SECOND DIRECTIONAL DERIVATIVE ZERO-CROSSING DETECTOR USING THE CUBIC FACET MODEL

Robert M. Haralick

Machine Vision International
325 E. Eisenhower Parkway
Ann Arbor, Michigan 48104

ABSTRACT

The cubic polynomial is analyzed and its translation invariant parameters are derived. These translation invariant parameters are scale and contrast and are related to the horizontal and vertical distance between relative extrema of the cubic. The implementation details of the facet model second directional derivative zero-crossing edge detector described previously by Haralick are then given in terms of the translation invariant parameters.

A variety of results are shown for a noiseless sample function having different kinds of discontinuities. The least square facet parameters of the approximating cubic are calculated under different window sizes and different amounts of Gaussian preaveraging. The results indicate that when the contrast threshold is set slightly smaller than the contrast difference across the discontinuity, no preaveraging or a Gaussian preaverage of $\sigma = .6$ yield identical perfectly placed edges for all odd window sizes between 5 and 11. When the standard deviation is as high as $\sigma = 1.5$, some edges are not detected, but those which are detected are correctly placed. If the contrast threshold is set too low, some false edges are detected.

Finally a comparison of the zero-crossing of Laplacian, a popular Mexican hat edge detector, shows that regardless of the standard deviation of the Gaussian, zero crossing slope thresholds which are too small yield some falsely detected edges and zero crossing slope thresholds which are too large yield some misdetected edges and some incorrectly placed ones. Furthermore, in contrast to the facet edge detector, there is no zero-crossing slope threshold for the Mexican hat edge detector which can provide perfect edge detection on the given noiseless sample function.

By edge we mean a configuration of gray tone intensity values which on each side of the edge have relatively small variation in value and which across the edge have relatively large variation in value. An ideal edge of this kind is a step edge whose gray tone intensity values on each side of the edge take a different constant value.

The key idea in detecting edges is to look for relatively large contrasts in small distances. Change in value, or contrast, divided by change in location which causes the value change is the essence of what a first derivative is. A large contrast in a small distance means a large first derivative. If there were to be many continguous points with large enough first derivatives, the natural one to choose would be the one which has the largest first derivative. If the first derivative is to be a relative maximum, then the second derivative must be zero and the third derivative must be negative if the edge is crossed from the lower value to the high value gray tone region.

In the second directional derivative zero crossing edge detector (Haralick, 1984), bivariate cubic function is fit to the central neighborhood of pixel. The fit produces the estimated bivariate function f:

$$f(r,c) = k_1 + k_2 r + k_3 c + k_4 r^2 + k_5 rc + k_6 c + k_7 r^3$$
$$k_8 r^2 c + k_9 rc^2 + k_{10} c^3$$

Based on the estimated coefficients $k_1, \ldots, k_{10}$ a decision is made to label the pixel as edge or non-edge. A pixel is labelled as an edge if the second directional derivative, taken in the direction of the gradient, has a negatively sloped zero crossing located near the center of the pixel.

The simplest way to think about directional derivatives is to cut the surface f(r,c) with a plane which is oriented in the desired direction and which is orthogonal to the row-column plane. By convention, we take the angle to be measured clockwise from the column axis. We define the desired direction to be the gradient direction at the center of the given pixel. Hence, the gradient angle θ, satisfies

$$\sin\theta = k_2/(k_2^2 + k_3^2)^{.5}$$

$$\cos\theta = k_3/(k_2^2 + k_3^2)^{.5}$$

The angle θ is well defined providing that $k_2^2+k_3^2>0$.

