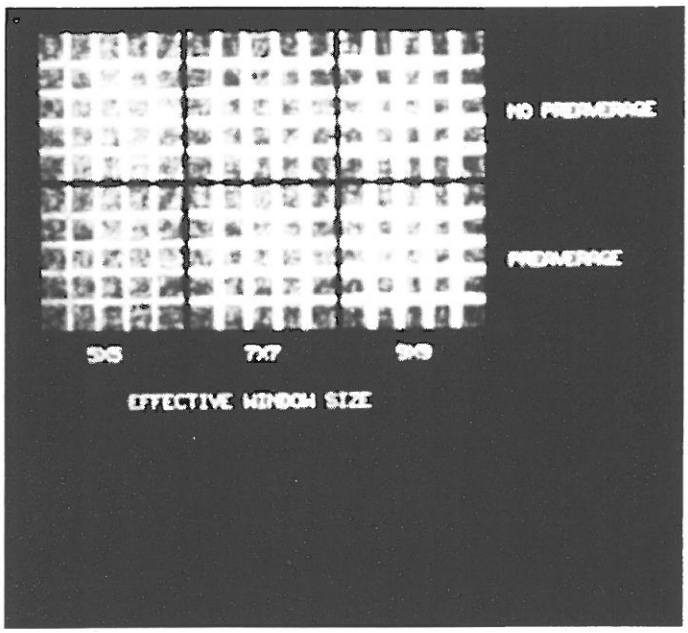

Figure 6 illustrates the Prewitt Operator done by using a least squares quadratic fit in the neighborhood versus doing preaveraging and using a small fitting neighborhood size. The no preaveraging and using a small fitting neighborhood size. The no preaveraging results show slightly contrast.

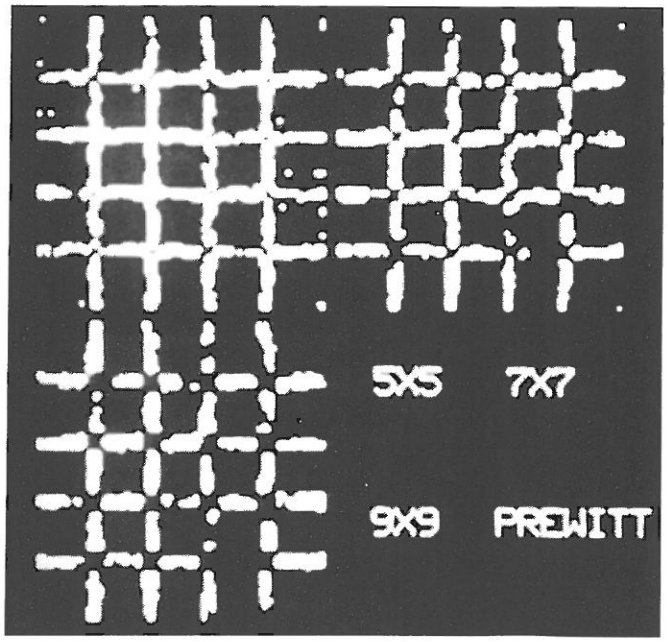

Figure 7 illustrates the edges obtained by thresholding the results of the Prewitt operator.

IV. 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 a 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 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 8 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 9 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 there being a zero gradient there.

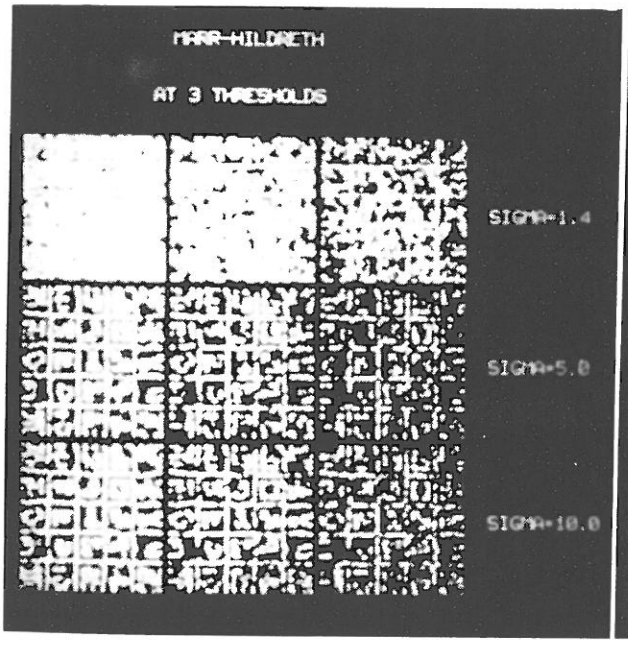

Figure 8 illustrates the edges obtained by the 11x11 Marr-Hildreth zero-crossing of Laplacian operator set for three different zero-crossing thresholds and three different standard deviations for the associated Mexican hat filter.

