PATTERN RECOGNITION IN PRACTICE II
E.S. Gelsema and L.N. Kanal (Editors)

CUBIC FACET MODEL EDGE DETECTOR AND RIDGE-VALLEY DETECTOR: IMPLEMENTATION DETAILS

Robert M. Haralick

Machine Vision International
325 E. Eisenhower Parkway
Ann Arbor, Michigan
U.S.A.

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 and facet model ridge and valley detector previously described by Haralick are then given in terms of these translation invariant parameters.

INTRODUCTION

In cubic facet model processing for edges and ridges and valleys, a bivariate cubic function is fit to the central neighborhood of each pixel. In a coordinate system whose origin is the center of the pixel, the fit produces the bivariate cubic function f which is an estimate of the underlying and unobserved function:

$$f(r,c) = k_1 + k_2r + k_3c + k_4r^2 + k_5rc + k_6c^2 + k_7r^3 + k_8r^2c + k_9rc^2 + k_{10}c^3$$

Based on the estimated coefficients $k_1,\ldots,k_{10}$ a decision is made to label the pixel as edge or non-edge, or ridge or valley or non-ridge and valley. 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. A pixel is labelled as a ridge or valley if the first directional image, taken in a direction which extremizes the second directional derivative, has a zero crossing located sufficiently near the center of the pixel.

The straightforward implementation of these definitions in terms of the fitting coefficients $k_1,\ldots,k_{10}$ can give rise to unnecessary artifacts. The purpose of this paper is to illustrate how to apply these definitions through the ''eyes'' of the fitted cubic. Section II discusses the edge detection case. Section III discusses the ridge valley detection case.

II THE FACET EDGE DETECTOR

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 enough first derivative. The natural one to choose would be the one which has 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), a bivariate cubic function is fit to the central neighborhood of pixel. The fit produces the estimated bivariate cubic function f:

$$f(r,c) = k_1 + k_2r + k_3c + k_4r^2 + k_5rc + k_6c^2 + k_7r^3 + k_8r^2c + k_9rc^2 + k_{10}c^3 \quad (1)$$

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 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

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

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 $f_\Theta(p)$.

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

Let

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

Then

$$f_\Theta(p) = C_0 + C_1p + C_2p^2 + C_3p^3 \quad (4)$$

from which it follows that the first, second and third directional derivatives are given by

$$\begin{aligned} f_\Theta'(p) &= C_1 + 2C_2p + 3C_3p^2 \\ f_\Theta''(p) &= 2C_2 + 6C_3p \\ f_\Theta'''(p) &= 6C_3 \end{aligned} \quad (5)$$

For a pixel to be labelled 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

$$f_\theta''(p) = 0 \quad \text{(this is the zero requirement)}$$

$$\text{and} \quad f_\theta'''(p) < 0 \quad \text{(this is the negative slope requirement).}$$

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_θ 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

$$\begin{aligned} g_\theta(p) &= f_\theta(p - C_2/3C_3) - (C_0 + 2C_2^3/27C_3^2 - C_1C_2/3C_3) \\ &= ((3C_1C_3 - C_2^2)/3C_3)p + C_3p^3 \end{aligned} \qquad (6)$$

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 $C_2 < 0$. 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

$$\begin{aligned} g_\theta(p) = {} & [(C_2^2 - 3C_1C_3)^{1.5}/3^{1.5}C_3^2][-3^{.5}C_3/(C_2^2 - 3C_1C_3)^{.5})p \\ & + (3^{1.5}C_3^3/(C_2^2 - 3C_1C_3)^{1.5}p^3] \end{aligned}$$

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

$$\begin{aligned} g_\theta(p) = {} & [C_2^2 - 3C_1C_3)^{1.5}/3^{1.5}C_3^2] \quad [(3C_3^2/C_2^2 - 3C_1C_3))]^{.5}p \\ & - (3C_3^2/(C_2^2 - 3C_1C_3))^{1.5}p^3 \end{aligned}$$

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

$$C = \frac{(C_2^2 - 3C_1C_3)^{1.5}}{3^{1.5}C_3^2} \qquad (7)$$

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

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,

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

$$g_\theta''(p) = 6CS^3p \qquad (8)$$

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-1)/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

$$\text{Edge Contrast} = \frac{2C}{3.16257} \tag{9}$$

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 mixes both scale and edge contrast together.

