

The Detection of Intensity Changes by Computer and Biological Vision Systems*

ELLEN C. HILDRETH

MIT Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, Mass. 02139

Received October 21, 1981; revised February 9, 1982

This article describes the implementation of a theory for the detection of intensity changes, proposed by Marr and Hildreth (*Proc. R. Soc. London, Ser. B* 207, 1980, 187-217). According to this theory, the image is first processed independently through a set of different size operators, whose shape is the Laplacian of a Gaussian, $\nabla^2 G(x, y)$. The loci, along which the convolution outputs cross zero mark the positions of intensity changes at different resolutions. These zero-crossings can be described by their position, slope of the convolution output across zero, and two-dimensional orientation. The set of descriptions from different operator sizes forms the input for later visual processes, such as stereopsis and motion analysis. There are close parallels between this theory and the early processing of information by the human visual system.

1. INTRODUCTION

The first goal in analyzing an image is to describe, locally, the significant intensity changes. These changes often correspond to what we intuitively call "edges," and are the consequence of a change in some physical property of a surface, such as reflectance, geometry, or incident illumination. The importance of edges in image analysis has led to extensive work on their automatic detection by computer vision systems. This article describes a theory of the detection of intensity changes which was motivated by a desire to understand early processing in human vision [1, 2]. Recent quantitative studies, from neurophysiology and psychophysics, have gone far in suggesting the type of operations that the human visual system first applies to an image. These studies have led us to develop a computational model for the detection of intensity changes, which suggests that these operations may be optimally suited for this task.

The basic theory has been presented in Marr and Hildreth [1]; the purpose of this article is to further detail the computational aspects of the theory, which are of greater interest to researchers in the computer vision community. The algorithm is described here in greater detail, and further examples of its performance on natural images are presented. We will illustrate the robustness of the initial operators in the presence of noise, and necessary limitations on the smallest size operator that can be used for the reliable detection of intensity changes. The computation of properties of intensity changes, such as contrast and width, will be discussed briefly. Section 2 provides a review of previous work on edge detection. The algorithm, and demonstration of its performance, is then presented in Section 3. The parallels between this

*Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643 and in part by National Science Foundation Grant MCS77-07569.

scheme, and early processing in the human visual system, will be described in Section 4.

2. RELATED WORK ON EDGE DETECTION

Many components of the scheme that we propose have been explored in previous work on edge detection. In this section, we review some of the early work which is closely related. Most edge detection schemes begin by applying small differential operators to an image, followed by a detection operation to locate small edge segments. A range of differential operators have been explored, utilizing directional first and second differences [3–12], and the rotationally symmetric Laplacian [13–15]. In the case of directional operators, they must be applied at several directions, in order to detect edges at all orientations in the image. A sharp change in intensity will give rise to a peak in the output of a first derivative operator, or zero-crossing in the output of the second derivative. Hence, the initial differentiation is generally followed by peak detection, or zero-crossing detection [5, 13], and thresholding and thinning operations are applied to localize boundaries, and remove edge segments that arise from noise in the imaging process (also see reviews in [16–18]).

Some operators incorporate a smoothing function, whose shape is typically rectangular (such as in [19, 20]), or Gaussian (such as in [6, 13]). The variety of types of intensity change that appear in an image has suggested the use of a range of operator sizes, with selection criteria for deciding which size best reflects the properties of the underlying edge (for example, [9, 19, 20]). A limitation of many of these schemes has been the use of extremely small operators, covering an area of 10 to 20 image elements. As a consequence, they have been very sensitive to noise and quantization, restricting their use to limited image domains.

More recently, Shanmugam, Dickey, and Green [21] proposed a frequency domain filter which is optimal for edge detection, in that it yields maximum energy in the vicinity of an edge in the image. Their filter also satisfies a small set of additional constraints, such as (1) limited frequency bandwidth, (2) small filter output when the input is constant or slowly varying, and (3) the one-dimensional operator is an even function, which will correspond to a two-dimensional operator which is rotationally invariant. For the special case of a step edge input, the optimal frequency domain filter corresponds to a spatial impulse response which is the second derivative of a Gaussian (for a given bandwidth) [50]. Shanmugam *et al.* follow their filtering stage with magnitude and thresholding operations. The model for detecting intensity changes that we will develop in this article leads to the rotationally invariant second derivative operator, the Laplacian of a Gaussian $\nabla^2 G$, but we will see that there are more appropriate operations to apply to its output, in order to locate the intensity changes.

3. THE DETECTION OF INTENSITY CHANGES

A distinction is drawn between the detection of *edges* and *intensity changes*: a change in intensity is the phenomenon that we will detect and describe in the image; edges are the physical changes that give rise to these intensity changes. Our scheme involves the detection of intensity changes at a range of image resolutions; in general, there may not exist a one-to-one correspondence between intensity changes detected at a particular resolution, and edges in the physical scene. Many open

questions remain for the final assessment of edges, and their type (for example, whether an edge corresponds to a reflectance change, shadow boundary, discontinuity in surface orientation, and so on). Later in this article, we will describe one heuristic that may allow us to explicitly determine that a particular intensity change corresponds to a physical edge.

Detecting Intensity Changes at Many Resolutions

Changes in intensity will occur in the image at a range of different scales. If we look at individual picture elements (pixels), we find intensity changing from pixel to pixel. Often, there will be uniform changes over some distance. Most edges in the real world are sharp edges; the intensity function will be composed of a few steep changes over a small number of pixels. Other edges, such as shading edges, are blurred; their corresponding intensity function will increase slowly over a large number of pixels. These different types of intensity change are not distinct in the image; for example, in an image of a textured surface, such as the leaf of Fig. 4a, it is common to find high frequency changes in intensity, due to the surface texture, superimposed on a more coarsely varying background intensity, possibly due to a slowly changing surface geometry. A single operator cannot be optimally suited for detecting intensity changes at different scales. It is this problem that motivated the use of different size operators by Marr [9], Binford [10], Rosenfeld and others [19, 20]. Small operators are used to detect intensity changes occurring at a resolution close to the image resolution. The intensity function is then smoothed over successively larger areas, to detect intensity changes that occur at coarser resolutions.

As we consider larger operators, the shape of the underlying smoothing filter becomes important; two issues constrain this shape. First, changes in the physical world are generally localized in space, so we desire that our initial operators also be spatially localized. Second, one of our goals for this operator is to restrict the scale at which intensity changes take place in the output of the operator. One method for achieving this is through bandlimiting in frequency. For example, to detect gross changes in intensity, we may utilize an operator whose frequency spectrum is localized about the low frequencies. Bandlimiting in frequency will, in general, conflict with the need for localization in space, but the two requirements are optimally satisfied by the Gaussian distribution [22] in that the Gaussian minimizes the product of bandwidth in space and frequency. To detect intensity changes at different scales, our initial operators will incorporate Gaussian smoothing functions of different sizes.

The Differential Operator

If an intensity change occurs along a particular orientation in the image, there will be a peak in the first directional derivative of intensity measured perpendicular to the change, or a zero-crossing in the second directional derivative. If we utilize directional operators, it becomes necessary to apply operators at a range of directions, in order to detect intensity changes at different orientations in the image. It would be computationally more efficient if we could apply a single nondirectional operator to the image; the simplest nondirectional linear differential operator is the Laplacian [23]. The elements in the output of the Laplacian operator, which signify the location of an intensity change, are the loci along which the output crosses zero [34].

The use of the Laplacian raises the following question: will the zero-crossings in the output of the Laplacian correctly capture the intensity changes that we want to detect? In [1, Appendix B] it is shown that if the variation of intensity along the orientation of an intensity change is at most linear, then the zeros of the Laplacian will coincide with the zeros of the second directional derivative, taken perpendicular to the local orientation of the intensity change. This establishes the conditions under which it is valid to use the Laplacian to detect intensity changes; its use rests on the assumption that these conditions are generally satisfied in natural images. If the intensity variation along an intensity change is highly nonlinear, the positions of the zero-crossings in the Laplacian output will deviate from those of the second directional derivative. Later in this section, we will show a direct comparison between the zero-crossings in the output of a Laplacian operator, and those of a directional second derivative operator, which provides empirical support for this assumption.

The ∇^2G Operator

The two operators, the Gaussian and Laplacian, can be combined into a single operator, so that one can now detect intensity changes occurring at a particular scale by locating the zero-crossings in the output of $\nabla^2G(x, y)$, the Laplacian of a Gaussian distribution, which is given by the following expression (up to a constant scaling factor):

$$\nabla^2G = \left[\frac{r^2}{\sigma^2} - 2 \right] \exp\left(-\frac{r^2}{2\sigma^2} \right),$$

where r is the distance from the center of the operator, and σ is the space constant for the Gaussian, $G(r) = \sigma^2 \exp(-r^2/2\sigma^2)$. The two-dimensional operator and its Fourier transform are illustrated in Figs. 1b and d. The one-dimensional analogue D^2G and its Fourier transform, are shown in Figs. 1a and c. (The operators illustrated in Fig. 1 differ from the above expression by a change in sign.) Examples of the application of this operator appear in Fig. 2. The convolution outputs for the original images of Fig. 2a are shown in Fig. 2b. Zero is represented by a medium grey, so that very positive values in the convolution output are white, and negative values black. In Fig. 2c, all positive values of the convolution output are shown in white, and negative values black. Finally, the loci of zero-crossings are shown in Fig. 2d.

The pictures are 320×320 pixels. The size of the operator is defined by w , the diameter of its positive central region; w varies linearly with the space constant σ of the underlying Gaussian $w = 2\sqrt{2}\sigma$. To construct the convolution operator, we have taken discrete samples of the continuous ∇^2G distribution, and scaled and quantized these values into 12-bit integers (a constraint of our convolution hardware). The operator is therefore truncated, and its values are then manipulated slightly to yield an operator whose discrete values sum to zero. This allows a zero response to constant input. This transformation to the discrete operator yields an error from the ideal convolution, but this error may not be significant for image understanding systems. A careful theoretical analysis of these issues would be valuable.

