MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 858 September, 1985

EDGE DETECTION

Ellen C. Hildreth

ABSTRACT. For both biological systems and machines, vision begins with a large and unwieldy array of measurements of the amount of light reflected from surfaces in the environment. The goal of vision is to recover physical properties of objects in the scene, such as the location of object boundaries and the structure, color and texture of object surfaces, from the two-dimensional image that is projected onto the eye or camera. This goal is not achieved in a single step; vision proceeds in stages, with each stage producing increasingly more useful descriptions of the image and then the scene. The first clues about the physical properties of the scene are provided by the *changes of intensity* in the image. The importance of intensity changes and edges in early visual processing has led to extensive research on their detection, description and use, both in computer and biological vision systems. This article reviews some of the theory that underlies the detection of edges, and the methods used to carry out this analysis.

© Massachusetts Institute of Technology 1985

Acknowledgments. This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. 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 N00014-80-C-0505.

TABLE OF CONTENTS

1. Introduction

2. The Detection of Intensity Changes

   2.1 The One-Dimensional Detection of Intensity Changes

   2.2 The Two-Dimensional Detection of Intensity Changes

3. Recovering Properties of the Physical World

1. INTRODUCTION

For both biological systems and machines, vision begins with a large and unwieldy array of measurements of the amount of light reflected from surfaces in the environment. The goal of vision is to recover physical properties of objects in the scene, such as the location of object boundaries and the structure, color and texture of object surfaces, from the two-dimensional image that is projected onto the eye or camera. This goal is not achieved in a single step; vision proceeds in stages, with each stage producing increasingly more useful descriptions of the image and then the scene. The first clues about the physical properties of the scene are provided by the *changes of intensity* in the image. For example, in Figure 1, the boundaries of the sculpture, the markings and bright highlights on its surface, and the shadows that the trees cast on the snow all give rise to spatial changes in light intensity. The geometrical structure, sharpness and contrast of these intensity changes convey information about the physical edges in the scene. The importance of intensity changes and edges in early visual processing has led to extensive research on their detection, description and use, both in computer and biological vision systems.

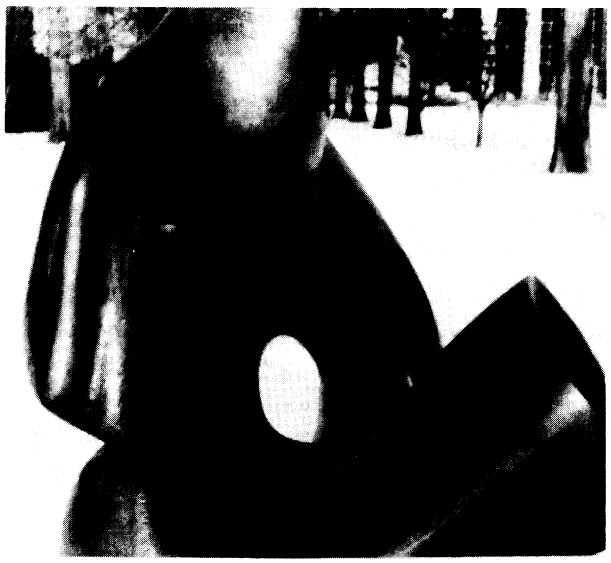


Figure 1. A natural image, exhibiting intensity changes due to many physical factors.

The process of edge detection can be divided into two stages: first, intensity changes in the image are detected and described; second, physical properties of edges in the scene are inferred from this image description. Section 2 concentrates on the first stage, about which more is known at this time. Section 3 briefly describes some areas of vision research that address the second stage. This article mainly reviews some of the theory that underlies the detection of edges, and the methods used to carry out this analysis. There is also some reference to studies of early processing in biological vision systems.

This article does not present a complete review of the edge detection literature; rather it introduces the reader to some of the basic issues that are considered central to the problem of edge detection.

2. THE DETECTION OF INTENSITY CHANGES

The most commonly used methods for detecting intensity changes incorporate three essential operations. First, the image intensities are either *smoothed* or approximated locally by a smooth analytic function. Second, the smoothed intensities are *differentiated*, using either a first or second derivative operation. Third, simple *features* in the result of this differentiation stage, such as peaks (positive and negative extrema) or zero–crossings (transitions between positive and negative values), are detected and described. This section first describes briefly the role of these operations in the detection of intensity changes and then presents in more detail, some of the methods used to carry out these operations.

The smoothing operation serves two purposes. First, it reduces the effect of noise on the detection of intensity changes. Second, it sets the resolution or scale at which intensity changes are detected. The sampling and transduction of light by the eye or camera introduces spurious changes of light intensity that do not correspond to significant physical changes in the scene. Smoothing of the intensities can remove these minor fluctuations due to noise. Figure 2a shows a one–dimensional intensity profile that is shown smoothed by a small amount in Figure 2b. Small variations of intensity, due in part to noise in the digitizing camera, do not appear in the smoothed intensities. Approximation of the intensity function by a smooth analytic function can serve the same purpose as a smoothing operation.

Significant changes in the image can also occur at multiple resolutions. Consider, for example, a leopard's coat. At a fine resolution, rapid fluctuations of intensity might delineate the individual hairs of the coat, while at a coarser resolution, the intensity changes might delineate only the leopard's spots. Changes at different resolutions can often be detected by smoothing the image intensities by different amounts. Figure 2c illustrates a more extensive smoothing of the intensity profile of Figure 2a, which preserves only the gross changes of intensity.

The differentiation operation accentuates intensity changes and transforms the image into a representation from which properties of these changes can be extracted more easily. A significant intensity change gives rise to a peak in the first derivative or a zero–crossing in the second derivative of the smoothed intensities, as illustrated in Figures 2d and 2e, respectively. These peaks or zero–crossings can be detected straightforwardly and properties such as the position, sharpness and height of the peaks capture the location, sharpness and contrast of the intensity changes in the image. The detection and description of these features in the smoothed and differentiated image provides a compact representation that captures meaningful information in the image. Marr (1) called this representation the *Primal Sketch* of the image. Later processes, such as binocular stereo, motion measurement and texture analysis, whose goal is to recover the physical properties of the scene, may then operate directly on this description of image features.