The scale of the edge can be defined by

$$\text{Edge Scale} = \frac{SN}{1.793157} \tag{10}$$

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 significantly 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.

III THE FACET RIDGE VALLEY OPERATOR

In the facet ridge valley detector (Haralick, 1983), a bivariate cubic function is fit to the central neighborhood of a pixel. The fit produces the estimated bivariate function f:

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

Based on the estimated coefficients $k_1,\ldots,k_{10}$ a decision is made to label the pixel as ridge of valley. A pixel is labelled as a ridge or valley if the first directional derviative, taken in a direction which extremizes the second directional derivative, has a zero crossing located sufficiently near the center of the pixel.

To determine a direction θ extremizing the second directional derivative, we proceed as before and cut the surface f(r,c) with a plane in the direction θ. Letting $r = p \sin\theta$ and $c = p\cos\theta$, we obtain the cubic curve $f_\theta(p)$ defined by equation 4. Using the definition for C_0, C_1, C_2, and C_3 from equation 3 and equation 4 expressing $f_\theta(p)$ a cubic polynomial having coefficients C_0, C_1, C_2, and C_3, we readily obtain that the second directional derivative in direction θ evaluated at the origin is given by

$$f_\theta''(0) = 2C_2 = (\sin\theta \ \cos\theta)\begin{pmatrix} 2k_4 & k_5 \\ k_5 & 2k_6 \end{pmatrix}\begin{pmatrix} \sin\theta \\ \cos\theta \end{pmatrix} \tag{11}$$

The directions extremizing $f_\theta''(0)$ are precisely the two orthogonal directions determined by the eigenvectors of

$$H = \begin{pmatrix} 2k_4 & k_5 \\ k_5 & 2k_6 \end{pmatrix} \tag{12}$$

which, in fact, is the Hessian of f, the matrix of the second partial derivatives of f. The eigenvalues of H are given by

$$\lambda = k_4 + k_6 \pm \sqrt{(k_6 - k_4)^2 + k_5^2} \tag{13}$$

Each eigenvector $(x,y)'$ of H must satisfy

$$2k_4 x + k_5 y = [(k_4 + k_6) \pm \sqrt{(k_6 - k_4)^2 + k_5^2}\,]x \tag{14}$$

Either $k_5 = 0$ or not. If $k_5 = 0$, then the eigenvectors of H are

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

If $k_5 \neq 0$, then equation 14 can be solved for y producing

$$y = \left(\frac{k_6 - k_4}{k_5}\right) + \left[\sqrt{\left(\frac{k_6 - k_4}{k_5}\right)^2 + 1}\right] x \tag{15}$$

The most numerically stable way of determining the eigenvectors is then to define

$$T = \text{Sign}\left(\frac{k_6 - k_4}{k_5}\right) \left[\left|\frac{k_6 - k_4}{k_5}\right| + \sqrt{\left(\frac{k_6 - k_4}{k_5}\right)^2 + 1}\right] \tag{16}$$

from which we obtain for eigenvectors

$$\frac{1}{\sqrt{1 + T^2}} \begin{pmatrix} 1 \\ T \end{pmatrix} \qquad \frac{1}{\sqrt{1 + T^2}} \begin{pmatrix} -T \\ 1 \end{pmatrix} \tag{17}$$