For the ∇^2G operator used in Fig. 2, $w = 9$ pixels, and the overall extent of the operator is 35×35 pixels. The subsequent zero-crossing computation involves

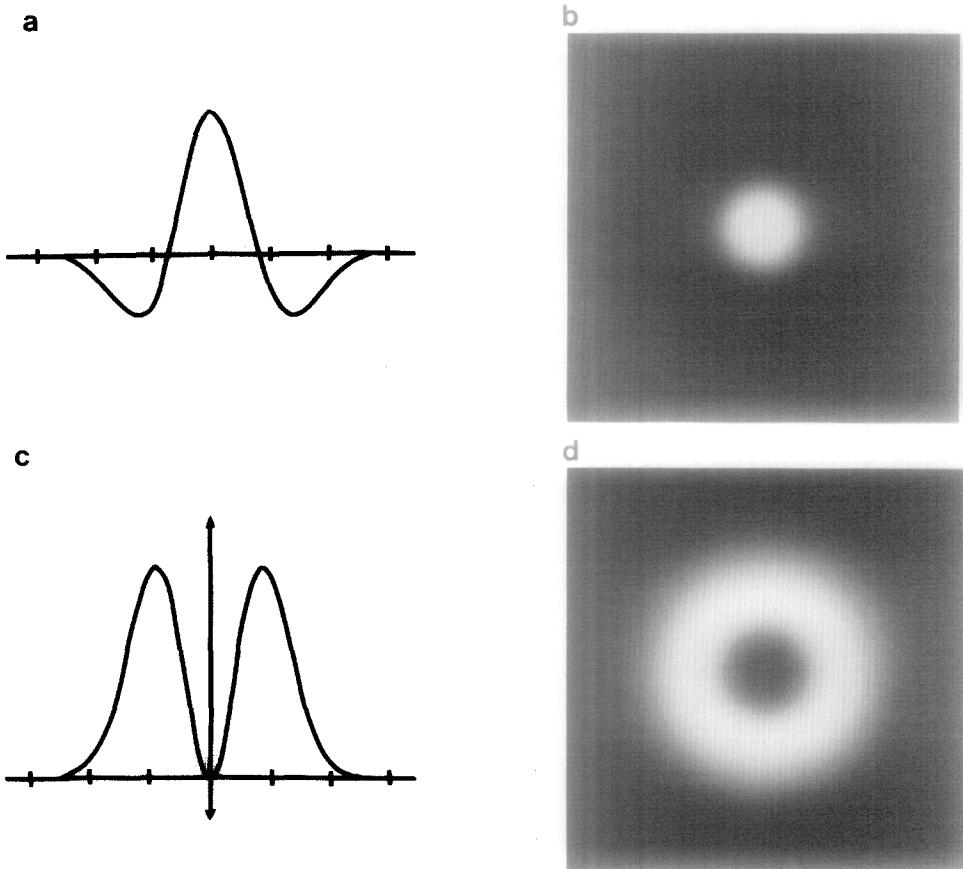


FIG. 1. The initial filters. (a) illustrates D^2G , the one-dimensional second derivative of a Gaussian, and (b) shows its two-dimensional counterpart, ∇^2G . (c) and (d) illustrate the one- and two-dimensional Fourier transforms.

detection of a change in sign of the convolution output between adjacent picture elements. (There is considerable flexibility available for the detection of zero-crossings; a different algorithm is suggested by their detection in the human visual system [1].) Figure 3 illustrates the zero-crossing contours from the output of three different size ∇^2G operators applied to an image; $w = 6, 12$, and 24 for Figs. 3b-d, respectively. A second example appears in Fig. 4. The image of Fig. 4a has been convolved with ∇^2G operators of size $w = 9$ and 18 pixels; the resulting zero-crossings are shown in Figs. 4b and c.

In Fig. 5, we show the comparison between the zero-crossings in the output of the Laplacian, and those of the output of a directional second derivative operator. The horizontal cross section of the directional operator is the one-dimensional second derivative of a Gaussian; its vertical cross section is a Gaussian

$$M(x, y) = \exp\left(-\frac{y^2}{2\sigma_v^2}\right)\left(\frac{x^2}{\sigma_x^2} - 1\right)\exp\left(-\frac{x^2}{2\sigma_x^2}\right),$$

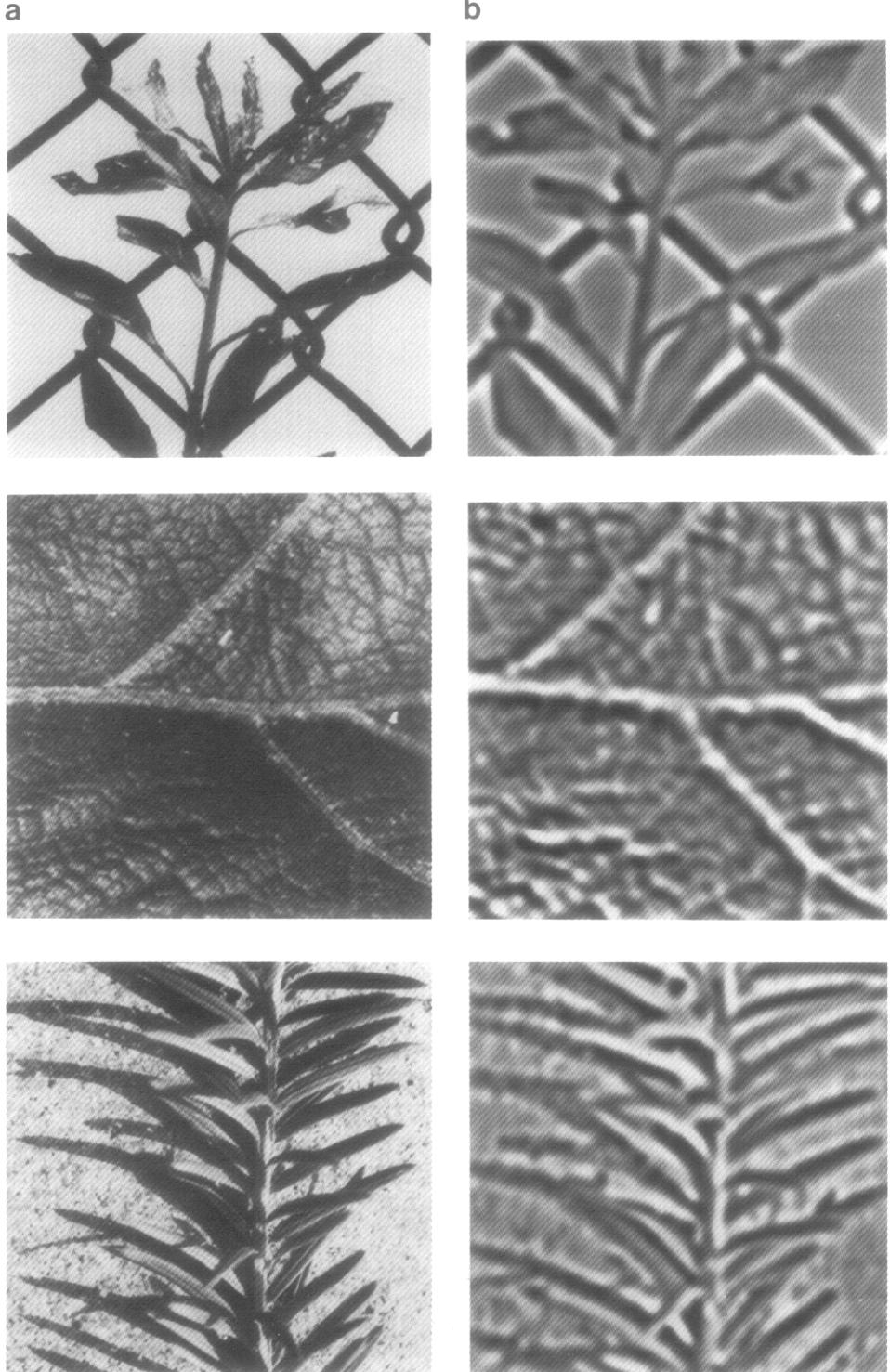


FIG. 2. Examples of zero-crossing detection using $\nabla^2 G$. Column (a) shows three images, and column (b) shows the convolutions with the $\nabla^2 G$ operator of Fig. 1 ($w = 9$ pixels), with zero being represented by a medium grey. In column (c), positive values of the convolution are shown in white, and negative values, black. The zero-crossing contours appear in column (d).