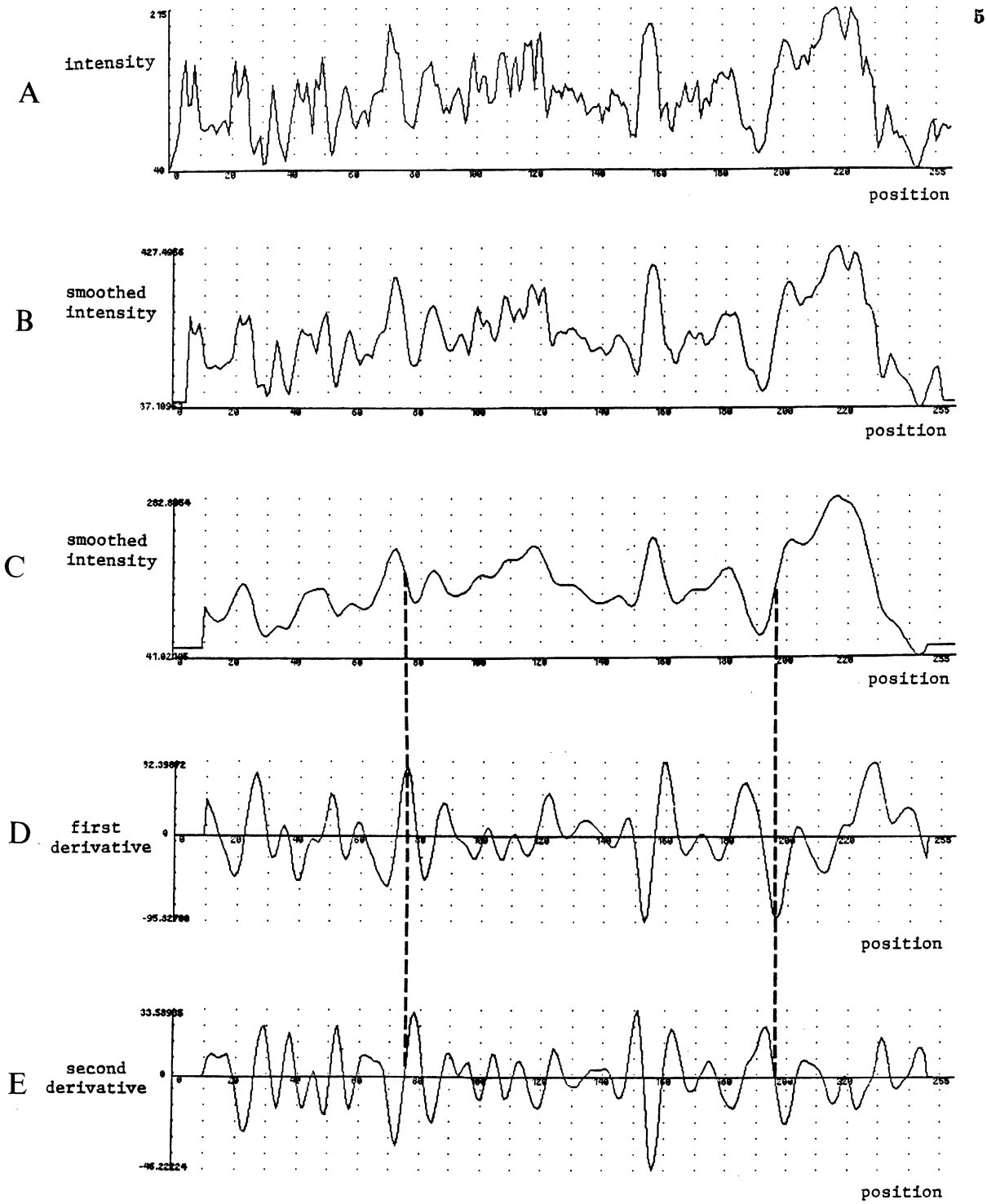


Figure 2. Detecting Intensity Changes. (a) One-dimensional intensity profile; the intensities along a horizontal scan line in an image are represented as a graph. (b) The result of smoothing the profile in (a). (c) The result of additional smoothing of (a). (d) and (e) The first and second derivatives, respectively, of the smoothed profile shown in (c). The vertical dashed lines indicate the peaks in the first derivative and zero-crossings in the second derivative that correspond to two significant intensity changes.

2.1 THE ONE-DIMENSIONAL DETECTION OF INTENSITY CHANGES

The theory that underlies the detection of intensity changes in two-dimensional images is based heavily on the analysis of one-dimensional signals. This section discusses three topics that have been addressed in this analysis: (1) the design of *optimal* operators for performing smoothing and differentiation, (2) the *information content* of the description of signal features such as zero-crossings, and (3) the relationship between features that are detected at *multiple resolutions*. Studies of these issues have used a variety of theoretical approaches that appear to yield similar conclusions.

Some of the early methods for detecting intensity changes incorporated only limited smoothing of the intensities and performed the differentiation by taking first or second differences between neighboring image elements (examples of this early work can be found in (2-8)). In one dimension, this is equivalent to performing a convolution of the intensity profile with operators of the type shown on the left in Figures 3b and 3c. Additional smoothing can be performed by increasing the spatial extent of these operators.

The operators in Figures 3b and 3c contain step-like changes. Other studies have employed Gaussian smoothing of the image intensities (for example, 9-13). Combined with the first and second derivative operations, Gaussian smoothing yields convolution operators of the type shown in Figures 3d and 3e. Several arguments have been put forth in support of the use of Gaussian smoothing. Marr and Hildreth (11, 12) argued that the smoothing function should have both limited support in space and limited bandwidth in frequency. In general terms, a limited support in space is important because the physical edges to be detected are spatially localized. A limited bandwidth in frequency provides a means of restricting the range of scales over which intensity changes are detected, which is sometimes important in applications of edge detection. The Gaussian function minimizes the product of bandwidths in space and frequency. The use of smoothing functions that do not have limited bandwidths in space and frequency can sometimes lead to poorer performance, reflected in a greater sensitivity to noise, the false detection of edges that do not exist, or a poor ability to localize the position of edges (see, for example, 11, 14).

Shanmugam, Dickey and Green (15) derived an optimal frequency domain filter for detecting intensity changes, using the criteria that the filter: (1) yields maximum energy in the vicinity of an edge in the image, (2) has limited frequency bandwidth, (3) yields a small output when the input is constant or slowly varying, and (4) is an even function in space. For the special case of detecting step changes of intensity, the optimal frequency domain filter corresponds to a spatial operator that is approximately the second derivative of a Gaussian (for a given bandwidth) shown in Figure 3e.