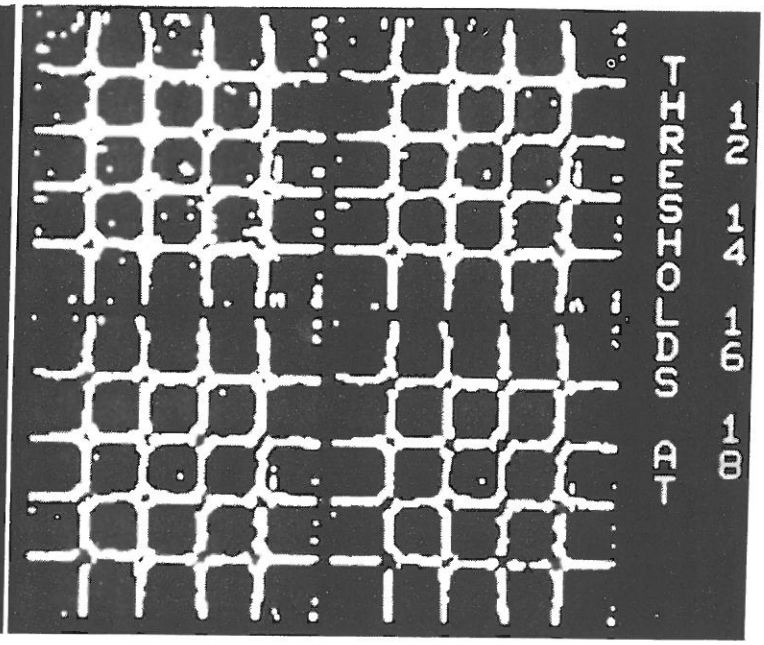

Figure 9 illustrates the directional derivative edge operator for 4 different thresholds.

Figure 10a illustrates an aerial photograph.

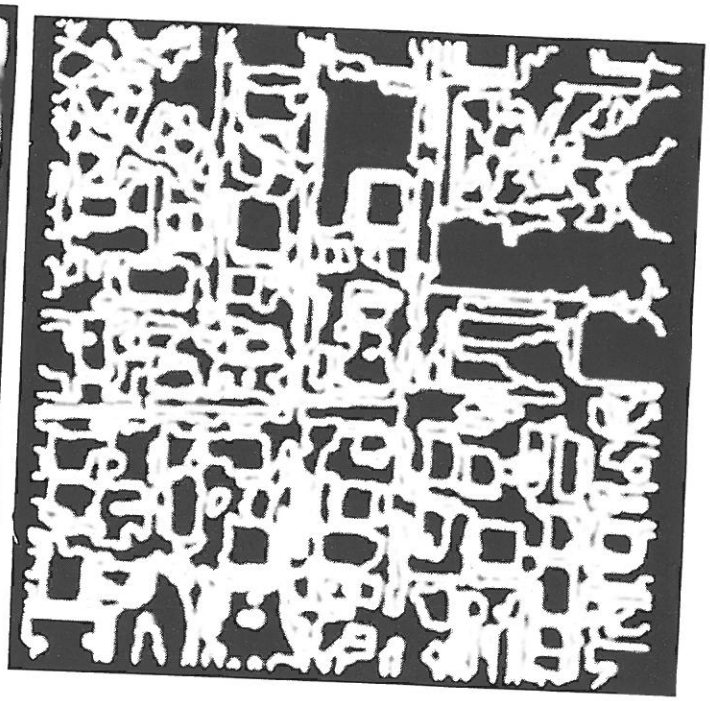

Figure 10b illustrates the directional derivative edges obtained from the aerial photograph by first 3x3 median filterning, then replacing each pixel by the closer of its 3x3 neighborhood minimum or maximum, then taking the directional derivative edges using a 7x7 window, then doing a connected components on the non-edge pixels, and removing all regions having fewer than 20 pixels, and then displaying any pixel neighboring an pixel different than it as an edge pixel.

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

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

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

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

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.

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