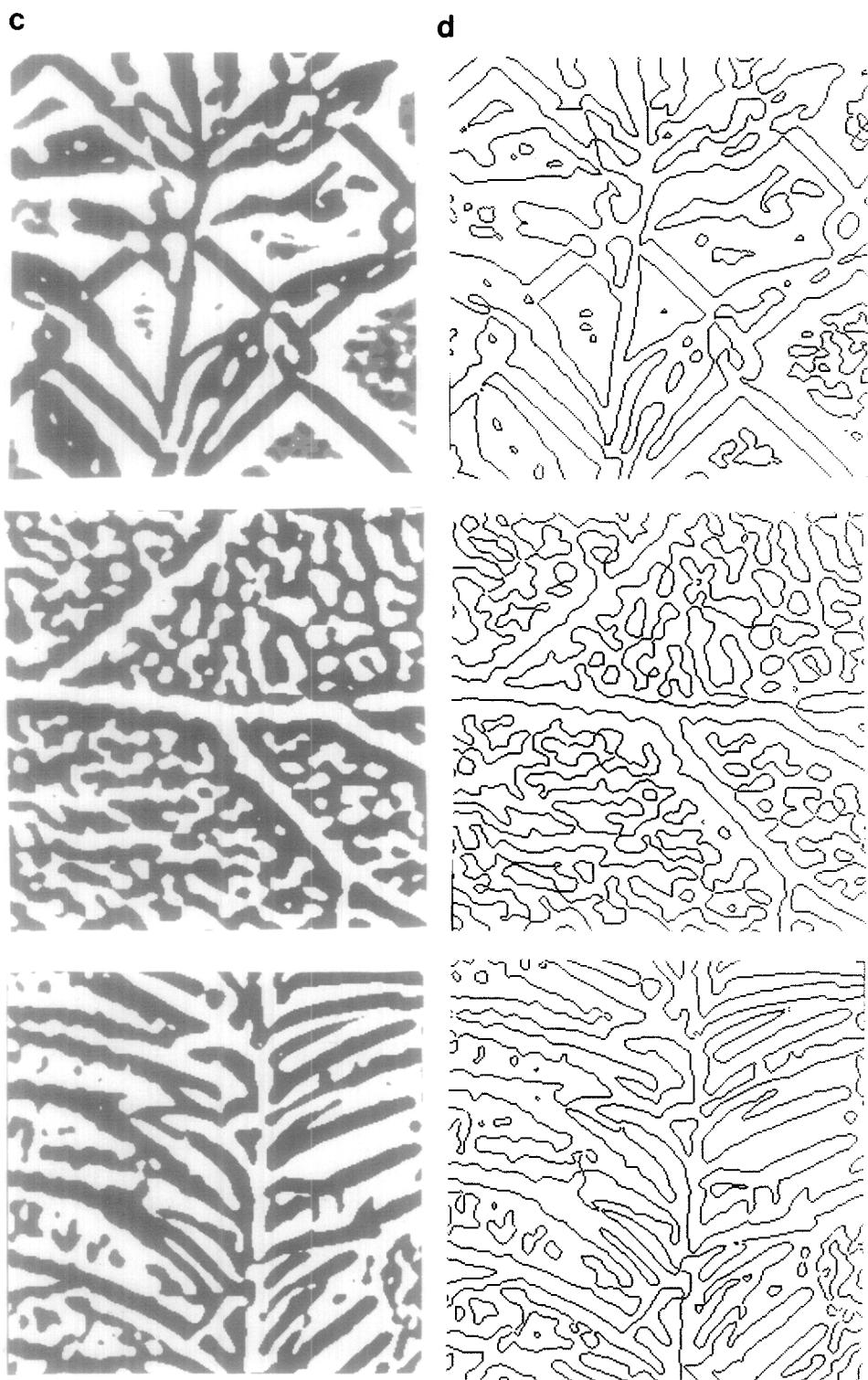


FIG. 2.—*Continued.*

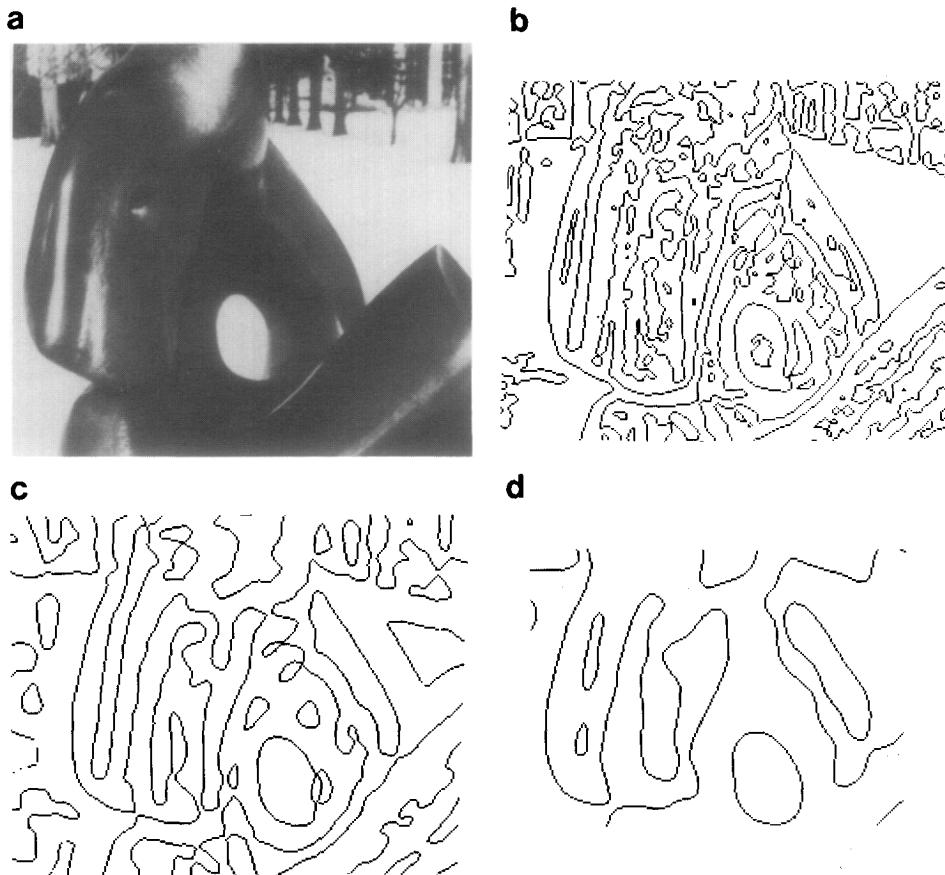


FIG. 3. Zero crossings from different size operators. The image in (a) has been convolved with $\nabla^2 G$ having $w = 6, 12$, and 24 pixels. (b)–(d) show the resulting zero-crossing contours.

where σ and σ_v are the space constants for the horizontal and vertical Gaussians. The zero crossings of the $\nabla^2 G$ output, with $w = 9$ pixels, are shown in Fig. 5a. In Fig. 5b are the zero crossings in the output of $M(x, y)$ applied to the image, where the width of the positive central region w_x , is also 9 pixels, and $\sigma_v = 3$. In Fig. 5c, the two zero-crossing descriptions are superimposed; the zero crossings of Fig. 5a are displayed in black, those of Fig. 5b are light grey, and zero crossings from the two outputs which exactly coincide are shown in medium grey.

In Fig. 5c, we can observe that for extended edges, whose orientation is not close to the horizontal, the zero-crossings from the two operators coincide. For this image, this comparison supports the assumption of linear variation of intensity along the orientation of intensity changes, on which the use of the Laplacian is based. This same comparison on other natural images yields similar results. Two comments about schemes for detecting intensity changes can be made. First, it is possible to use directional second derivative operators, with the appropriate direction determined locally by the orientation of an intensity change. In such a scheme, it is unnecessary to assume linear variation of intensity along the orientation of the change. Second,

there are situations in which this assumption is violated. An example is at corners, where the orientation of the intensity change is varying rapidly. Havens and Strikwerda [24] have developed an edge detection operator, based on the definition of edges as the loci of points of maximum gradient of intensity in the direction of the gradient. They show that their operator provides a better description of the position of the intensity change around a corner. While in general there may exist situations in which other edge detection schemes yield a more precise description of the position of intensity changes in the image, by making the assumption of linear variation, we gain computational simplicity and efficiency. A simple linear convolution with $\nabla^2 G$ can be applied uniformly across the image. We will see in Section 4 that the human visual system utilizes the difference of two Gaussian approximation to the $\nabla^2 G$ operator, so is implicitly making the assumption of linear variation.

The Slope and Orientation of the Zero-crossings

There are two additional properties which can be computed for the zero-crossings at a particular scale: slope and local orientation. The slope of a zero-crossing is the rate at which the convolution output changes as it crosses zero, and is related to the contrast and width of the intensity change. In the case of a single, isolated intensity change in one dimension, there will be a single zero-crossing at the location marking the middle of the edge in each operator output, as shown in Fig. 6, whose slope s is related to the contrast and width of the edge and size of the one-dimensional operator by the equation

$$s = c \exp\left(-\frac{d^2}{8\sigma^2}\right),$$

where c and d are the edge contrast and width, and σ is the space constant for the Gaussian. The two-dimensional case is a straightforward extension; in this case, slope is related to the parameters of the two-dimensional intensity change by the expression

$$s = c \exp\left(-\frac{d^2}{4\sigma^2}\right).$$

In the isolated case, the slopes of zero-crossings from two different size operators could be used to compute the contrast and width of the intensity change explicitly. In Fig. 7b, the magnitude of the slope of the zero crossings of the convolution of the image in Fig. 7a with $\nabla^2 G$ operator of size $w = 9$ pixels, is displayed as intensity; sharp, high contrast edges yield darker zero-crossing contours. All the zero-crossing contours are shown in Fig. 7c.

The local orientation of a zero-crossing contour is taken to be the direction of the tangent along the contour in the image plane, and can be computed from the gradient of the $\nabla^2 G$ output. In the present implementation, we simply compute the gradient of the $\nabla^2 G$ output, over a local 3×3 neighborhood around a zero-crossing point. Brady [25] has shown that the computation of orientation varies very little over a wide range of local gradient operators. The explicit computation of orientation is important for later analysis by processes such as stereo [26] and motion [27].

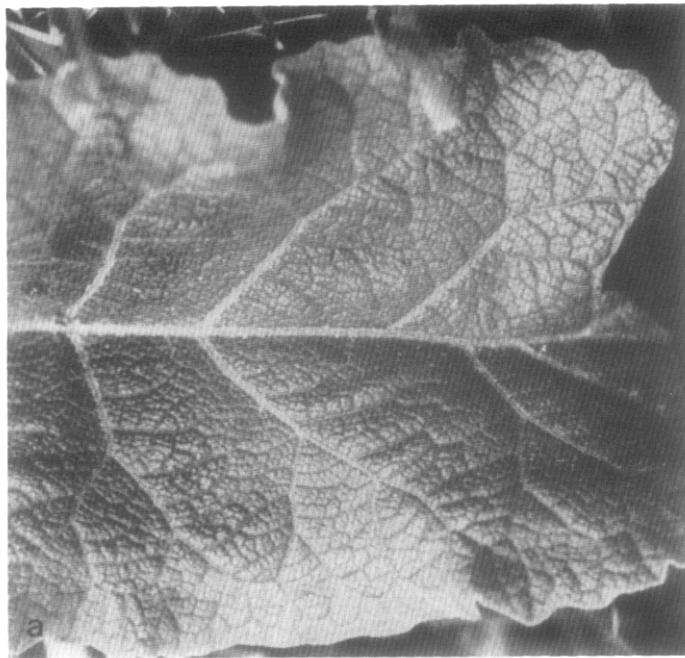


FIG. 4. Further demonstration of zero crossings from different size operators. The image in (a) has been convolved with $\nabla^2 G$ having $w = 9$ and 18 pixels. (b) and (c) show the resulting zero-crossing contours.

