CMPT 412

Robot Vision

Berthold Klaus Paul Horn

The MIT Press
Cambridge, Massachusetts London, England

McGraw-Hill Book Company
New York St. Louis San Francisco Montreal Toronto

8

Edges & Edge Finding

In this chapter we discuss edge detection and localization. Edges are curves in the image where rapid changes occur in brightness or in the spatial derivatives of brightness. The changes in brightness that we are particularly interested in are the ones that mirror significant events on the surface being imaged. These might be places where surface orientation changes discontinuously, where one object occludes another, where a cast shadow line appears, or where there is a discontinuity in surface reflectance properties. In each case, we hope to locate the discontinuity in image brightness, or its derivatives, in order to learn something about the corresponding feature on the object being imaged. We show in this chapter how differential operators can be used to accentuate those image features that help us locate places in the image where a fragment of an edge can be found. This is done first in the continuous domain, and then the results are applied to the discrete case.

Naturally, noise in the brightness measurements limits our ability to uncover edge information. We find trade-offs between sensitivity and accuracy and discover that to be detectable, short edges must have higher contrast than long edges. Edge detection can be considered complementary to image segmentation, since edges can be used to break up images into regions that correspond to different surfaces.

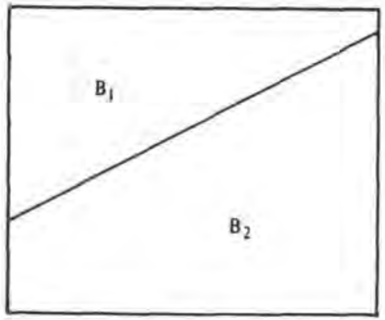

Figure 8-1. An idealized edge is a line separating two image regions of constant brightness.

8.1 Edges in Images

Intuitively, a simple edge is a border between two regions, each of which has approximately uniform brightness. Edges in images often result from *occluding contours* of objects. In this case the two regions are images of two different surfaces. Edges also arise from discontinuities in surface orientation and from discontinuities in surface reflectance properties. If we take a cross section of the image brightness along a line at right angles to an edge, we might hope to see a step discontinuity. In practice, the transition will not be abrupt because of blurring and limitations of the imaging device. Also, some edge transitions are better modeled as step transitions in the first derivative of brightness, rather than in the brightness itself.

For now, we shall use the simple model to gain some insight into operators that might enhance edges in images by producing high values near them. The edge-enhanced images must be processed further to extract lines or curves. Most effort has so far been concentrated on the edge-enhancement step, less on detection and localization of edge fragments, but this has changed recently. Least work has been done on the organization of the edge fragments into larger entities, that is, lines and curves in the image.

8.2 Differential Operators

A simple model for an edge in an image is a straight line separating two regions of contrast brightness (figure 8-1). We use the unit step function $u(t)$ defined by

$$u(z) = \begin{cases} 1, & \text{for } z > 0; \\ 1/2, & \text{for } z = 0; \\ 0, & \text{for } z < 0, \end{cases}$$

noting that it is the integral of the one-dimensional unit impulse

$$u(z) = \int_{-\infty}^{z} \delta(t)\, dt.$$

As on p117

Suppose that the edge lies along the line

$$x \sin\theta - y\cos\theta + \rho = 0.$$

As on p50–51.

Then we can write the image brightness in the form

$$E(x,y) = B_1 + (B_2 - B_1)\, u(x\sin\theta - y\cos\theta + \rho).$$

The partial derivatives are

Brightness change. i.e. Are you on the edge.

Orientation dependence crossing the edge maximizes.

$$\frac{\partial E}{\partial x} = +\sin\theta\,(B_2 - B_1)\,\delta(x\sin\theta - y\cos\theta + \rho),$$

$$\frac{\partial E}{\partial y} = -\cos\theta\,(B_2 - B_1)\,\delta(x\sin\theta - y\cos\theta + \rho).$$

These differential operators are directional, producing results that depend on the orientation of the edge. The vector $(\partial E/\partial x, \partial E/\partial y)^T$ is called the *brightness gradient*. The brightness gradient is *coordinate-system-independent* in the sense that it maintains its magnitude and orientation relative to the underlying pattern when the pattern is rotated or translated.