To cut the surface f(r,c) with a plane in the direction θ we just require that $r = p\sin\theta$ and $c = p\cos\theta$ where p is the independent variable. This requirement produces the cubic curve $f_\theta(p)$.

$$f_\theta(p) = k_1 + (k_2\sin\theta + k_3\cos\theta)p + (k_4\sin^2\theta + k_5\sin\theta\cos\theta + k_6\cos^2\theta)p^2 + (k_7\sin^2\theta + k_8\sin^2\theta\cos\theta + k_9\sin\theta\cos\theta + k_{10}\cos^3\theta)p^3$$

CH2145-1/85/0000/0672$01.00©1985 IEEE

Let

$$C_0 = k_1$$
$$C_1 = k_2 \sin\theta + k_3 \cos\theta = (k_2^2 + k_3^2)^{.5}$$
$$C_2 = k_4 \sin^2\theta + k_5 \sin\theta\cos\theta + k_6 \cos^2\theta$$
$$C_3 = k_7 \sin^3\theta + k_8 \sin^2\theta\cos\theta + k_9 \sin\theta\cos^2\theta + k_{10}\cos^3\theta$$

Then $f_\theta(p) = C_0 + C_1 p + C_2 p^2 + C_3 p^3$ from which it follows that the first, second and third directional derivatives are given by

$$f_\theta'(p) = C_1 + 2C_2 p + 3C_3 p^2$$
$$f_\theta''(p) = 2C^2 + 6C_3 p$$
$$f_\theta'''(p) = 6C_3$$

For a pixel to be labeled as an edge pixel, the second directional derivative must have a negatively sloped zero crossing sufficiently near the center of the pixel. In this case, with the origin taken as the center of the pixel, there must be a p sufficiently small in magnitude satisfying

$f_\theta''(p) = 0$ (this is the zero requirement)

and $f_\theta'''(p) < 0$ (this is the negative slope requirement)

For $f_\theta'''(p) < 0$ we must determine that $C_3 < 0$. If $C_3 < 0$, then $C_3 \neq 0$ and a p having the value $-C_2/3C_3$ exists which makes $f_\theta''(p) = 0$. If $|C_2/3C_3| < p_0$, where we take p_0 to be somewhat less than a pixel length, we can label the pixel as an edge. In essence, this is the procedure given by Haralick (1984).

If our ideal edge is the step edge, then we can refine the above detection criteria by insisting that the cubic polynomial $f_\theta(p)$ have coefficients which make $f_\theta(p)$ a suitable polynomial approximation of the step edge. Now a step edge does not change in its essence if it is translated to the left or right or if it has a constant added to its height. Since the cubic polynomial is representing the step edge, we must determine what it is about the cubic polynomial which is its fundamental essence after an ordinate and abscissa translation.

To do this, we translate the cubic polynomial so that its inflection point is at the origin. Calling the new polynomial g, we have

$$g_\theta(p) = f_\theta(p - C_2/3C_3) - (C_0 + 2C_2^3/27C_3^2 - C_1C_2/3C_3)$$
$$= ((3C_1C_3 - C_2^2)/3C_3)p + C_3p^3$$

In our case since $C_1 = (k_2^2 + k_3^2)^{.5}$ we know $C_1 > 0$. If a pixel is to be an edge the second derivative zero crossing slope must be negative. Hence, for edge pixel candidates $C_3 < 0$. This makes $-3C_1C_3 + C_2^2 > 0$ which means that $g_\theta(p)$ has relative extrema. The parameters of the cubic which are invariant under translation relate to these relative extrema. The parameters are the distance between the relative extrema in the abscissa direction and in the ordinate direction.

We develop these invariants directly from the polynomial equation for $g_\theta(p)$. First we factor out the term

$$\frac{(C_2^2 - 3C_1C_3)^{1.5}}{3^{1.5}\, C_3^2}$$

This produces

$$g_\theta(p) = \left(\frac{(C_2^2 - 3C_1C_3)^{1.5}}{C^{1.5}C_3^2}\right)\left(\frac{(-3^{.5}C_3)\ p}{(C_2^2 - 3C_1C_3)^{.5}} + \frac{(3^{1.5}C_3^3)\ p^3}{(C_2^2 - 3C_1C_3)^{1.5}}\right)$$

For candidate edge pixels, $C_3 < 0$. This permits a rewrite to

$$g_\theta(p) = \left(\frac{(C_2^2 - 3C_1C_3)^{1.5}}{3^{1.5}C_3^2}\right)\left[\left(\frac{3C_3^2}{C_2^2 - 3C_1C_3}\right)^{.5} p - \left(\frac{3C_3^2}{C_2^2 - 3C_1C_3}\right)^{1.5} p^3\right]$$