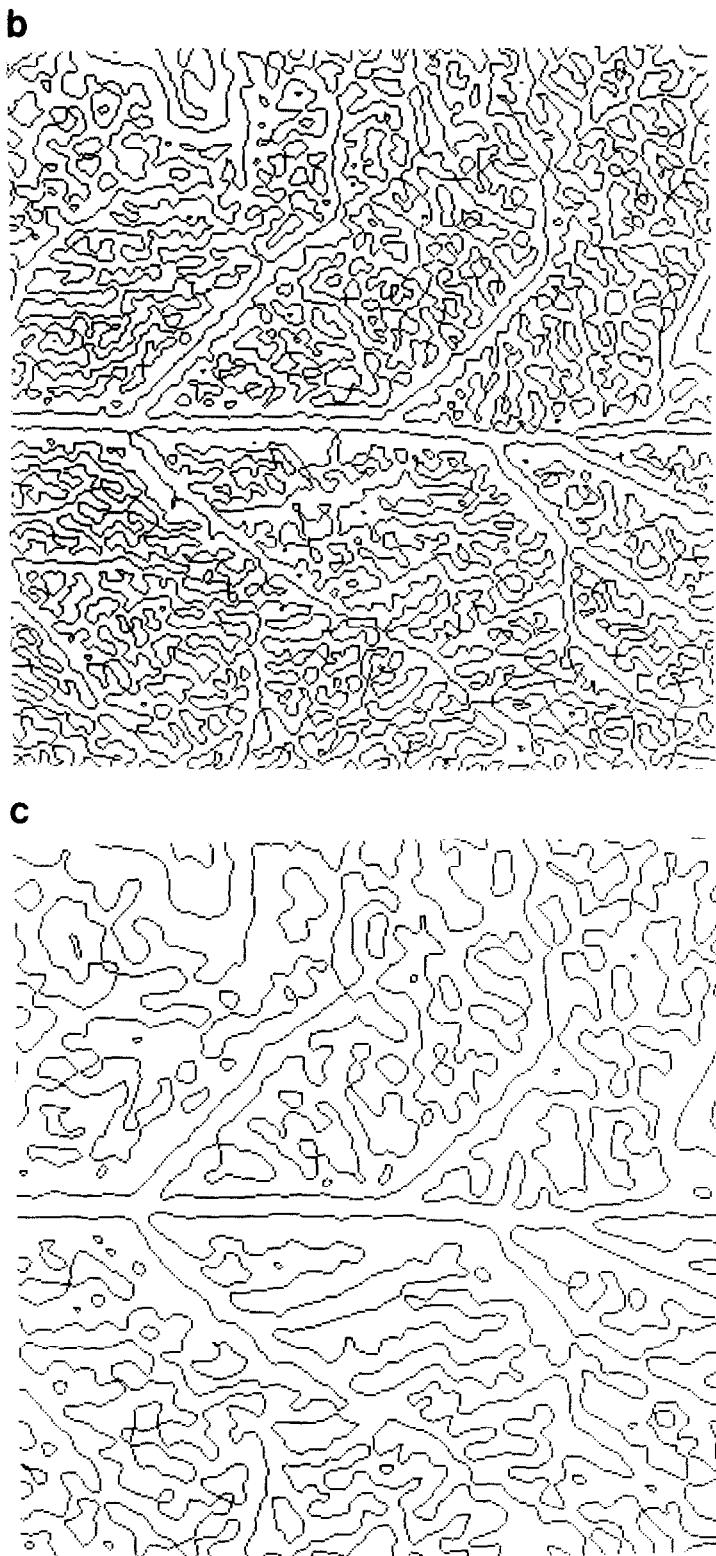
The Smallest Operator Size

In previous work on edge detection, designers have often been limited by practical engineering constraints to methods that were fast and used little memory. This led to the use of very small operators, covering a total area of 10 to 20 image elements. In the case of second derivative operators, the width of the central positive region was typically one or two elements. This contrasts sharply with the operators used in Fig. 2, where $w = 9$ pixels, and the total support of the operator covers over 1000 image elements. These larger operators yield more reliable detection of intensity changes in the presence of noise, or poorly quantized intensity.

Noise will generally be restricted to very high spatial frequencies; by restricting the smallest size operator, we are limiting the high spatial frequencies allowed to pass through the filters. Figure 8 demonstrates the need for this lower limit. Fig. 8a shows a simple two-dimensional step change in intensity, to which three levels of Gaussian noise have been added. A similar demonstration, used in the evaluation of other edge detection schemes, can be found in Pratt [18]. The signal-to-noise ratio is defined by the expression

$$\text{SNR} = c^2/\sigma_n^2,$$

where c is the edge contrast, and σ_n is the standard deviation for the added noise. (This particular definition of the signal-to-noise ratio was chosen to allow a direct comparison between our results and those of Pratt [18].) The three edges in Fig. 8a

FIG. 4.—*Continued.*

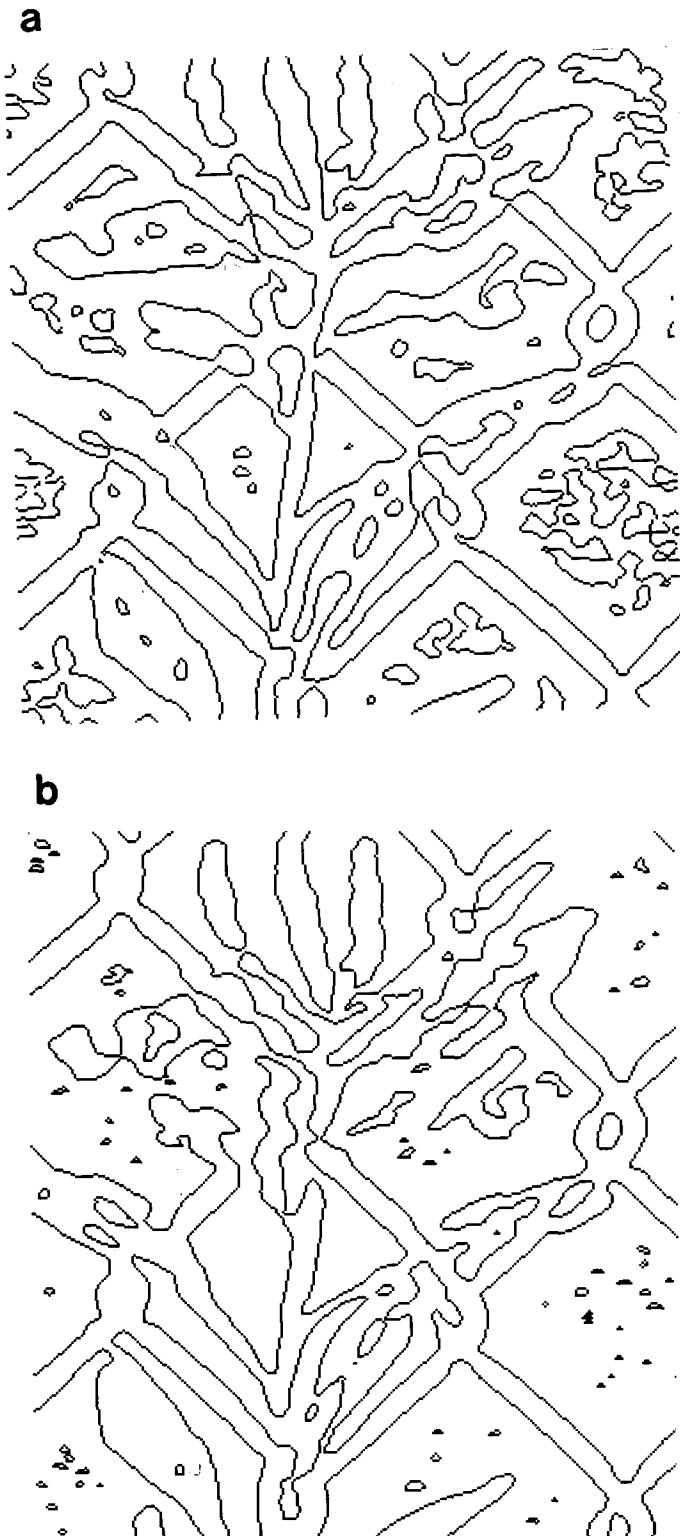


FIG. 5. Comparison between the zero-crossings in the output of the $\nabla^2 G$ operator, and $M(x, y)$ given in the text. (a) shows the zero crossings of the output of $\nabla^2 G$, with $w = 9$ pixels. (b) shows the zero crossings of the output of $M(x, y)$, with $w_x = 9$ pixels, and $\sigma_v = 3$. The two zero crossing descriptions are superimposed in (c); the zero crossings of (a) are displayed in black, those of (b) are light grey, and points at which the zero crossings exactly coincide are medium grey.

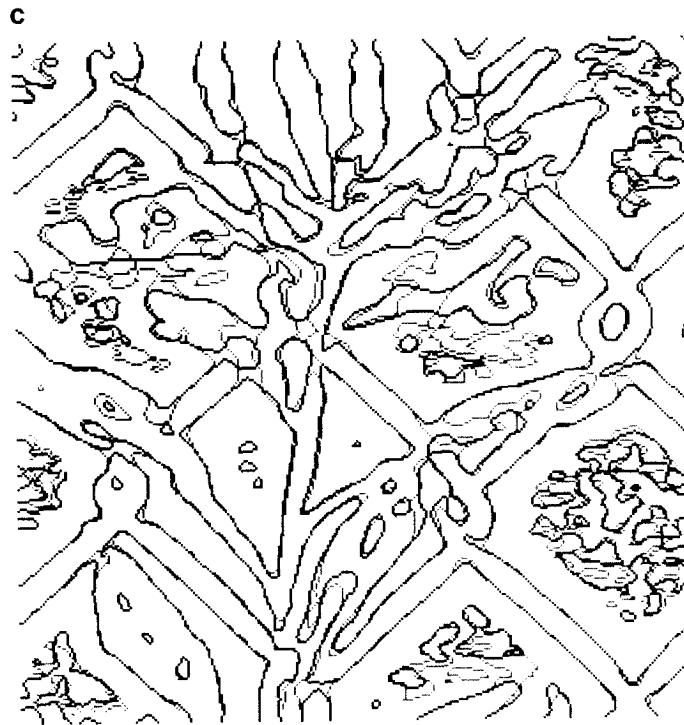


FIG. 5.—Continued.

have a signal-to-noise ratio of 100, 10, and 1, respectively. In Figs. 8b-d are the zero-crossings in the output of the three edges convolved with ∇^2G operators with $w = 1, 4$, and 8 pixels. The zero-crossings with maximum slope in each case are displayed at maximum intensity; the slope values are not correlated between outputs. For a signal-to-noise ratio of 100, all sizes are capable of detecting the primary edge, and respond more strongly to this edge than those due to the added noise. However, the smallest operator breaks down very fast, and for a signal-to-noise ratio of 1, it is unable to detect the primary edge behind the noise. The largest operator here corresponds to the smallest size presently used in our implementation; in this example, it is capable of responding more strongly to the underlying edge, even in the presence of large amounts of noise. For a signal-to-noise ratio of 100, slopes of many of the zero-crossings along the central edge are larger than those of the noisy edges.