In a later study, Canny (14) used the following criteria to derive an optimal operator: (1) good detection ability, that is, there should be low probabilities of failing to detect real edges and falsely detecting edges that do not exist, (2) good localization ability, that is, the position of the detected edge should be as close as possible to the true position of the edge, and (3) uniqueness of detection, that is, a given edge should be detected only once. The first two criteria are related by an uncertainty principle; as detection ability increases, localization ability decreases, and vice versa. The analysis also assumed

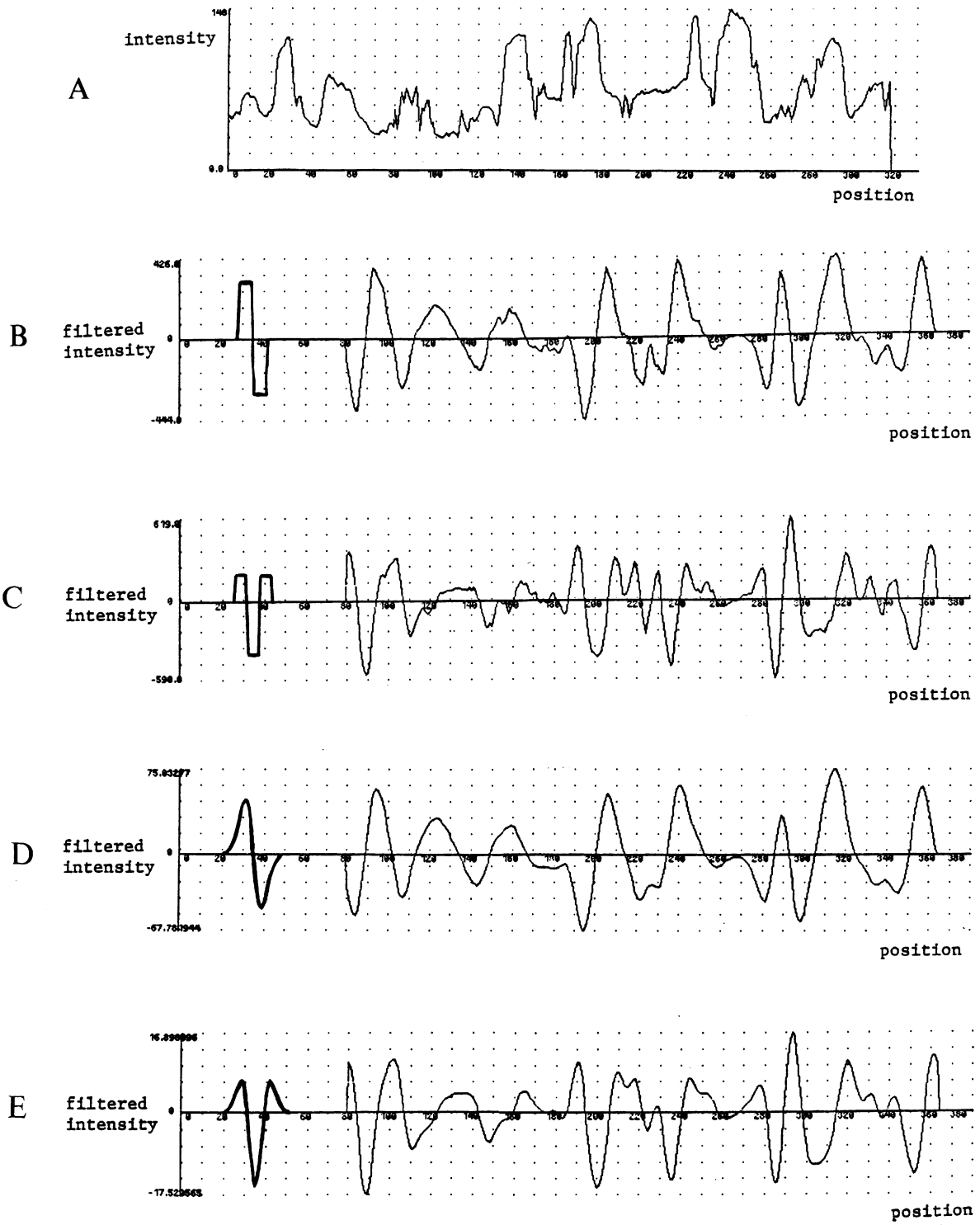


Figure 3. Smoothing and Differentiation. (a) A one-dimensional intensity profile. (b) A first difference operator is shown in heavy lines on the left and the result of its convolution with the profile in (a) is shown on the right. (c) A second difference operator is shown in heavy lines on the left and its convolution with (a) on the right. (d) The first derivative of a Gaussian (left) and its convolution with (a) (right). (e) The second derivative of a Gaussian (left) and its convolution with (a) (right). (Note that in (b) through (e), the position of the filtered intensity profile is shifted relative to the position of the original intensity profile shown in (a).)

that extrema in the output of the operator indicate the presence of an edge. For the particular case in which an "edge" is defined as a step change of intensity, the operator that optimally satisfies these criteria is a linear combination of four exponentials, which can be approximated closely by the first derivative of a Gaussian shown in Figure 3d.

Poggio, Voorhees and Yuille (16) and Torre and Poggio (17) derived an optimal smoothing operator, using the tools of *regularization theory* from mathematical physics. They began with the observation that numerical differentiation of the image is a mathematically *ill-posed* problem (18), because its solution does not depend continuously on the input intensities (this is equivalent to saying that the solution is not robust against noise). The smoothing operation serves to *regularize* the image, making the differentiation operation mathematically well-posed. In the case where the image intensities are assumed to contain noise, the following method was used to regularize the image. First, let $I(x)$ denote the continuous intensity function, which is sampled at a set of discrete locations x_k, $1 \leq k \leq n$, and let $S(x)$ denote the smoothed intensity function to be computed. It was assumed that $S(x)$ should both fit the sampled intensities as closely as possible and be as smooth as possible. Using the tools of regularization theory, this was formulated as the computation of the function $S(x)$ that minimizes the following expression:

$$\sum_{k=1}^{n}(I(x_k) - S(x_k))^2 + \lambda \int \|S''(x)\|^2 dx.$$

The first term measures how well $S(x)$ fits the sampled intensities and the second term measures the smoothness of $S(x)$. The constant λ controls the trade-off between these two measures. Poggio, Voorhees and Yuille showed that the solution to this minimization problem is equivalent to the convolution of the image intensities with a cubic spline that is very similar to the Gaussian. Torre and Poggio (17) further expanded upon the theoretical properties of a broad range of smoothing filters, from the perspective of regularizing the image intensities for differentiation.

Another approach to the smoothing stage is to find an analytic function that best models or approximates the local intensity pattern. An early representative of this approach was the Hueckel operator (5, 7). Surface-fitting methods used a variety of basis functions to perform the approximation, including planar functions (19) and quadratic functions (20). More recently, Haralick (21, 22) used the discrete Chebychev polynomials to approximate the image intensities. In these methods, a differentiation operation is then performed analytically on the polynomial approximation of the intensity function. The method of approximation used by Haralick (21, 22) is roughly equivalent to smoothing the image by convolution with spatial operators such as those derived by Canny (14) and Poggio, Voorhees and Yuille (16). A rigorous comparison between the performance of surface-fitting versus direct smoothing methods has not yet been made.