The extremizing directions θ_1 and θ_2 are given by

$$\sin \theta_1 = \frac{1}{\sqrt{1 + T^2}} \qquad \sin \theta_2 = \frac{-T}{\sqrt{1 + T^2}}$$
$$\cos \theta_1 = \frac{T}{\sqrt{1 + T^2}} \qquad \cos \theta_2 = \frac{1}{\sqrt{1 + T^2}} \tag{18}$$

Having determined an extremizing direction, the cubic $f_\theta(p)$ must be analyzed to determine where its extrema are. The extrema are located at the zero crossings of $f'_\theta(p)$. If there is a value of p^* sufficiently close to 0 which makes

$$f'_\theta(p^*) = C_1 + 2C_2p^* + 2C_3p^{*2} = 0$$

and for which $f''_\theta(p^*) = 2C_2 + 6C_3p^{*2} \neq 0$, then the central pixel of the fitting neighborhood can be classified as a peak or pit in direction θ. It is a peak in direction θ if $f''_\theta(p^*) < 0$ and a pit in direction θ if $f''_\theta(p^*) > 0$. The central pixel is classified as a ridge if in one of the extremizing directions it is a dominant peak and in the other extremizing direction it is neither a peak or pit or if it is a peak or pit it is a relatively weak one. It is classified as a valley if in one of the extremizing directions it is a dominant pit and in the other extremizing direction it is neither a peak or a pit or if it is a peak or a pit it is a relatively weak one. Relative strength of pit or peak can be measured by the ratio of $\min\{|f_1''|, |f_2''|\}$ to $\max\{|f_1''|, |f_2''|\}$. When this ratio is close to 0 one extrema dominates the other.

To determine a value of p^*, if any, which makes $f_\theta(p^*) = C_1 + 2C_2p^* + C_3p^{*2} = 0$, there are three cases to consider. If $C_3 = 0$ and $C_2 = 0$, then $f_\theta'' = 0$ and there are no extrema. If $C_3 = 0$ and $C_2 \neq 0$, f_θ is quadratic and has one extrema located at $p^* = -C_1/2C_2$. If $C_3 \neq 0$, f_θ is cubic and has extrema only if

$C_2^2 - 3C_1C_3 > 0$. If it has extrema, the most numerically stable way of computing their location is by

$$p^*_{large} = \text{Sign}\ (C_2)(|C_2| + \ C_2^2 - 3C_1C_3\ ] \tag{19}$$

$$p_*^{small} = \frac{C_1}{3C^3 p^*_{large}} \tag{20}$$

The extrema closest to origin is located at p^*_{small}. If $|p^*_{small}| < $ radius, then a peak or pit in direction θ can be declared for the pixel.

However, there are some complications to this basic procedure. The first complications revolve around the fact that the fitted f_θ is, in general, a cubic. The cubics of interest have extrema. Such a fitted cubic arising from data which is as simple as piecewise constant with one jump have relative extrema even though the data does not. This is an artifact of using a cubic fit. It may or may not be significant; that is, they may or may not correspond to extrema in the data. To illustrate, consider a one-dimensional data pattern which is a constant with a jump change at one end. We analyze it in terms of the fitted cubic dynamic range, the relative depth of its extrema, and its inflection point.

The dynamic range of the cubic segment is defined by

$$\text{range} = \max_p f_\theta(p) - \min_p f_\theta(p)$$

The relative depth is defined as follows. Suppose the extrema is a relative minimum. Let a represent the left end point of the interval, b represent the right end point of the interval, i represent the location of the inflection point and v represent the location of the relative minimum. If the relative minimum occurs to the left of the relative maximum, the depth of the minimum is defined by

$$\text{depth} = \min\ \{f_\theta(a),\ f_\theta(i)\} - f_\theta(v)$$

If the relative minimum occurs to the right of the relative maximum, the depth of the minimum is defined by

$$\text{depth} = \min\ \{f_\theta(b),\ f_\theta(i)\} - f_\theta(v)$$

The relative depth of the minimum is then defined by

$$\text{relative depth} = \text{depth/range}$$

Data	Position of Extrema Closest to Origin	Position of Inflection Point	Relative Depth
00009	.241	-.571	.0408
0000009	.533	-.714	.0610
000000009	.811	-.857	.0731
000000099	-.755	-2.14	.0205
00000000099	.0237	-1.75	.0338
0000000000099	.519	-1.71	.0474