Let the contrast be C and the scale be S. They are defined by

$$C = \frac{(C_2^2 - 3C_1C_3)^{1.5}}{3^{1.5}C_3^2}$$

$$S = \frac{(3C_3^2)^{.5}}{(C_2^2 - 3C_1C_3)^{.5}}$$

Finally, we have

$$g_\theta'(p) = C(Sp - S^3p^3)$$

In this form it is relatively easy to determine the character of the cubic. Differentiating,

$$g_\theta'(p) = C(S - 3S^3p^2)$$
$$g_\theta''(p) = 6CS^3p$$

The locations of the relative extrema only depend on S. They are located at $\pm 1/(3^{.5}S)$. The height difference between relative extrema depends only on the contrast. Their heights are $\pm 2C/(3^{1.5})$. Other characteristics of the cubic depend on both C and S. For example, the magnitude of the curvature at the extreme is $2(3^{.5})CS^2$ and the derivative at the inflection point is CS.

Of interest to us is the relationship between an ideal perfect step edge and the representation it has in the least squares approximating cubic whose essential parameters are contrast C and scale S. We take an ideal step edge centered in an odd neighborhood size N to have (N-2)/2 pixels with value -1, a center pixel with value 0, and (N-1)/2 pixels with value +1. Using neighborhood sizes of from 5 to 23 we find the following values for contrast C and scale S of the least squares approximating cubic.

Neighborhood Size N	Contrast C	Scale S
5	3.0867	.37796
7	3.1357	.26069
9	3.1566	.20000
11	3.1673	.16253
13	3.1734	.13699
15	3.1773	.11844
17	3.1799	.10434
19	3.1817	.09325
21	3.1830	.08430
23	3.1841	.076924

The average contrast of the approximating cubic is 3.16257. The scale S(N) appears to be inversely related to N; S(N) = S/N. The value of S minimizing the relative error

$$\left(\frac{S(N) - S/N}{S(N)}\right)$$

is 1.793157.

These two relationships

$$C = 3.16257$$
$$S = 1.793157/N$$

for ideal step edges having a contrast of 2 can help provide additional criteria for edge selection. For example the contrast across an arbitrary step edge can be estimated by

$$\text{Edge Contrast} = \frac{2C}{3.16257}$$

If the edge contrast is too small, then the pixel is rejected as an edge pixel. We have found that in many kinds of images, too small means smaller than 5 percent of the image's true dynamic range. Interestingly enough, edge contrast C depends on the three coefficients C_1, C_2, C_3 of the representing cubic. First derivative magnitude at the origin, a value used by many edge gradient magnitude detection techniques, only depends on the coefficient C_1. First derivative magnitude at the inflection point is precisely CS, a value which mixed both scale and edge contrast together.

The scale of the edge can be defined by

$$\text{Edge Scale} = \frac{SN}{1.793157}$$

Ideal step edges, regardless of their contrast, will produce least squares approximating cubic polynomials whose Edge Scale is very close to unity. Values of Edge Scale larger than one have the relative extrema of the representing cubic closer together than expected for an ideal step edge. Values of Edge Scale smaller than one have the relative extrema of the representing cubic further away from each other than expected for an ideal step edge. Values of Edge Scale which are significatnly different from unity may be indicative of a cubic representing a data value pattern very much different from a step edge. Candidate edge pixels with an edge scale very different from unity can be rejected as edge pixels.