Figure 9 provides a second demonstration, now in one dimension. Figure 9a contains an ideal intensity profile for a series of bars; in Fig. 9b, a high level of Gaussian noise has been added. This example is intended to emphasize that in evaluating the output of various size operators, we are interested in the position and slope of the zero-crossings, and how well they reflect the significant changes in our input profile. In Fig. 9c, the outputs of the one-dimensional D^2G operator, with $w_x = 6$ pixels, convolved with the ideal bar profile and noisy profile are superimposed. Although there is deviation in the overall noise output, the position and slope of the zero-crossings are well preserved. As we decrease the operator size, the error in localizing the edges of the bars does not change significantly, but the error in

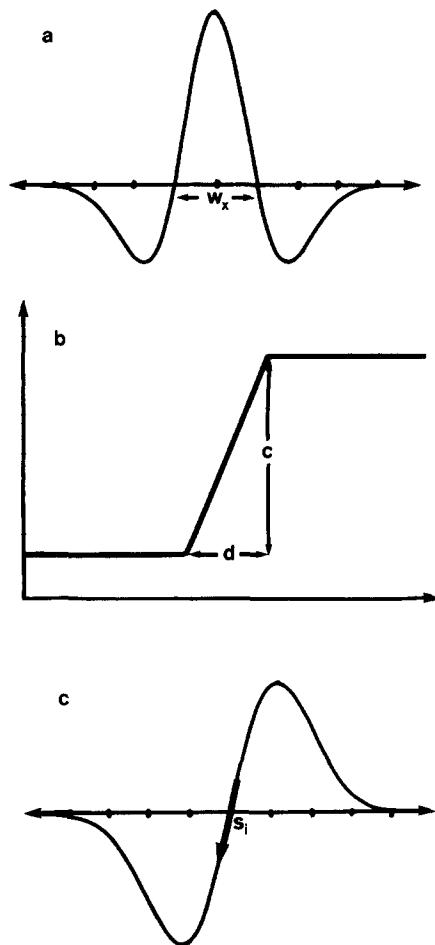


FIG. 6. The slope of a zero crossing as a function of the contrast and width of an intensity change. Displayed in (a) is the one-dimensional D^2G operator; the space constant of the Gaussian σ is related to w_x by the expression $w_x = 2\sigma$. The one-dimensional profile of an intensity change, indicating its contrast c and width d , is shown in (a). The output of the convolution of the profile in (b) with the operator in (a) is shown in (c), indicating the slope s of the zero crossing.

measuring slope increases, and additional zero-crossings, reflecting the noisy edges, are introduced. This is shown in Figs. 9d and e, where we have the outputs of the convolution of the ideal and noisy profiles with D^2G operators of size $w_x = 2$ pixels. The slopes associated with these noisy zero-crossings are comparable to those associated with significant edges, so that a simple thresholding technique cannot distinguish between the two. In the zero-crossing descriptions presented throughout this article, all zero-crossing contours are shown; there is no thresholding to remove zero-crossings whose slope is small.

If it is desired for a particular application, the position of the zero-crossings may be localized with a precision finer than the initial image sampling. The ∇^2G output is roughly linear across zero, so interpolation schemes as simple as linear interpolation may be used for this localization [2]. In Fig. 10 we show an example of a D^2G

output which has been reconstructed at a resolution which is five times the image resolution. In Fig. 10a we have the reconstruction of the output with the ideal sinc function. In Figs. 10b-d the same output is reconstructed with a truncated sinc, linear interpolation, and a Gaussian, respectively. In each case, the new reconstruction is superimposed on the profile of Fig. 10a. Linear interpolation successfully localized the zero-crossings at this precision. This observation that linear interpolation may be used to localize edges has also been made by MacVicar-Whelan and Binford [28].

The lavish use of large operator sizes is essential for modeling the operators in the human visual system, which analyzes its input at an extremely fine resolution (roughly one receptor per 25–27 sec. of visual arc for the central fovea [29, 30]), and utilizes operators in the fovea which cover areas of several hundred photoreceptors. In addition, we will see in Section 4 that the human visual system processes information through four or five different operator sizes, everywhere across the visual field, with adjacent sizes separated by about one octave.

Finally, I would like to note that we are no longer limited by the same practical engineering constraints of early researchers in machine vision. Using present technology, it has become possible to implement in real time [31–33] the type of operations for detecting intensity changes that are proposed in this article. With the use of special purpose hardware, the performance of edge detection by machine vision systems can now approach that of the human visual system.

Integrating the Zero-crossings from Different Operator Sizes

The integration of zero-crossing descriptions from different operator sizes remains an open research question. Some processes, such as stereopsis, appear to use the descriptions independently. The goal of stereopsis is to compute a match between elements in the left image, and those in the right, which correspond to the same element on a surface in the scene. The relative positions of the left and right elements may then be used to compute the depth of the corresponding physical element. What makes the matching process difficult is the false targets problem; for a given element in one image, there may be many candidate matches in the other image. Marr and Poggio ([34], see also [26]) present a theory of human stereo vision, in which they propose that the elements being matched are the zero crossings in the output of a range of different size $\nabla^2 G$ operators. Their algorithm takes advantage of the range of sizes to allow a wide range of disparity in the images, to be computed at a fine resolution. From the larger operators, the sparseness of features simplifies the matching process, and yields rough disparity information. Via vergence eye movements, this rough information is used to shift corresponding elements closer together, so that their relative positions are within the range that the smaller operators can match successfully. The smallest operator ultimately allows the computation of fine resolution disparity information; the size of the largest operator constrains the overall range of disparity that can be detected, for a particular eye position. A similar type of scheme may underly the analysis of visual motion; here, the different operator sizes are capable of measuring a wide range of velocities, at a fine resolution [35].

Marr and Hildreth [1] suggested that we may be able to use the zero-crossing descriptions more directly, to assert the presence of physical edges. The zero-crossing contours, particularly those from larger operators, do not religiously follow the

a



b



FIG. 7. The slope of zero crossings. Zero-crossings in the output of the $\nabla^2 G$ operator of size $w = 9$ pixels, applied to the image of (a), are shown in (b), with the slope of the zero crossing displayed as intensity; that is, a zero crossing across which the convolution output changes more sharply is displayed with a darker line. All the zero crossing contours are shown in (c).

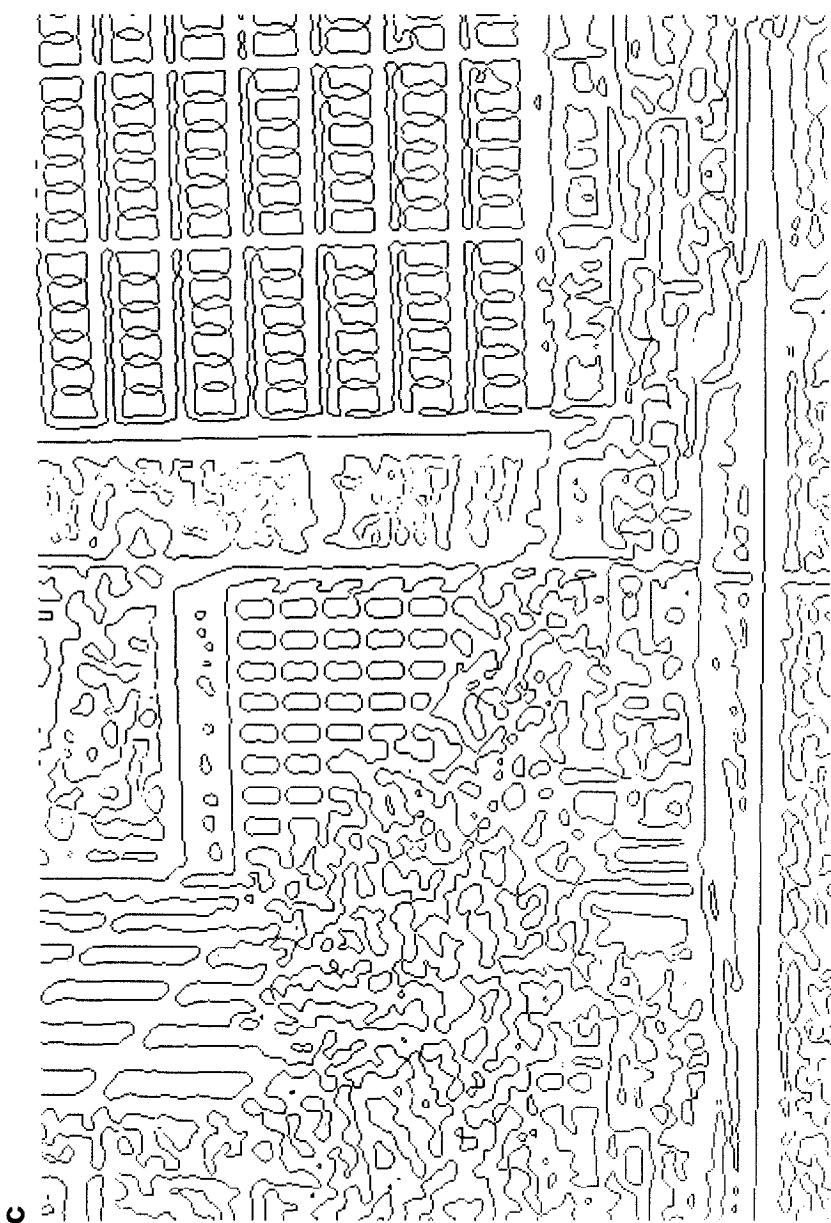


FIG. 7.—Continued.