Consider now the squared gradient,

$$\left(\frac{\partial E}{\partial x}\right)^2 + \left(\frac{\partial E}{\partial y}\right)^2 = \big((B_2 - B_1)\,\delta(x\sin\theta - y\cos\theta + \rho)\big)^2.$$

This operator, while nonlinear, is rotationally symmetric, treating edges at all angles equally.

The derivative of the unit impulse is called the unit *doublet*, denoted δ'. Using this notation, we have

(see p…

$$\frac{\partial^2 E}{\partial x^2} = \sin^2\theta\,(B_2 - B_1)\,\delta'(x\sin\theta - y\cos\theta + \rho),$$

$$\frac{\partial^2 E}{\partial x \partial y} = -\sin\theta\cos\theta\,(B_2 - B_1)\,\delta'(x\sin\theta - y\cos\theta + \rho),$$

$$\frac{\partial^2 E}{\partial y^2} = \cos^2\theta\,(B_2 - B_1)\,\delta'(x\sin\theta - y\cos\theta + \rho).$$

The Laplacian of the image $E(x,y)$ is

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} = (B_2 - B_1)\,\delta'(x\sin\theta - y\cos\theta + \rho),$$

a quantity that is also rotationally symmetric. Finally, the *quadratic variation*,

$$\left(\frac{\partial^2 E}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 E}{\partial x \partial y}\right)\left(\frac{\partial^2 E}{\partial y \partial x}\right) + \left(\frac{\partial^2 E}{\partial y^2}\right)^2$$
$$= \left((B_2 - B_1)\,\delta'(x\sin\theta - y\cos\theta + \rho)\right)^2,$$

is rotationally symmetric, too, as expected. In the case of our idealized edge image, the quadratic variation happens to be equal to the square of the Laplacian. Note that of the three rotationally symmetric operators considered, only the Laplacian retains the sign of the brightness difference across the edge, $B_2 - B_1$. This allows us to determine which side of the edge is brighter from the edge-enhanced image. Thus the Laplacian is the only one of the three operators that might permit reconstruction of the original image from the edge image. It is also the only one of the three that is linear.

8.3 Discrete Approximations

Consider a 2×2 group of picture cells:

$E_{i,j+1}$	$E_{i+1,j+1}$
$E_{i,j}$	$E_{i+1,j}$

The derivatives at the center of this group can be estimated as follows:

$$\frac{\partial E}{\partial x} \approx \frac{1}{2\epsilon}\left((E_{i+1,j+1} - E_{i,j+1}) + (E_{i+1,j} - E_{i,j})\right),$$
$$\frac{\partial E}{\partial y} \approx \frac{1}{2\epsilon}\left((E_{i+1,j+1} - E_{i+1,j}) + (E_{i,j+1} - E_{i,j})\right),$$

where ϵ is the spacing between picture cell centers. Each estimate is the average of two finite-difference approximations.

A finite-difference approximation of a derivative always applies to a particular point. For example, the familiar first-difference formula gives an estimate that is unbiased for a point midway between the two places where the function is evaluated. The formulae shown above for estimating the partial derivatives are used because they are unbiased at the same point, namely the corner in the middle of the four picture cells. The squared

gradient can now be approximated by

$$\left(\frac{\partial E}{\partial x}\right)^2 + \left(\frac{\partial E}{\partial y}\right)^2 \approx \left((E_{i+1,j+1} - E_{i,j})^2 + (E_{i,j+1} - E_{i+1,j})^2\right).$$

If we perform this simple computation all over the image, we obtain high values at places where the brightness changes rapidly. In regions of constant brightness the output is zero. (If noise is present, the output is nonzero, but fairly small.) The results can be written into a new image array, in which the edges will be strongly highlighted.

The squared gradient does not tell us anything about the direction of the edge. This information is in the gradient itself, which points in the direction of most rapid increase in brightness. The edge is at right angles to the gradient since

$$\left(\frac{\partial E}{\partial x}, \frac{\partial E}{\partial y}\right)^T = (\sin\theta, -\cos\theta)^T\ (B_2 - B_1)\,\delta(x\sin\theta - y\cos\theta + \rho),$$

so that $\partial E/\partial x$ is proportional to $\sin\theta$ and $\partial E/\partial y$ is proportional to $-\cos\theta$. The gradient will point in a direction at right angles to the edge even if the edge transition is more gradual, perhaps due to blurring. Naturally, a discrete approximation of the gradient may not produce very accurate estimates of edge direction, since the picture cells through which the edges pass have intermediate values of brightness.