The determination of how far from unity is different enough requires an understanding of what sorts of non-edge situations yield cubics with a high enough contrast and with an inflection point close enough to the neighborhood center. We have found that such non-edge situations occur when a step like jump occurs at the last point in the neighborhood. For example, suppose all the observed values are the same except the value at an endpoint. If N is the neighborhood size then the inflection point of the approximating cubic will occur at $\pm$ (N + 3)/14, the plus sign corresponding to a different left endpoint and the minus sign corresponding to a different right endpoint. Hence, for neighborhood sizes of N = 5, 7, 9, or 11 the inflection point occurs within a distance of 1 from the center point of the neighborhood. So providing the contrast is high enough, the situation would be classified as an edge if scale were ignored. For neighborhood sizes of N = 5, 7, 9, 11, and 13, however, the scale of the approximating cubic is 1.98, 1.81, 1.74. 1.71, and 1.68, respectively. This suggests that scales larger than 1 are significally more different from unity scale than corresponding scales smaller than 1. We have found that in many images restricting edge scale to be between .4 and 1.1 works well.

RESULTS

Figure 1 draws a sample data set which has three obvious step edges with jumps of 150, 100, and 100 respectively. We will preprocess this data set using no preaveraging and preaveraging with a Gaussian having a standard deviation of .6 and 1.5. We use fitting windows of 5, 7, 9, and 11 points and edge contrast thresholds of 30 and 75. These results are shown in Figures 2 and 3.

It is apparent from these results that as the amount of preaveraging increases, the tendency to lose an edge increases if the edge contrast threshold remains the same. As the edge contrast threshold increases, the tendency to eliminate false edges increases if the amount of preaveraging remains constant.

For all cases where the edge is marked correctly, the position of the edge is correct. Those edges which are two pixels wide have the right boundary point of the left segment marked. Those edges which are one pixel wide have only one of the boundary points from the left or right segment marked.

These results also suggest that for thresholds a small fraction below the edge jump value, little or no preaveraging gives a better result from a lot of preaveraging. This holds for all fitting window sizes tried. Thresholds which are a small fraction of the edge jump value have the chance of incorrectly assigning some inflection points as edges. This tendeny can be mitigated somewhat by a large amount of preaveraging.

Finally, a comparison is made with the zero-crossing of Laplacian edge detector. The Mexican hat kernel is given by sampling the second derivative of a Gaussian having standard deviation. The support window for the kernel is large enough so that the magnitude of the value of the kernel on the boundary is 1/1000 of the value in the center. Any pixel where the magnitude of the difference between itself and a neighboring pixel of different sign is greater than a specified threshold is marked as an edge. We ran experiments for standard deviations of &sin.= .6, 1.0, 1.5, 2.5, and 5.0 and zero-crossing slope thresholds of 1, 10, and 20. The results shown in Figure 4 indicate that thresholds which are too low yield some falsely detected edges. Thresholds which are too large yield some misdetected edges and falsely placed edges. We tried all values of threshold between what was too small and what was too large and there was no threshold for all standard deviations which produced perfect edge detection.

CONCLUSIONS

In the one dimensional example we illustrated, the first difference between the facet edge detector and the Mexican hat edge detector is the way derivatives are estimated. The facet model uses a least squares estimate and produces estimates which are evidently more stable or robust than those produced by the Mexican hat filter. The second difference is that the facet model recognizes that the derivatives are estimated based on a model and that model must be taken into account in the processing. Hence, if the model is a cubic polynomial, the discontinuities must be understood through the eyes of the cubic polynomial. The implementation of the facet model recognized this and interprets discontinuities of step edges through the scale and contrast parameters of the cubic. On the other hand, there is no model of derivative estimation behind the Mexican hat edge detector. Finally we showed that even with the facet edge detector, preaveraging with a Gaussian filter with standard deviation just larger than one pixel width can yield misdetections. These results are similar to those of Leclerc and Zucker (1984). Standard deviations smaller than one pixel width do not adversely affect results.