A second issue that bears on the choice of operator for the smoothing and differentiation stages is the *information content* of the subsequent description of image features. That is, to what extent does a representation of only the significant changes of intensity capture all of the important information in an image? This question led to a number of theoretical studies of the reconstruction of a signal from features such as its zero-crossings. Although the goal of vision is not to *reconstruct* the visual image, these results

are important because they suggest that an image can be transformed into a compact representation of its features with little loss of information.

An early study by Logan (23) that interested many vision researchers addressed the information content of the zero-crossings of a signal. Logan proved that if a signal has (1) a frequency bandwidth of less than one octave and (2) no zeros in common with its Hilbert transform, then the signal can be entirely reconstructed from the positions of its zero-crossings, up to a multiplicative constant. The second condition is almost always satisfied for physical signals. This result has also been extended to two dimensions (1). This analysis is interesting, because it shows that the zero-crossings of a signal are very rich in information. Its direct relevance to vision is limited, however, because the initial smoothing and differentiation of an image is typically performed by operators that are not one-octave bandpass in frequency.

Other studies have addressed the information content of features of signals that are more relevant to visual processing. For example, Yuille and Poggio (24) proved some interesting results regarding the zero-crossings (or more generally, the level-crossings*) of an image that is convolved with the second derivative of a Gaussian, over a continuous range of scales. Before stating the results, we introduce the *scale-space* representation of zero-crossings used by Witkin (25), illustrated in Figure 4. First, let the one-dimensional Gaussian function be defined as follows, where σ is the standard deviation of the Gaussian:

$$G(x) = \frac{1}{\sigma} e^{-\frac{x^2}{2\sigma^2}}.$$

The second derivative of the Gaussian function is then given by the following expression:

$$G''(x) = \frac{d^2G(x)}{dx^2} = \frac{1}{\sigma^3}\left(\frac{x_2}{\sigma^2} - 1\right)e^{-\frac{x^2}{2\sigma^2}}.$$

Suppose that a one-dimensional signal $I(x)$ is convolved with $G''(x)$ for a continuous range of standard deviations σ and the positions of the zero-crossings are marked for each size or scale. Figure 4 shows an intensity profile (Figure 4a) that is convolved with a $G''(x)$ function with large σ (Figure 4b). The positions of the zero-crossings are marked with heavy dots. In the scale-space representation of Figure 4c, the vertical dimension represents the value of σ and the horizontal dimension represents position in the signal. For each value of σ, the positions of the zero-crossings of $I(x) * G''(x)$ are plotted as points along a horizontal line in this diagram. For example, points along the dashed line at $\sigma = \sigma_l$ indicate the positions of the zero-crossings of the signal in Figure 4b. The scale-space representation of zero-crossings illustrates the behavior of these features across scales. For small σ, the zero-crossings capture all of the changes in the original intensity function. At coarser scales (larger σ), the positions of the zero-crossings capture only the gross changes of intensity.

The scale-space representation is visually suggestive of a *fingerprint.* In fact, in much the same way that a fingerprint uniquely identifies a person, the scale-space representation may uniquely identify an image. Yuille and Poggio (24) proved that for almost all one-dimensional signals, the scale-space map of the zero-crossings of the signal convolved with $G''(x)$ over a continuum of scales determines the signal uniquely, up

*The level-crossings of a signal are the points at which a value v is crossed by the signal, where v may be non-zero.