Now consider a 3×3 group of picture cells:

$E_{i-1,j+1}$	$E_{i,j+1}$	$E_{i+1,j+1}$
$E_{i-1,j}$	$E_{i,j}$	$E_{i+1,j}$
$E_{i-1,j-1}$	$E_{i,j-1}$	$E_{i+1,j-1}$

To estimate the Laplacian at the center cell we use the approximations

$$\frac{\partial^2 E}{\partial x^2} \approx \frac{1}{\epsilon^2}\left(E_{i-1,j} - 2E_{i,j} + E_{i+1,j}\right),$$
$$\frac{\partial^2 E}{\partial y^2} \approx \frac{1}{\epsilon^2}\left(E_{i,j-1} - 2E_{i,j} + E_{i,j+1}\right),$$

so that

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} \approx \frac{4}{\epsilon^2}\left(\frac{1}{4}(E_{i-1,j} + E_{i,j-1} + E_{i+1,j} + E_{i,j+1}) - E_{i,j}\right).$$

Here we subtract the value of the central picture cell from the average of the neighbors. The result is clearly zero in regions of constant brightness. This is true even in areas where brightness varies linearly.

Such approximations to differential operators are used in the finite-difference solution of partial difference equations. Recall what was said earlier about stencils. The coefficient by which a value is multiplied is called a *weight*. The pattern of weights, arranged spatially to indicate which picture cells they apply to, is called a *stencil* or *computational molecule*. The stencil in our case is

p152 -

$\frac{1}{\epsilon^2}$

	1	
1	−4	1
	1	

where the term on the left is a multiplier to be applied to all weights. Remember that the application of the Laplacian is equivalent to convolution with a generalized function defined by a sequence of functions that feature a central depression surrounded by a positive wall. The above discrete approximation should remind you of the functions in that sequence.

p122 -

On a squared grid it is hard to come up with a stencil that approximates the Laplacian and is symmetric. Earlier, when we tried to find a consistent definition of connectivity for binary images, we had to make a decision about which of the neighbors of a picture cell were to be considered connected to that picture cell. Here we are again faced with the issue of whether we should include only the edge-adjacent picture cells, or the corners as well. On a hexagonal grid this problem does not occur; all six neighbors are weighted equally (figure 8-2).

One way to proceed is to consider a coordinate system rotated 45° with respect to the xy-coordinate system. If we label the axes in the new coordinate system x' and y', we can use the approximation

$$\frac{\partial^2 E}{\partial x'^2} \approx \frac{1}{2\epsilon^2}(E_{i+1,j+1} - 2E_{i,j} + E_{i-1,j-1}),$$

$$\frac{\partial^2 E}{\partial y'^2} \approx \frac{1}{2\epsilon^2}(E_{i-1,j+1} - 2E_{i,j} + E_{i+1,j-1}),$$

At 45° the effective spacing is ($\sqrt{2}\epsilon$)

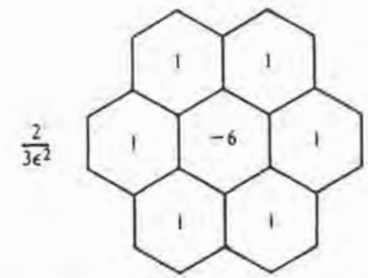

Figure 8-2. An excellent approximation of the Laplacian operator on a hexagonal grid is obtained by subtracting the value in the center from the average of the six neighboring cells. The result is multiplied by a constant that depends on the grid spacing ϵ.

so that

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} \approx \frac{2}{\epsilon^2}\left(\frac{1}{4}(E_{i+1,j+1} + E_{i-1,j+1} + E_{i+1,j-1} + E_{i-1,j-1}) - E_{i,j}\right).$$

The corresponding stencil is

$\frac{1}{2\epsilon^2}$

1		1
	−4	
1		1

Clearly, linear combinations of the two stencils shown also produce estimates of multiples of the Laplacian. A popular combination that, as we show in exercise 8-8 provides a particularly accurate estimate of the Laplacian, is

$\frac{1}{6\epsilon^2}$

1	4	1
4	−20	4
1	4	1

This is produced by adding two-thirds of the first to one-third of the second estimate introduced above. We show in the exercise that this operator can be written in the form

$$\nabla^2 + \frac{\epsilon^2}{12}\nabla^2(\nabla^2) + e,$$

where ∇^2 is the Laplacian and e contains terms of sixth and higher order multiplied by ϵ^4 and higher powers of ϵ.