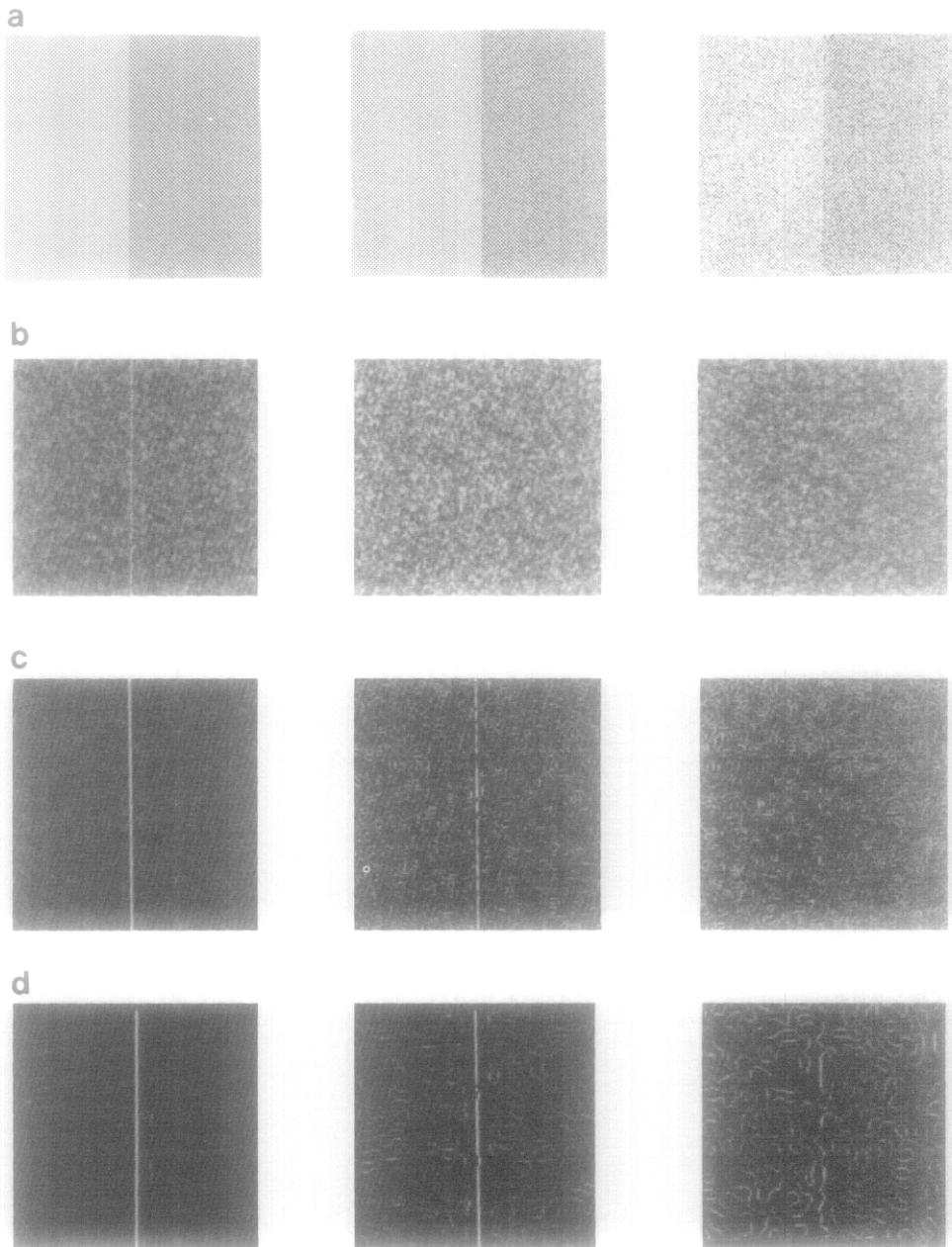


FIG. 8. The relationship between operator size and its sensitivity to noise. Row (a) illustrates three step changes in intensity, to which different levels of Gaussian noise have been added. The images were then convolved with $\nabla^2 G$ operators having $w = 1, 4$, and 8 pixels. The zero crossings of these outputs are displayed in rows (b)–(d), respectively.

projection of physical edges in the scene onto the image. Is it possible to use the zero-crossing information from different operator sizes to determine when a contour reflects the presence of a physical edge? We may take advantage of the following observation: many edges, such as those due to reflectance changes, give rise to spatially localized, sharp, broad-band changes in the intensity function. We would therefore expect a range of operator sizes to detect them, in roughly the same location in the image. This suggests that using the spatial coincidence of zero-crossing contours from nearby operator sizes may be used to assert the presence of an edge. Figure 11 illustrates the coincidence of zero-crossing contours from $\nabla^2 G$ operators whose sizes are separated by an octave; in this case, $w = 9$ and 18 pixels. Zero-crossings from the smaller operator are displayed in black, those of the larger operator in light grey. Points at which the zero-crossings exactly coincide are shown in medium grey. The smaller operator detects edges with much finer detail, but for strong physical edges, both operators detect the edge, with zero-crossing contours which roughly coincide.

The integration of zero-crossing descriptions from different size operators is a difficult and unsolved problem. The fact that multiple channels are necessary to compute the physical properties of an edge is implied by our equation relating the slope of a zero-crossing to the contrast and width of an intensity change. This point is emphasized in Fig. 12, which illustrates two intensity profiles, which yield zero-crossings that are identical in position and slope, for a $D^2 G$ operator of size $w_x = 8$ pixels. The slopes of zero-crossings from a second operator size could be used to resolve the parameters of each intensity change, provided that the operator is sufficiently small that its response to adjacent intensity changes do not interfere. The extent to which spatial coincidence might be used as part of the detection of physical edges, and computation of their properties, remains an open question.

4. EDGE DETECTION IN THE HUMAN VISUAL SYSTEM

Neurophysiological studies of the primate visual system, together with psychophysical studies, reveal many parallels between the ideas presented here and early processing in the human visual system. Cell recordings in the retina of a cat and monkey (for example, [36–41]) have uncovered two classes of retinal ganglion cells whose axons form the optic nerve fibers along which visual information is carried to the lateral geniculate nucleus before reaching visual cortex. These two cell types, termed *X* and *Y* cells, both have receptive fields with an antagonistic center-surround organization, whose shape is the difference of two Gaussians (DOG). Marr and Hildreth [1] proposed that the output of the *X* cells represents the output of the convolution of the DOG with the input image on the photoreceptors. More recently, Richter and Ullman [42] have proposed a detailed model for how this function arises from interactions between cell layers of the retina. At the visual cortex there is again a variety of cell types. The class of primary interest to us are the simple cells (see, for example, [43, 44]), which are selective to an edge or bar of a particular orientation and contrast sign moving across their receptive fields. Marr and Hildreth proposed that the function of one class of simple cells is to detect segments of the zero-crossing contours in the output of the DOG convolution that forms the input to these cells. This idea is presently under neurophysiological investigation (Kobi Richter, personal communication).

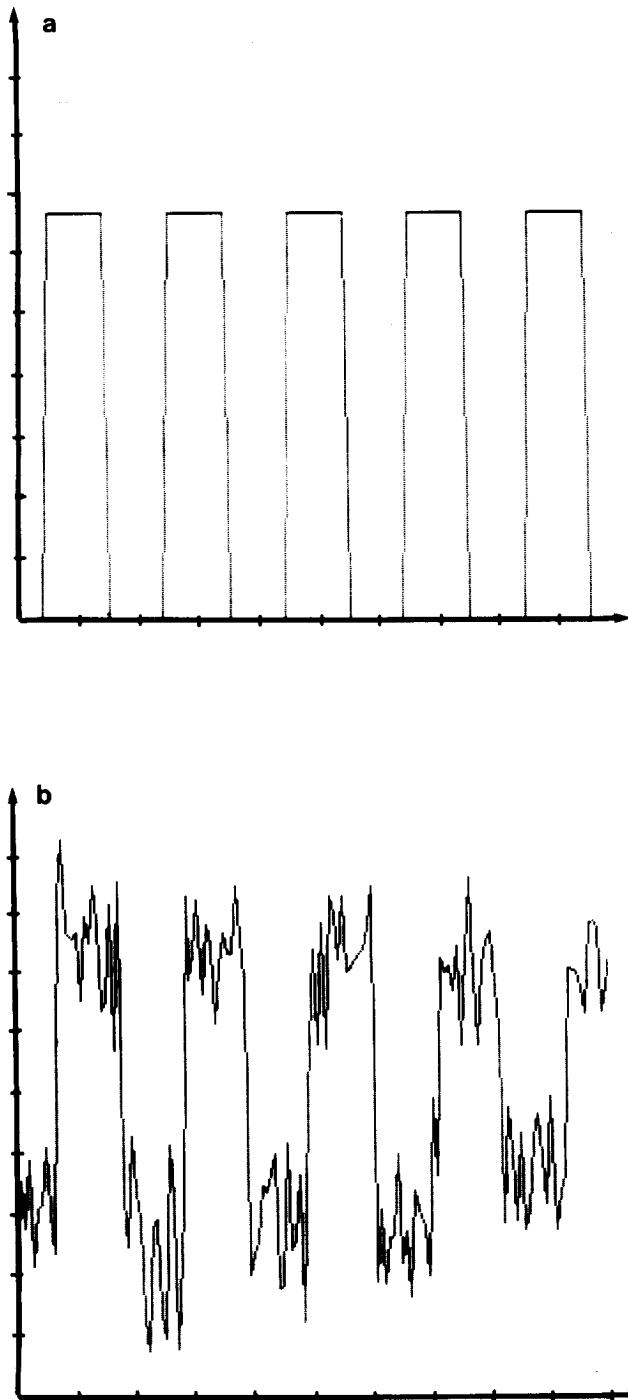


FIG. 9. One-dimensional demonstration of the relationship between operator size and its sensitivity to noise. (a) and (b) illustrate the ideal profile for a set of bars, and the bar profile with a large amount of Gaussian noise added. In (c), the outputs of the convolution of profiles (a) and (b) with a D^2G operator of size $w_x = 6$ pixels are shown superimposed. In (d) and (e) are the outputs of an operator with $w_x = 2$ pixels, applied to the ideal and noisy profiles, respectively.

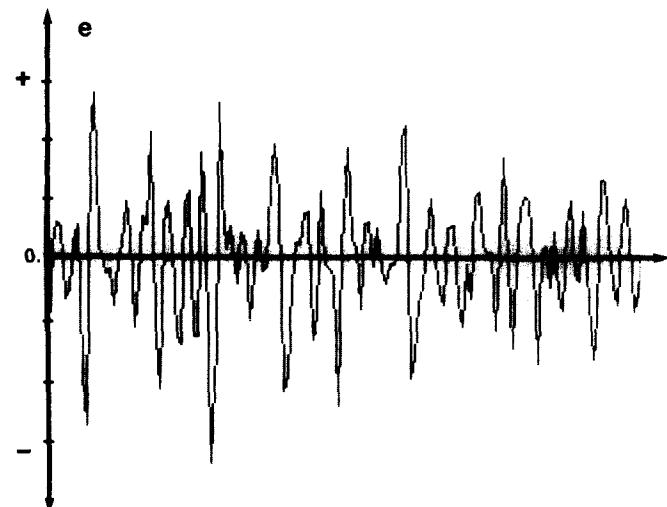
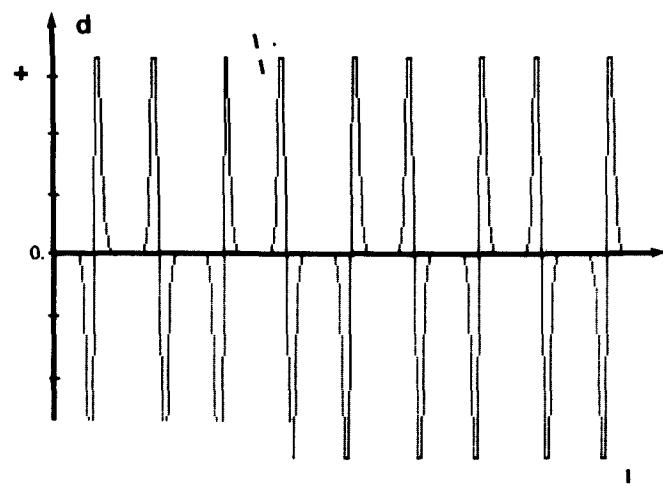
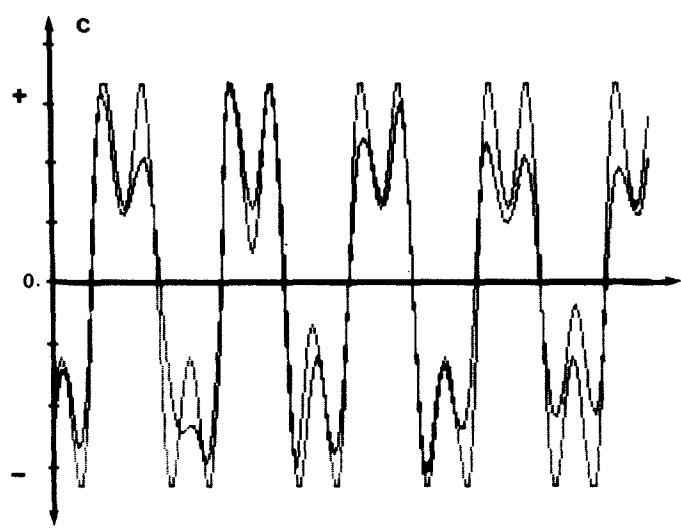


FIG. 9.—Continued.