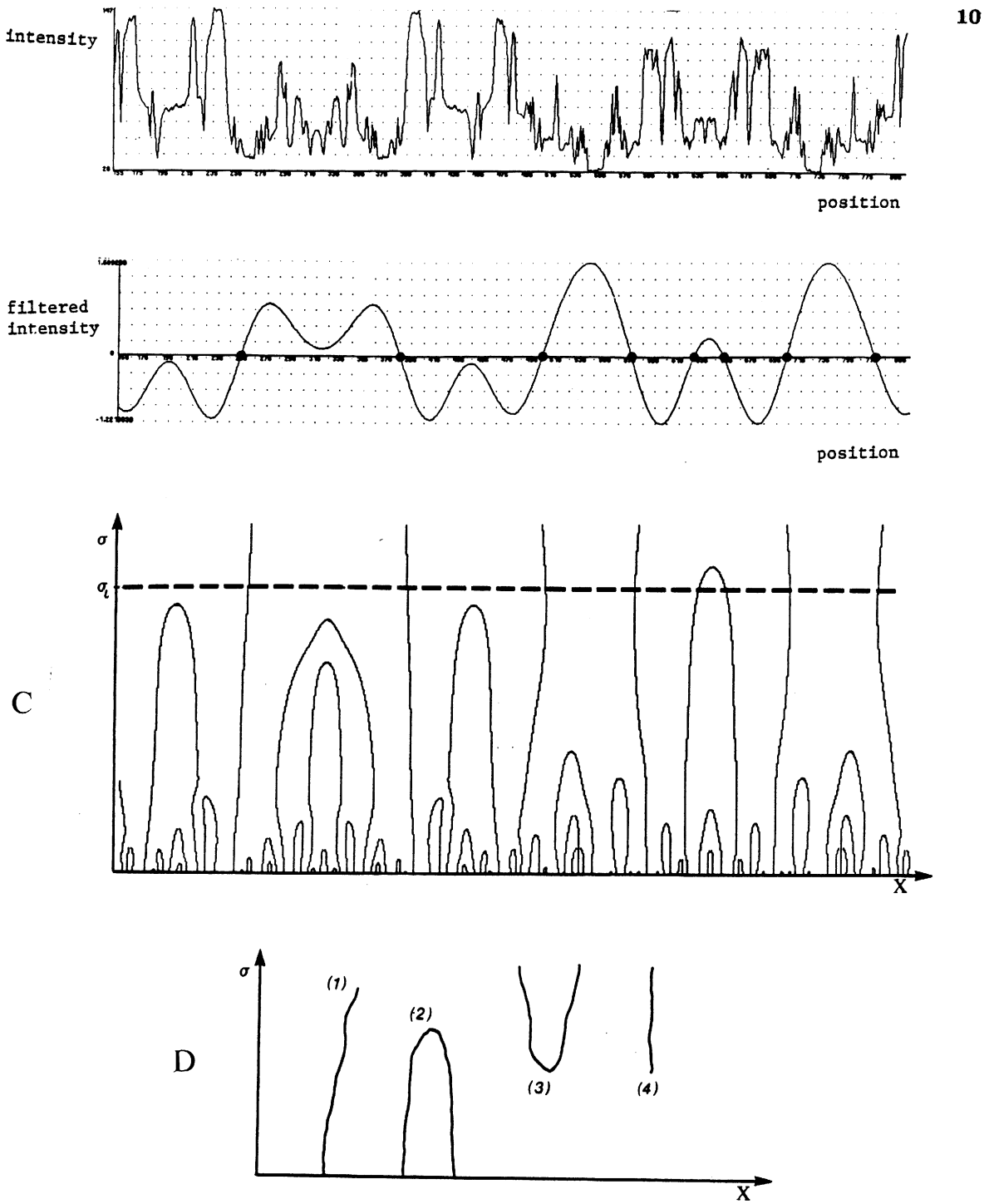


Figure 4. The Scale-Space Representation. (a) An extended one-dimensional intensity profile. (b) The result of convolving the profile in (a) with a $G''(x)$ operator with large σ. The zero-crossings are marked with heavy dots. (c) The scale-space representation of the positions of the zero-crossings over a continuous range of scales (sizes of σ). The zero-crossings of (b) are plotted along the dashed horizontal line at $\sigma = \sigma_l$. (d) Contours of the type labelled (1) and (2) are commonly found in the scale-space representation, while those of the type labelled (3) and (4) are never found.

to a multiplicative constant and an additional harmonic function. The proof provides a method for reconstructing a signal $I(x)$ from knowledge of how the zero-crossings of $I(x) * G''(x)$ change across scales. The use of Gaussian smoothing is critical to the completeness of the subsequent feature representation, but the basic theorem applies to zero-crossings and level-crossings of the result of applying any linear differential operator to the Gaussian-filtered signal. Yuille and Poggio also derived a two-dimensional extension to this result.

Careful observation of the contours in the scale-space representation of Figure 4c reveals that the contours either begin at the smallest scale and continue as a single, isolated contour through larger scales as shown in Figure 4d(1), or they form closed, inverted bowl-like shapes as shown in Figure 4d(2). Additional zero-crossings are never created as scale increases; that is, there are no contours in the scale-space representation of the type shown in Figures 4d(3) and 4d(4). This observation has been supported by a number of theoretical studies (26-28), which have also shown that the Gaussian function is the only smoothing function that yields this behavior of subsequent features across scale. This observation applies to zero-crossings and level-crossings of the result of applying any linear differential operator to the Gaussian-smoothed signal. This behavior of features across scale has been exploited successfully in the qualitative analysis of one-dimensional signals (25).

To summarize, the analysis of one-dimensional signals has been important for developing a solid theoretical foundation on which to base methods for detecting intensity changes in an image. Several theoretical studies attempted to derive an optimal operator for detecting intensity changes, using a variety of criteria for evaluating the performance of the operator. All of these operators essentially perform a smoothing and differentiation of the image intensities. Furthermore, the one-dimensional analyses all point to operators whose spatial shape is roughly the first or second derivative of a Gaussian function. Mathematical studies also addressed the information content of representations of image features and the behavior of these features across multiple scales. These latter studies also stressed the importance of Gaussian smoothing.* Interestingly, the initial filters in the human visual system also appear to perform a spatial convolution of the image with a function that is closely approximated by the second derivative of a Gaussian (29). It is also well known that the human visual system initially analyzes the retinal image through a number of spatial filters that differ in the amount of smoothing that is performed in space and in time (29).

2.2 THE TWO-DIMENSIONAL DETECTION OF INTENSITY CHANGES

The problems that were addressed in the one-dimensional analysis of intensity signals also arise for the detection of intensity changes in two-dimensional images, although their solution is more complex. The design of optimal operators for performing the

*It should be noted again that some edge detection methods that perform an analytic approximation of the intensity function may be equivalent to those performing a direct smoothing operation with a Gaussian function.

smoothing and differentiation stages, for example, is complicated by a larger selection of possible derivative operations that can be performed in two dimensions. Many of the mathematical results regarding the information content of image features and behavior of features across scale have been extended to two dimensions, but the algorithms for extracting and describing these features in the image are also more complex than their one-dimensional counterparts. This section reviews some of the techniques used to detect and describe intensity changes in two-dimensional images.

Early work on edge detection primarily used directional first and second derivative operators for performing the two-dimensional differentiation (2–10, 19, 20, 30–32). A change of intensity that is extended along some orientation in the image gives rise to a peak in the first derivative of intensity taken in the direction perpendicular to the orientation of the intensity change, or a zero-crossing in the second directional derivative. The simplest directional operators are formed by extending one-dimensional cross-sections such as those shown in Figure 3 along some two-dimensional direction in the image. Directional operators have differed in the shape of their cross-sections, both perpendicular to and along their primary orientations. Macleod (9) and Marr and Poggio (10), for example, used directional derivatives that embodied Gaussian smoothing.

In principle, the computation of the derivatives in two directions, such as the horizontal and vertical directions, is sufficient to detect intensity changes at all orientations in the image. Several algorithms, however, use directional operators at a large number of discrete orientations (for example, see (4, 7, 8, 14, 32)). A given intensity change is detected by a number of directional operators in this case and the output of the directional operator that yields the largest response is typically used to describe the local intensity change. Two examples of algorithms of this type are those of Nevatia and Babu (32) and Canny (14). An example of the results of Canny's algorithm is shown in Figure 5. The contours of Figure 5b represent only the positions of the significant intensity changes in Figure 5a.

Other related differential operators that are used in two dimensions are the first and second derivatives in the direction of the gradient of intensity (14, 17, 22). The intensity gradient, defined as follows:

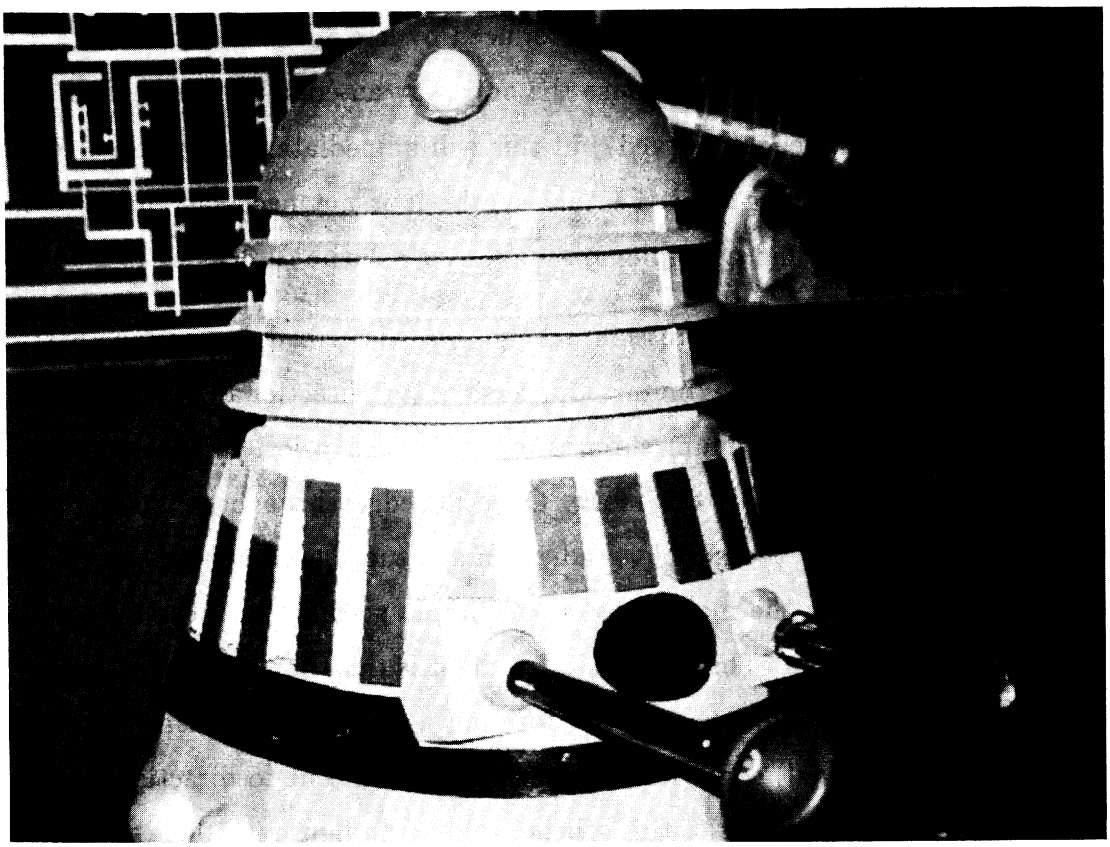
$$\nabla^2 I = (\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y})$$

is a vector that indicates the direction and magnitude of steepest increase in the two-dimensional intensity function. Let $\mathbf{n}$ denote the unit vector in the direction of the gradient. The differential operators $\frac{\partial}{\partial \mathbf{n}}$ and $\frac{\partial^2}{\partial \mathbf{n}^2}$ are non-directional operators, in the sense that their value does not change when the image is rotated. They are also nonlinear operators, and unlike the linear differential operators, cannot be combined with the smoothing function in a single filtering step. Methods such as those of Nevatia and Babu (32) and Canny (14) essentially use the directional derivative along the gradient for extracting features.

A second non-directional operator that is used for detecting intensity changes is the Laplacian operator, ∇^2 (1, 5, 11–13, 15, 33):

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}.$$

A



B

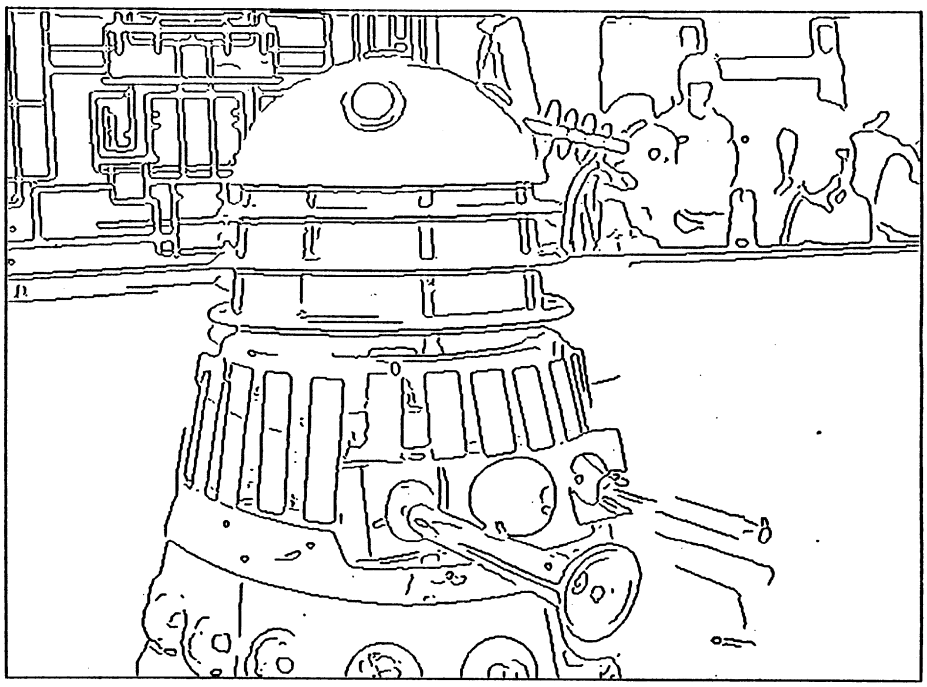


Figure 5. Canny's Edge Detection Algorithm. (a) A natural image. (b) The positions of the intensity changes detected by Canny's algorithm. *(Courtesy of J. F. Canny.)*

Combined with a two-dimensional Gaussian smoothing function,

$$G(r) = \frac{1}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}},$$

the Laplacian yields the function $\nabla^2 G$ given by the following expression:

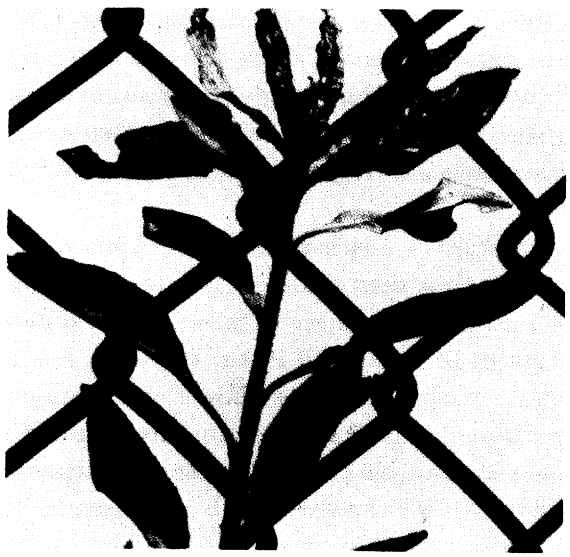
$$\nabla^2 G = \frac{1}{\sigma^2}\left[\frac{r^2}{\sigma^2} - 2\right] e^{-\frac{r^2}{2\sigma^2}}.$$

r denotes the distance from the center of the operator and σ is the standard deviation of the two-dimensional Gaussian. The $\nabla^2 G$ function is shaped something like a mexican hat in two dimensions. Figure 6 shows an example of the convolution of an image (Figure 6a) with a $\nabla^2 G$ operator (Figure 6b). The Laplacian is a non-directional second derivative operation; the elements in the output of the Laplacian that correspond to the location of intensity changes in the image are therefore the zero-crossings. The zero-crossing contours derived from Figure 6b are shown in Figure 6c. In this case, the zero-crossing contours were located by detecting the transitions between positive and negative values in the filtered image, by scanning in the horizontal and vertical directions.* A single convolution of the image with the non-directional $\nabla^2 G$ operator allows the detection of intensity changes at all orientations, for a given scale. The two-dimensional orientation of a local portion of the zero-crossing contour can be computed from the gradient of the filtered image (12).