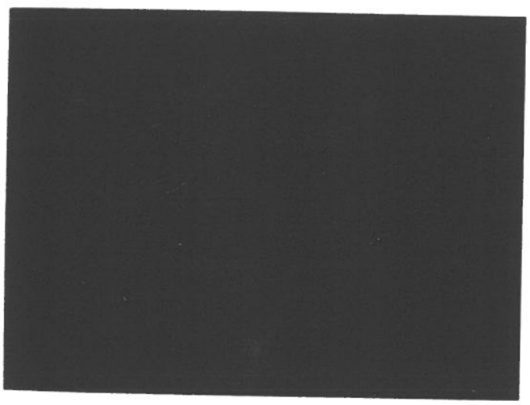
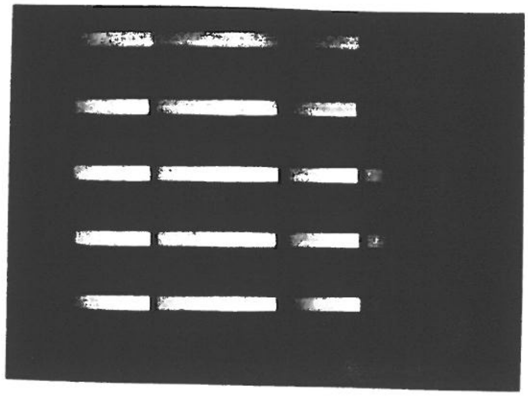
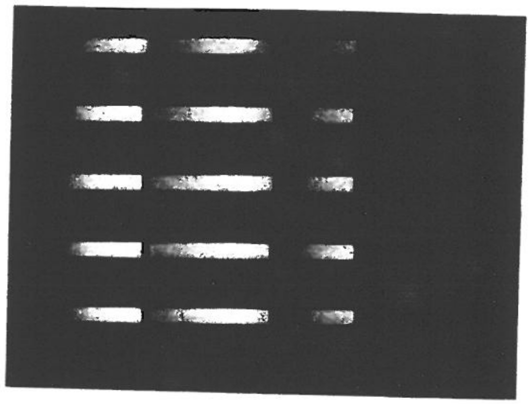


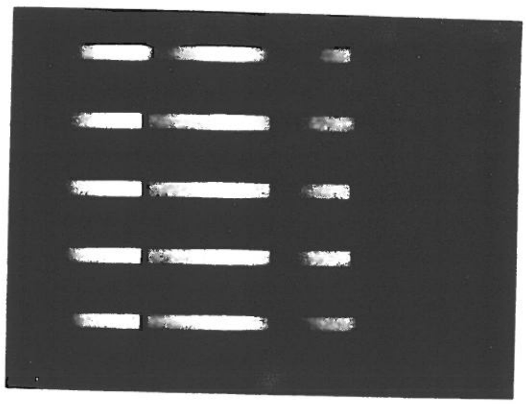
Figure 1. Figure 1 shows the point plot of the original data set.



(A)

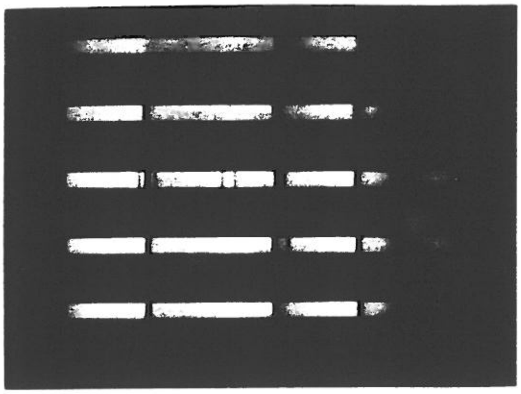


(B)

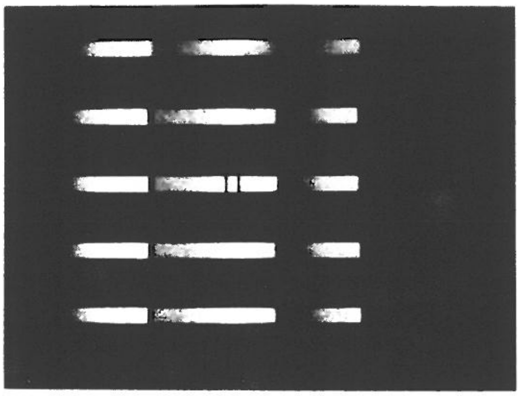


(C)