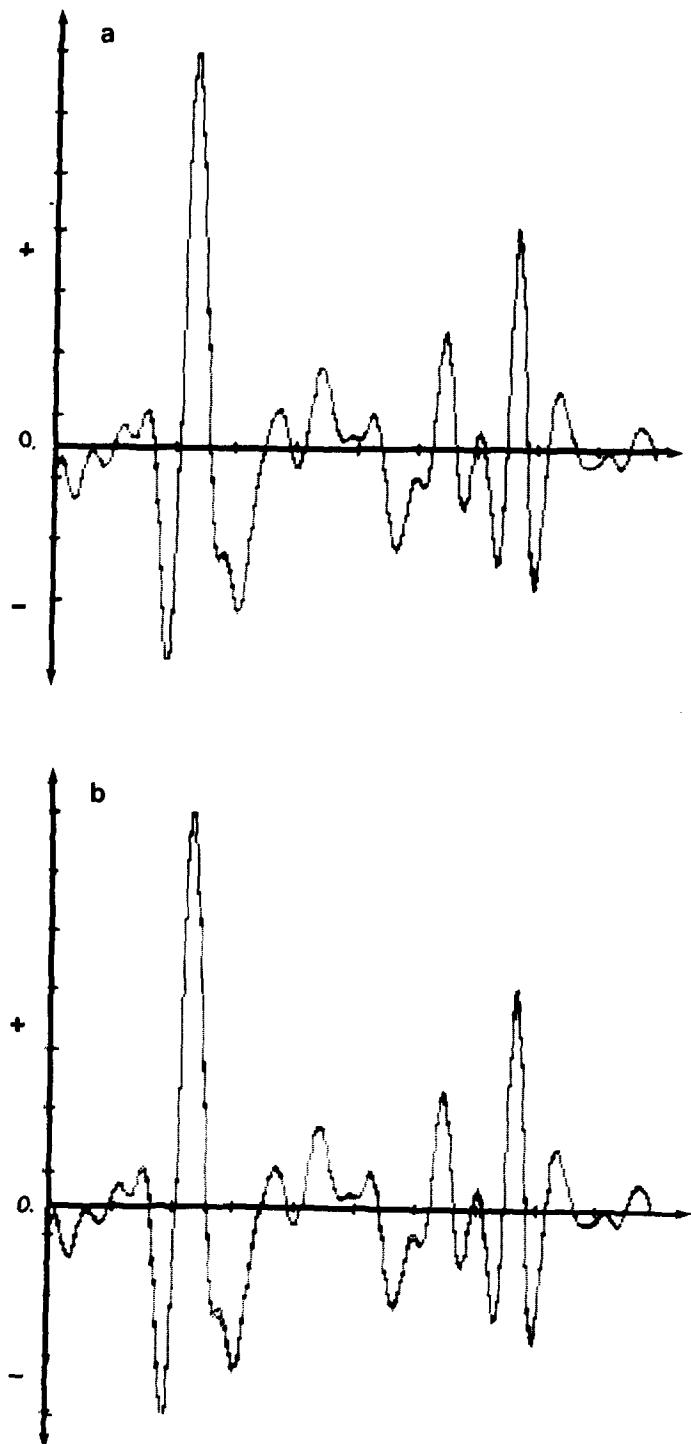


FIG. 10. Localizing the position of zero crossings. (a) illustrates the output of a D^2G operator of size $w_x = 8$ pixels convolved with an intensity profile, and reconstructed at five times the initial image resolution, using the ideal sinc function. In (b)-(d), the D^2G output has been reconstructed with a truncated sinc, linear interpolation, and Gaussian, respectively, and superimposed on the output of (a).

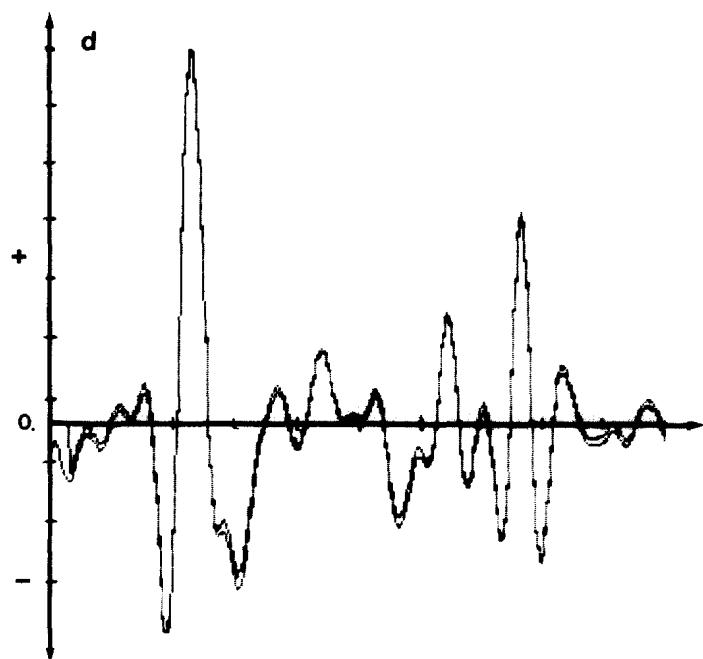
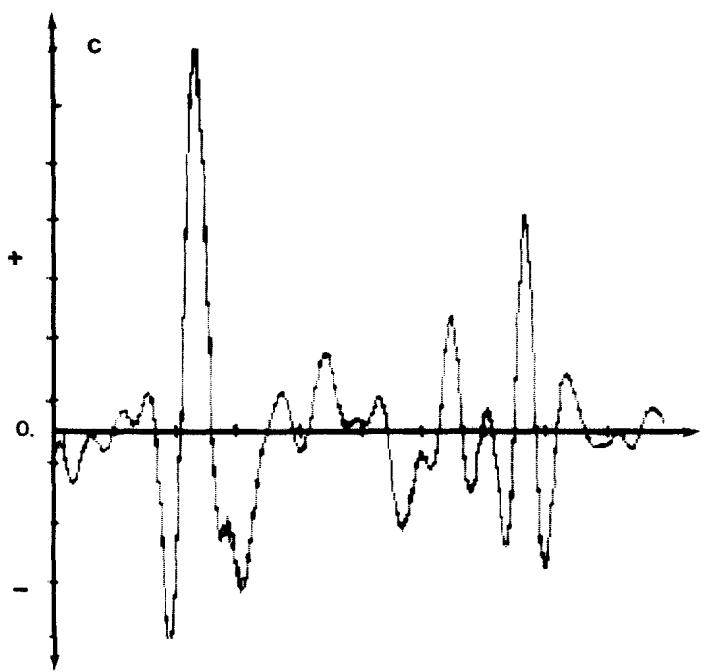


FIG. 10.—Continued.

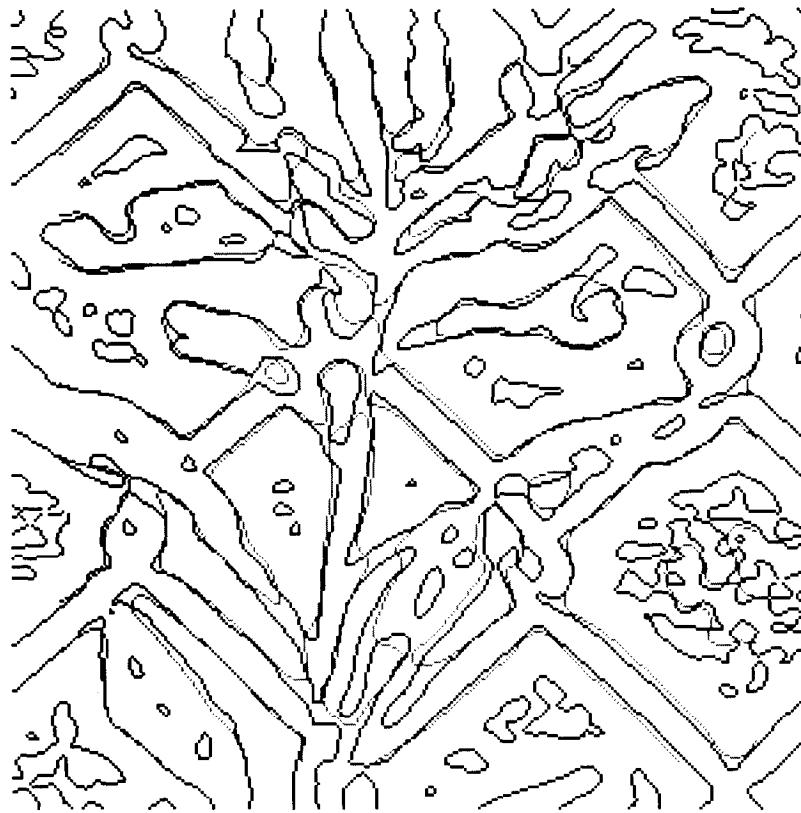


FIG. 11. The spatial coincidence of zero crossings from different size channels. The zero crossings obtained from the convolution of the first image of Fig. 2a with $\nabla^2 G$ operators of size $w = 9$ and 18 pixels are shown superimposed, with zero crossings of the smaller operator displayed in black, and zero crossings of the larger operator in light grey. Points at which the zero crossings exactly coincide are shown in medium grey.