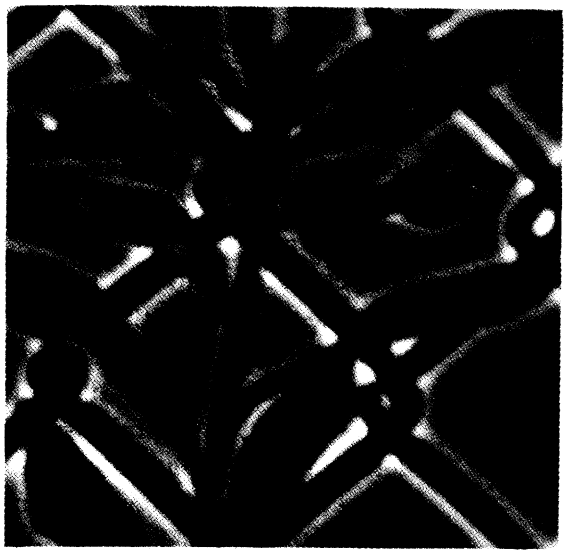
It is not yet clear whether directional or non-directional operators are most appropriate for detecting intensity changes. Both have advantages and disadvantages. The use of the Laplacian is simpler and requires less computation than the use of either directional derivatives or derivatives in the direction of the gradient. The directional operators, however, yield somewhat better localization of the position of intensity changes (14, 22), particularly in areas where the orientation of an edge is changing rapidly in the image (34, 35). Features such as the zero-crossing contours, when derived with non-directional operators, generally form smooth, closed contours, while features obtained with directional operators generally do not have such special geometric properties (17). Marr and Hildreth (11) showed that if the intensity function along the orientation of an intensity change varies at most linearly, then the zero-crossings of the Laplacian exactly coincide with the zero-crossings of a directional operator taken in the direction perpendicular to the orientation of the intensity change. Torre and Poggio (17) characterized more formally, the relationship between the zeros of the Laplacian and those of the second derivative in the direction of the gradient, in terms of the geometry of the two-dimensional intensity surface. With regard to the use of directional versus non-directional derivative operators, it is interesting to note that physiological studies reveal that the retina analyzes the visual image through a circularly-symmetric filter whose spatial shape is given by the difference of two Gaussian functions (see, for example, 36, 37), which is closely approximated by the $\nabla^2 G$ function.

Mathematical results regarding the information content and behavior across scales of image features have some bearing on the choice of differential operators. For example, Yuille and Poggio (28) showed that in two dimensions, the combination of Gaussian

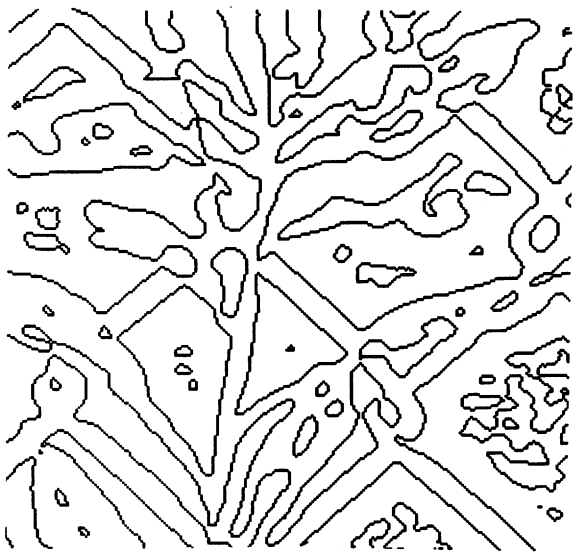
*The design of robust methods for detecting zero-crossings remains an open area of research in edge detection.



A



B



C

Figure 6. Detecting Intensity Changes with the $\nabla^2 G$ Operator. (a) A natural image. (b) The result of convolving the image with a $\nabla^2 G$ operator. The most positive values are shown in white and most negative values in black. (c) The zero-crossings of the convolution output shown in (b).

smoothing with any linear differential operator yields zero-crossings or level-crossings that behave well with increasing scale, in that no features are created as the size of the Gaussian is increased. In the case of the second derivative along the gradient, Yuille and Poggio proved that there is no smoothing function that avoids the creation of zero-crossings with increasing scale. The completeness of the scale-space representation of zero-crossings or level-crossings in two dimensions also requires the use of linear differential operators (24).

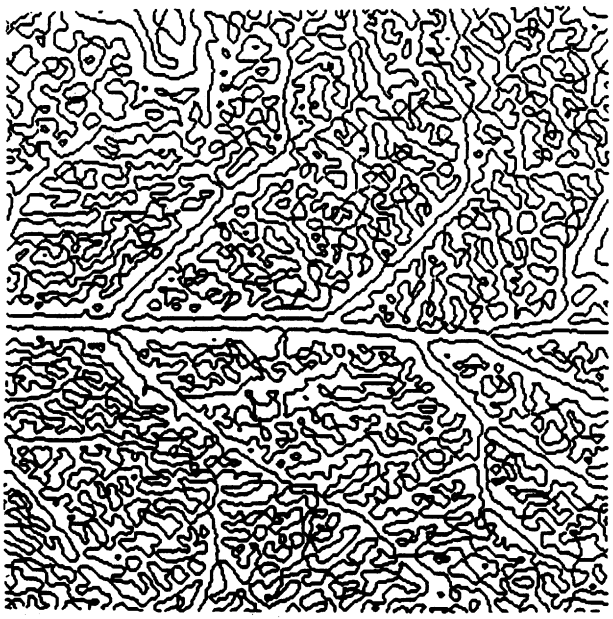
The analysis of intensity changes across multiple scales is a difficult problem that has not yet found a satisfactory solution. There is a clear need to detect intensity changes at multiple resolutions (2). Important physical changes in the scene take place at different scales. Spatial filters that allow the description of fine detail in the intensity function generally miss coarser structures in the image, while those that allow the extraction of coarser features generally smooth out important detail. At all resolutions, some of the detected features may not correspond to real physical changes in the scene. For example, at the finest resolutions, some of the detected intensity changes may be a consequence of noise in the sensing process. At coarser resolutions, spurious image features might arise as a consequence of smoothing together nearby intensity changes. The problems of sorting out the relevant changes at each resolution and combining them into a representation that can be used effectively by later processes are difficult and unsolved problems. We mention here some of the research that has attempted to address these problems.