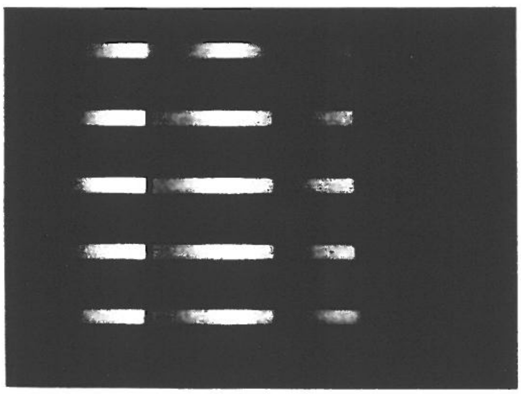
Figure 2. Figure 2 (A) shows processing of the data set using no initial preaveraging, (B) a Gaussian preaverage with σ = .6, and a Gaussian preaverage with a σ = 1.5 (C). The edge contrast threshold is 75 for fitting windows of 5, 7, 9, and 11 pixels wide.



(A)

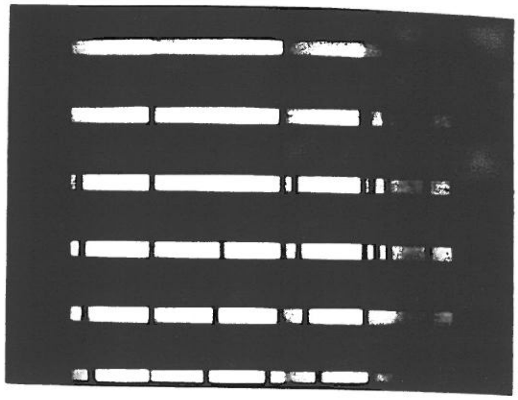


(B)

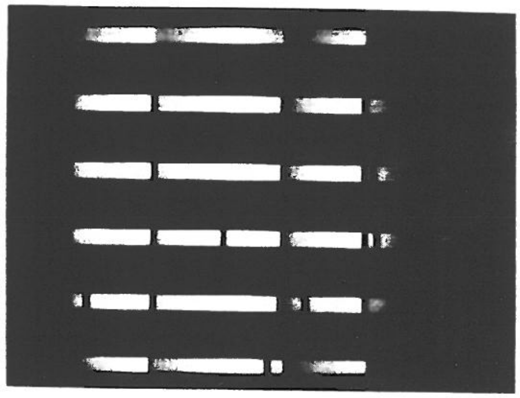


(C)

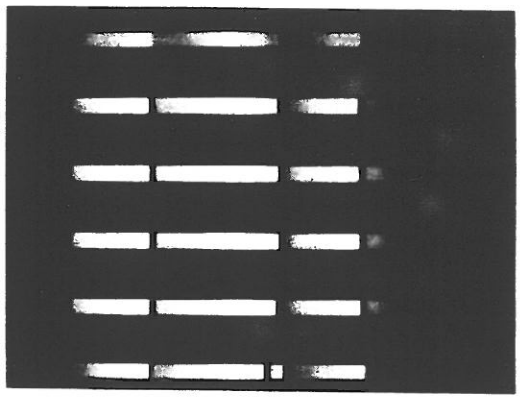
Figure 3. Figure 3 shows processing of the data set using no initial preaveraging (A), a Gaussian preaverage with σ = .6, (B) and a Gaussian preaverage with a σ = 1.5. (C). The edge contrast threshold is 30 for fitting window of 5, 7, 9, and 11 pixels wide.



(A)



(B)



(C)

Figure 4. Figure 4 shows the zero crossing of Laplacian operator with zero crossing slope threshold of 2 (A), 20 (B), and 40 (C) for Gaussian presmoother having standard deviation of .6, 1.0, 1.5, 2.5, and 5.0.

REFERENCES

Haralick, Robert M., "Digital Step Edges from Zero Crossings of Second Directional Derivative", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-6, No. 1, January 1984, pp. 58-68.

Leclerc, Yvan and Steven Zucker, "The Local Structure of Image Discontinuities in One Dimension", Seventh International Conference on Pattern Recognition, Montreal, Canada, July 30 - August 2, 1984, pp. 46-48.