It has long been known, through psychophysical experimentation, that there exist separate orientation sensitive channels through which visual information is processed, each tuned to different ranges of spatial frequency (for example [45–48]). There still remains considerable controversy over the number and frequency bandwidth of these channels. A recent quantitative study by Wilson and Bergen [49], suggests that at each location in the visual field, there are four fairly broadly tuned channels, whose shape is a difference of two Gaussians. The sizes of the underlying two-dimensional operators are separated by an octave, with a central diameter w ranging from roughly 4.5 to 30 min. of visual arc in the central fovea, and increasing linearly with distance from the fovea. These psychophysical channels correspond closely to the $\nabla^2 G$ operators [1, 50]. Studies of human visual acuity suggest that there is a smaller channel in the fovea, whose size w is roughly 2.1 min. of visual arc [51]. Finally, the use of zero-crossings allows for simple interpolation schemes to be used to provide human hyperacuity, which refers to the ability of the human visual system to accurately localize the position of an edge to a resolution finer than that of the foveal receptors. We noted previously that the $\nabla^2 G$ output is roughly linear across zero, so that simple interpolation schemes, such as linear interpolation, can accu-

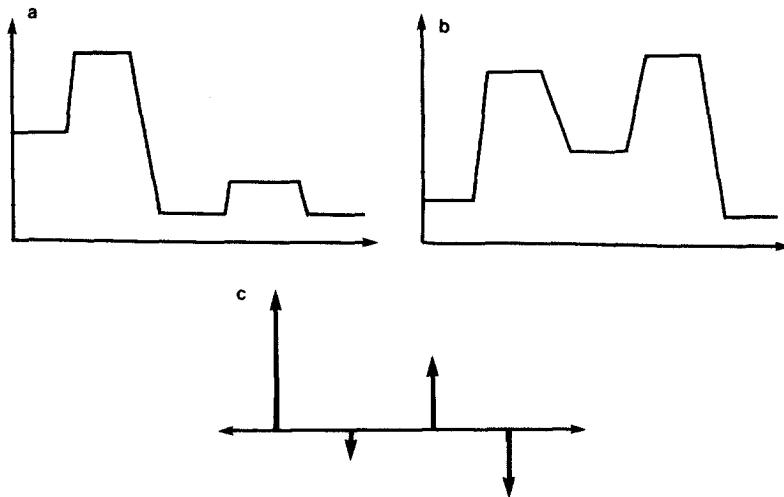


FIG. 12. A single operator size is not sufficient for computing the properties of an intensity change in the image. The two profiles of (a) and (b), though quite different in shape, yield zero crossings which are identical in position and slope, when convolved with a D^2G operator of size $w_x = 8$ pixels. (c) shows the resulting zero crossings; the arrows mark their position, with the sign and magnitude of the arrows representing the sign and slope of the zero crossings.

rately localize the position of a zero-crossing to a resolution finer than the image sampling [2, 51].

5. SUMMARY

We have presented a theory of the detection of intensity changes which suggests that processing begins with the convolution of the image with a range of operator sizes, whose shape is the Laplacian of a Gaussian ∇^2G . In the output of these operators, we detect the zero-crossings, and compute the local properties of slope and orientation. An important implication of this work is that it supports the idea that the initial detection stage can involve low-level processes, driven strictly from information in the image. Higher level processes are not necessary for the guidance of its operation. Second, these zero-crossing descriptions may be the lowest level description of the image from which later processes, such as stereopsis and motion analysis, may operate. That is, these later processes may not require access to the initial intensity information from which these descriptions were derived. Although these ideas were largely motivated by studies of the human visual system, they can be supported by computational arguments alone, and hence are relevant to early processing in computer vision systems.

ACKNOWLEDGMENTS

The author wishes to thank Shimon Ullman for valuable comments on this article. This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology.

REFERENCES

1. D. Marr and E. C. Hildreth, Theory of edge detection, *Proc. R. Soc. London Ser. B*, **207**, 1980, 187–217.

2. E. C. Hildreth, Implementation of a Theory of Edge Detection, M.I.T.A.I. Tech. Rep. TR-579, 1980.
3. L. G. Roberts, Machine perception of three dimensional solids, in *Optical and Electro-Optical Information Processing* (J. Tippett, D. Berkowitz, L. Clapp, C. Koester, and A. Vanderburgh, Eds.), M.I.T. Press, Cambridge, Mass., 1965.
4. A. Rosenfeld, R. Thomas, and Y. Lee, Edge and Curve Enhancement in Digital Pictures, Univ. of Maryland Tech. Rep., 93, 1969.
5. A. Herskovitz and T. O. Binford, On Boundary Detection, M.I.T.A.I. Memo, No. 183, 1970.
6. I. D. Macleod, Comments on Techniques for edge detection, *Proc. Inst. Electr. Eng.* **60**, 1972, 344.
7. B. K. P. Horn, The Binford-Horn Line-Finder, M.I.T.A.I Memo, No. 285, 1973.
8. E. Persoon, A new edge detection algorithm and its applications, *Computer Graphics and Image Processing* **5**, 1976, 425-446.
9. D. Marr, Early processing of visual information, *Philos. Trans. R. Soc. London Ser. B* **275**, 1976, 483-524.
10. T. O. Binford, Inferring surfaces from images, *Artif. Intell.*, 1981.
11. R. Nevatia and R. Babu, Linear feature extraction and description, *Computer Graphics and Image Processing* **13**, 1980, 257-269.
12. L. S. Davis and A. Mitiche, Edge detection in textures—maxima selection, *Computer Graphics and Image Processing* **16**, 1981, 158-165.
13. B. K. P. Horn, The Application of Fourier Transform Methods to Image Processing, M. S. Thesis, submitted to the Dept. of Electrical Engineering and Computer Science, M.I.T., Cambridge, Mass., 1968; MIT Artificial Intelligence Lab. Working Paper, No. 100, 1968.
14. K. C. Hayes and A. Rosenfeld, Efficient Edge Detectors and Applications, Univ. of Maryland Tech. Rep., No. 207, 1972.
15. H. Weschler and K. S. Fu, Image processing algorithms applied to rib boundary detection in chest radiographs, *Computer Graphics and Image Processing* **7**, 1978, 375-390.
16. L. Davis, A survey of edge detection techniques, *Computer Graphics and Image Processing* **4**, 1975, 248-270.
17. A. Rosenfeld and A. Kak, *Digital Picture Processing*, Academic Press, New York, 1976.
18. W. Pratt, *Digital Image Processing*, Wiley, New York, 1978.
19. A. Rosenfeld and M. Thurston, Edge and curve detection for visual scene analysis, *IEEE Trans. Comput.* **C-20**, 1971, 562-569.
20. A. Rosenfeld, M. Thurston, and Y. Lee, Edge and curve detection: further experiments, *IEEE Trans. Comput.* **C-21**, 1971, 677-715.
21. K. S. Shanmugam, F. M. Dickey, and J. A. Green, An optimal frequency domain filter for edge detection in digital pictures, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-1**, 1979, 37-49.
22. R. Leipnik, The extended entropy uncertainty principle, *Inf. Control* **3**, 1960, 18-25.
23. J. M. Brady and B. K. P. Horn, Rotationally Symmetric Operators for Surface Interpolation, M.I.T.A.I. Memo, No. 654, 1981.
24. W. S. Havens and J. C. Strikwerda, An Improved Operator for Edge Detection, in preparation.
25. J. M. Brady, Edge Orientation in Convolved Images, in preparation.
26. W. E. L. Grimson, *From Images to Surfaces: A Computational Study of the Human Early Visual System*, M.I.T. Press, Cambridge, Mass., 1981.
27. S. Ullman, *The Interpretation of Visual Motion*, M.I.T. Press, Cambridge, Mass., 1979.
28. P. J. MacVicar-Whelan and T. O. Binford, Intensity discontinuity location to subpixel precision, *Proc. Seventh International Joint Conference on Artificial Intelligence*, pp. 752-754, 1981.
29. B. O'Brien, Vision and resolution in the central retina, *J. Opt. Soc. Amer.* **41**, 1951, 882-894.
30. A. W. Snyder and W. H. Miller, Photoreceptor diameter and spacing for highest resolving power, *J. Opt. Soc. Amer.* **67**, 1977, 696-698.
31. P. Marks, Low-level vision using an array processor, *Computer Graphics and Image Processing* **14**, 1980, 281-292.
32. G. R. Nudd, S. D. Fouse, T. A. Nussmeier, and P. A. Nygaard, Development of custom-designed integrated circuits for image understanding, *Proc. Image Understanding Workshop*, November 1979, 1-9.
33. K. Nishihara and N. Larson, Towards a real time implementation of the Marr and Poggio stereo matcher, *Proc. Image Understanding Workshop*, April 1981, 114-120.
34. D. Marr and T. Poggio, A computational theory of human stereo vision, *Proc. R. Soc. London Ser. B* **204**, 1979, 301-328.

35. D. Marr and S. Ullman, Directional selectivity and its use in early visual processing, *Proc. R. Soc. London Ser. B* **211**, 1981, 151–180.
36. S. W. Kuffler, Discharge patterns and functional organization of the mammalian retina, *J. Neurophysiol.* **16**, 1953, 37–68.
37. R. W. Rodieck and J. Stone, Analysis of receptive fields of cat retinal ganglion cells, *J. Neurophysiol.* **28**, 1965, 833–849.
38. C. Enroth-Cugell and J. G. Robson, The contrast sensitivity of retinal ganglion cells of the cat, *J. Physiol. (London)* **187**, 1966, 517–552.
39. B. G. Cleland, M. W. Dubin, and W. R. Levick, Sustained and transient neurones in the cat's retina and lateral geniculate nucleus, *J. Physiol. (London)* **217**, 1971, 473–496.
40. F. M. deMonasterio, Properties of concentrically organized *X* and *Y* ganglion cells of macaque retina, *J. Neurophysiol.* **41**, 1978, 1394–1417.
41. F. M. deMonasterio, Center and surround mechanisms of opponent-color *X* and *Y* ganglion cells of retina macaques, *J. Neurophysiol.* **41**, 1978, 1418–1434.
42. K. Richter and S. Ullman, A model for the temporal organization of *X*-type and *Y*-type receptive fields in the primate retina, *Biol. Cybern.* **43**, 1982, 127–145; M.I.T.A.I. Memo, No. 573A, 1981.
43. D. H. Hubel and T. N. Wiesel, Receptive fields and functional architecture of monkey striate cortex, *J. Physiol. (London)* **195**, 1968, 215–243.
44. P. H. Schiller, B. L. Finlay, and S. F. Volman, Quantitative studies of single-cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields, *J. Neurophysiol.* **39**, 1976, 1288–1319.
45. F. W. Campbell and J. G. Robson, Applications of Fourier analysis to the visibility of gratings, *J. Physiol. (London)* **197**, 1968, 551–556.
46. C. Blakemore and F. W. Campbell, On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images, *J. Physiol.* **203**, 1969, 237–260.
47. J. D. Cowan, Some remarks on channel bandwidths for visual contrast detection, *Neurosci. Res. Program Bull.* **15**, 1977, 492–517.
48. N. Graham, Visual detection of aperiodic spatial stimuli by probability summation among narrow-band channels, *Vision Res.* **17**, 1977, 637–652.
49. H. R. Wilson and J. R. Bergen, A four mechanism model for threshold spatial vision, *Vision Res.* **19**, 1979, 19–32.
50. M. E. Jernigan and R. W. Wardell, Does the eye contain optimal edge detection mechanisms? *IEEE Trans. Syst., Man, Cybern. SMC-11*, 1981, 441–444.
51. D. Marr, T. Poggio, and E. C. Hildreth, Smallest channel in early human vision, *J. Opt. Soc. Amer.* **70**, 1980, 868–870.