Finally, to obtain a discrete approximation of the quadratic variation we need the cross derivative $\partial^2 E/\partial x \partial y$. A suitable stencil for computing it is

$$\frac{1}{4\epsilon^2}\begin{array}{|c|c|c|}\hline -1 & & 1 \\ \hline & & \\ \hline 1 & & -1 \\ \hline\end{array}$$

It should be clear how to compute the quadratic variation now:

$$\left(\frac{\partial^2 E}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 E}{\partial x \partial y}\right)\left(\frac{\partial^2 E}{\partial y \partial x}\right) + \left(\frac{\partial^2 E}{\partial y^2}\right)^2$$
$$\approx \frac{1}{\epsilon^4}\Big((E_{i-1,j} - 2E_{i,j} + E_{i+1,j})^2 + (E_{i,j-1} - 2E_{i,j} + E_{i,j+1})^2$$
$$+ \frac{1}{8}(E_{i+1,j+1} - E_{i+1,j-1} + E_{i-1,j-1} - E_{i-1,j+1})^2\Big).$$

8.4 Local Operators and Noise

In practice, the local operators for accentuating edges introduced so far are not directly useful, mainly because their output is seriously degraded by noise in the image. Noise, being independent at different picture cells, has a flat spectrum. The high-frequency components of noise are greatly amplified by the simple differential operators, as we saw when we discussed image processing. Edges of low contrast are simply lost in the noise.

We can apply the optimal filtering methods explored in chapter 6. From our discussion there it should be apparent that the optimal filter for recovering the Laplacian, for example, is

$$H' = -\rho^2 \frac{1}{1 + \Phi_{nn}/\Phi_{bb}}.$$

We need to make some assumptions about the images in order to estimate the power spectrum Φ_{bb}. If an image has a flat power spectrum, little can be done. However, images tend to have spectra that fall off with frequency, as we discussed in chapter 6. The transform of a single step function, for

example, falls off as the inverse of frequency. Images in which regions of uniform brightness are separated by sharp edges can be expected to have similar properties.

Suppose that $\Phi_{bb} = S^2/\rho^2$ and $\Phi_{nn} = N^2$. Then

$$H' = -\frac{\rho^2 S^2}{S^2 + \rho^2 N^2},$$

so that the optimal filter behaves like the Laplacian at low frequencies, but does not amplify high frequencies as much. The maximum gain is $-S^2/N^2$, equal to (minus) the signal-to-noise ratio.

The optimal filter can be decomposed into two filters. The first recovers the image as well as possible in a least-squares sense. This is done by multiplying the transform by $S^2/(S^2 + \rho^2 N^2)$. The second filter computes the desired output from the result produced by the first. In the case discussed here, this is accomplished by applying the Laplacian, that is, by multiplying by $-\rho^2$. It is clear that an image with little information content above a certain frequency is first lowpass filtered by the optimal filter. If the power spectrum rolls off at some known frequency, the optimal filter will tend to roll off near that frequency. Alternatively, we can apply this filter to the differential operator, since convolution is associative. The optimal estimator of the Laplacian of an image is an operator with nonzero spatial support, since it is obtained by applying the Laplacian to the optimal filter for estimating the image. This means that it is not local. In practice, the region over which the operator has substantial weight can be quite large.

If our objective is to obtain the best estimate of the image gradient, then we find the gradient of the estimate of the original image produced by the optimal filter. Alternatively we can take the x and y derivatives of the point-spread function of the optimal filter and use these directly in convolution with the computed image.