In each of these cases the cubic fitted to the data ''rings'' around the large constant region where the data is zero. One extrema of this ringing is close to the origin. This extrema is an artifact. Notice that in each of these cases, the inflection point is not too far from the origin indicating a relatively high frequency ringing. Also for all these cases, the relative depth, which measures the significance of the ringing, is close to zero.

Compare this situation to the case where the data clearly has an extrema near the origin as in the following one-dimensional patterns.

Data	Position of Extrema Closest to Origin	Position of Inflection Point	Relative Depth
01300	-.527	1.00	.279
0013000	-.569	2.14	.419
000130000	-.589	3.67	.604
00001300000	-.600	5.57	.821
0000013000000	-.607	7.86	.849
01900	-.235	2.71	.803
0019000	-.243	5.57	.874
000190000	-.246	9.38	.906
00001900000	-.247	14.14	.925
0000019000000	-.248	19.85	.937

Notice that in these cases the inflection points tend to be much further away from the extrema indicating the ''ringing'' behavior has a lower frequency. Also the relative depth tends to be much larger in the case of a true extrema.

What all this means is that in using f_θ to evaluate derivatives, we must understand and interpret the data through the ''eyes'' of a cubic polynomial. Not all extrema of the cubic are significant. A cubic extrema is significant in the sense that it reflects an extrema in the fitted data only if it passes the following tests

(1) |position of extrema from origin| < radius threshold

(2) |position of inflection from origin| > distance

(3) |distance between roots| > 1.756* Size of Interval

(4) relative depth > .2

(5) $|f_\theta''(p*)|$ > curvature threshold

Test (1) guarantees that the extrema is close enough to the origin. Tests (2) and (3) guarantee that the ''ringing'' behavior has a long enough period, test (3) taking into account that for true extrema, the period increases with the size of the fitting interval. Test (4) guarantees that the relative extrema have a significant enough height compared to the dynamic range of the fitted cubic segment. Test (5) guarantees that the curvature at the extrema is sufficiently high.

REFERENCES

[1] Haralick, R.M., ''Ridges and Valleys on Digital Images'', Computer Vision, Graphics, and Image Processing, Vol. 22, (1983), pp. 28-38.

[2] Haralick, R.M., ''Digital Step Edges from Zero Crossing of Second Directional Derivatives'', IEEE Transactions on on Pattern Analysis and Machine Intelligence, Vol. PAMI-5, No. 1, January 1984, p. 58-68.

DISCUSSION:

Smeulders:
Is it necessary that the cubic spline goes right through the pixel point?

Haralick:
Yes, the estimation is done for each neighbourhood independently. These neighbourhoods are highly overlapping and the places where the grey values originated from are the centers of the pixels.

Smeulders:
Is that not in conflict with your initial assumption that you are considering noisy digitized images. Could you not take a looser sort of fit, for instance the least squares fit?

Haralick:
Well, the fit is a least squares fit. If you do not want to assume that the points come on a regular grid you increase the computations by a tremendous amount.

Choudry:
Can't you use something different from cubics.

Haralick:
The reason for the cubic model is that it is of just one higher degree of complexity than absolutely necessary to solve the problem. That extra degree actually gave me a specific location in the pixel which was not necessarily the center of the pixel. Suppose you use a higher order model. Then, in the case of the edge, instead of trying to solve a linear equation you are going to ask me to solve a quadratic equation. The complexity goes up. And if you ask me to change the basis into a spline basis or a discrete cosine basis, then again to compute the zero crossings is computationally more complex.