If, for example, the optimal filter for estimating the original image $E(x,y)$ is a Gaussian,

$$h(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2}\frac{x^2+y^2}{\sigma^2}},$$

then the point-spread function of the optimal filters for recovering E_x and E_y are

$$h_x(x,y) = -\frac{x}{2\pi\sigma^4} e^{-\frac{1}{2}\frac{x^2+y^2}{\sigma^2}} \quad \text{and} \quad h_y(x,y) = -\frac{y}{2\pi\sigma^4} e^{-\frac{1}{2}\frac{x^2+y^2}{\sigma^2}},$$

respectively. The support of these operators is large, quite unlike the local operators considered earlier. Thus they compute weighted averages over large areas. As a result, noise contributions are reduced, while the

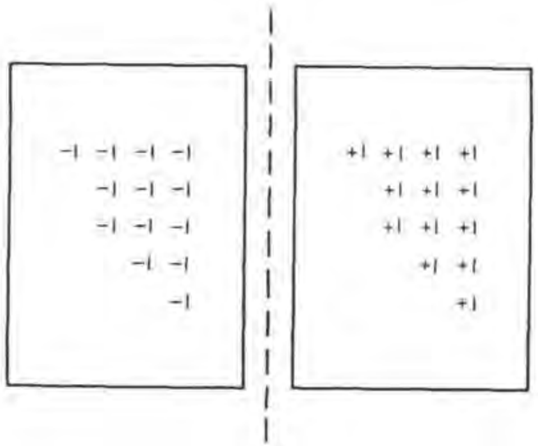

Figure 8-3. A simple edge operator subtracts the sum of the gray-levels obtained over one large area from the sum over an adjacent large area.

signal—the difference between the brightnesses of neighboring regions—is maintained.

We can proceed from this intuition by considering a simple method for estimating the difference in brightness across an edge (figure 8-3). Suppose we take the difference between the average of N picture cells on one side and the average of N picture cells on the other. The mean of the result is the brightness difference. The standard deviation of the noise in the result, on the other hand, is $\sigma/\sqrt{N/2}$, if the standard deviation in each measurement is σ. Thus the larger we make the patches on each side of the edge, the less will the result be affected by noise.

If the patches become very large, however, other image regions may be included (figure 8-4). The measurement will then be incorrect, unless the image contains just one edge.

8.5 Edge Detection and Localization

If the edge-enhanced signal is substantially above the noise, we might conclude that there is an edge at a certain point in the image. Such a decision is not perfectly reliable since the noise may just happen to add up sufficiently at that point. All we can do is try to reduce the probability of this happening by choosing the threshold applied to the edge-enhanced image so that the number of false edge points is acceptable.

If the threshold is too high, weak edges will be missed. So there is a trade-off between two kinds of errors. By increasing the size of the patches

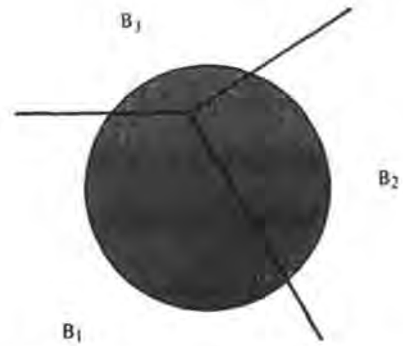

Figure 8-4. A limit to the size of the support region of an edge operator is set by the spacing between image features. If it is too large, other image detail will be included. In some places—for example, near joints in an image of a polyhedral scene—any finite support area will lead to inclusion of unwanted areas.

over which averages are taken, or equivalently by lowering the rolloff frequency above which image components are suppressed, we can reduce the effect of noise and make weak edges easier to detect. We face another trade-off immediately, however, since other edges in the image might be included in the larger patches. We note that short edges must have higher contrast in order to be detectable.

The enhanced image will have high values not just for the picture cells right on the edge, but also for a few neighboring ones. This effect is particularly pronounced if the image has been smoothed to reduce the effects of noise. Thus the problem of locating the edge arises. If it were not for noise, we would expect to find the largest values right on the edge. These large values could then be used to suppress neighboring large values.

In the case of the squared gradient, the edge-enhanced images will have a ridge for each edge with height proportional to the square of the brightness difference. In the case of the Laplacian and the quadratic variation, there will be two parallel ridges, one on each side of the edge. For the Laplacian the two ridges will have opposite polarity, and the edge can be located where the sign changes.

8.6 Conclusion and Examples

Eventually it will be necessary to identify the physical cause of the edge. What aspect of the three-dimensional scene created the changes in image brightness picked up by the edge detector? Little is understood about this