Marr and Hildreth (11) explored the combination of zero-crossing descriptions that arise from convolving an image with $\nabla^2 G$ operators of different size. An example of these descriptions is illustrated in Figure 7. The zero-crossings from the smaller $\nabla^2 G$ operator primarily detect the bumpy texture on the surface of the leaf, whereas the zero-crossing contours from the larger operator also outline some of the highlights on the leaf surface that are due to changing illumination (the arrows point to one example). Marr and Hildreth suggested the use of *spatial coincidence* of zero-crossings across scale as a means of indicating the presence of a real edge in the scene. Strong edges such as object boundaries often give rise to sharp intensity changes in the image that are detected across a range of scales and in roughly the same location in the image. In the one-dimensional scale-space representation*, these edges give rise to roughly vertical lines. The existence of contours in the scale-space representation that are roughly vertical and extend across a range of scales could be used to infer the presence of a significant physical change at the corresponding location in the scene.

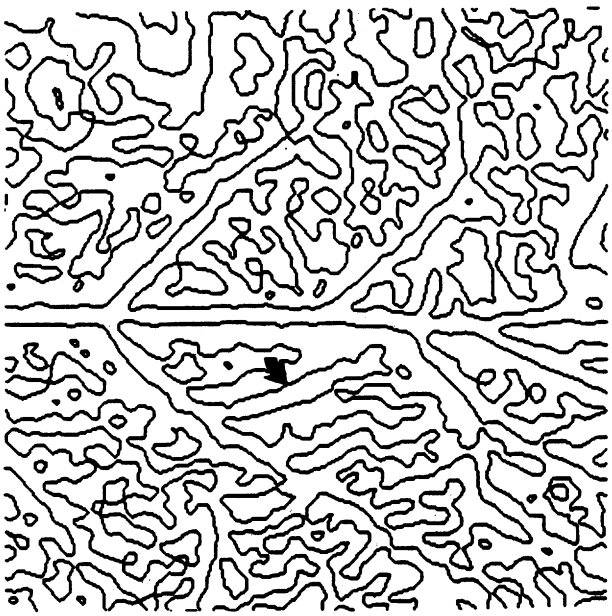
Witkin (25) developed a method for constructing qualitative descriptions of one-dimensional signals that uses the scale-space representation. The method embodied two basic assumptions: (1) the *identity* assumption, that zero-crossings detected at different scales, which lie on a common contour in the scale-space description, arise from a single physical event; and (2) the *localization* assumption, that the true location of a physical event that gives rise to a contour in the scale-space description is the contour's position as σ tends to zero. Coarser scales were used to identify important events in the signal and finer scales used to localize their position. Events that persisted over large changes in scale

*The scale-space representation can be extended to two dimensions, in which the positions of the zero-crossings on the $x - y$ plane are represented across multiple operator sizes.

A



B



C

Figure 7. Multiple Operator Sizes. (a) A natural image. (b) and (c) The zero-crossings that result from convolving the image with $\nabla^2 G$ operators whose central positive region has a diameter of 6 and 12 image elements, respectively. The arrows in (a) and (c) indicate a highlight in the image that is detected by the larger operator.

also had special significance. Witkin's method, called *scale-space filtering*, begins with the scale-space description and collapses it into a discrete tree structure that represents the qualitative behavior of the signal. Some of the heuristics embodied in this analysis may be useful for analyzing two-dimensional images.

Canny (14) used a different approach to combining descriptions of intensity changes across multiple scales. Features were first detected at a set of discrete scales. The finest scale description was then used to predict the results of the next larger scale, assuming that the filter used to derive the larger scale description performs additional smoothing of the image. In a particular area of the image, if there was a substantial difference between the actual description at the larger scale and that predicted by the smaller scale, then it was assumed that there is an important change taking place at the larger scale that is not detected at the finer scale. In this case, features detected at the larger scale were then added to the final feature representation. Empirically, Canny found that most features were detected at the finest scale and relatively few were added from coarser scales.

Poggio, Voorhees and Yuille (16) have also begun to explore the issue of detecting intensity changes across scales, using the methods of regularization theory. Recall that their approach was to find a smoothed intensity function $S(x)$, given the sampled intensities $I(x_k)$, which minimizes the following expression:

$$\sum_{k=1}^{n}(I(x_k) - S(x_k))^2 + \lambda \int \|S''(x)\|^2 dx.$$

The parameter λ controls the scale at which intensity changes are detected. That is, if λ is small, $S(x)$ closely approximates $I(x_k)$, and as λ increases, $S(x)$ becomes increasingly more smooth. Regularization theory may suggest methods for choosing the optimal λ for a given set of data, which may be useful for analyzing changes across multiple scales (16).

To summarize, there has been considerable progress on the detection and description of intensity changes in two-dimensional images, but there still exists many open questions. A large body of theoretical and empirical work has addressed the question of what operators are most appropriate for performing the smoothing and differentiation stages. Emerging from this work is a better understanding of the advantages and disadvantages of various operators and the relationship between alternative approaches. It is unlikely that a single method will be most appropriate for all tasks. The choice of operators depends in part on the application, the nature of the later processes that use the description of image features, and the available computational resources. Some interesting work has begun to address the problem of detecting and integrating intensity changes across multiple scales, but a satisfactory solution to this problem still eludes vision researchers. A problem that was not discussed here is the computation of properties such as contrast and sharpness of the intensity changes. There has been some work on this problem, but it has not yet received a rigorous analytic treatment.

3. RECOVERING PROPERTIES OF THE PHYSICAL WORLD

In the introduction, it was noted that the goal of vision is to recover the physical properties of objects in the scene, such as the location of object boundaries and the structure,

color and texture of object surfaces, from the two-dimensional image that is projected onto the eye or camera. The detection of intensity changes in the image represents only a first, meager step toward achieving this goal. This section briefly mentions some of the areas of vision that address the recovery of physical properties of edges in the scene.

The property of edges that is perhaps most important and most studied is their three-dimensional structure. The structure of edges is conveyed through many sources. For example, the relative locations of corresponding edges in left and right stereo views conveys information about the location of the edges in three-dimensional space, and relative movement between edges in the image can be used to assess their relative position in space. Three-dimensional structure can also be inferred from the shape of the two-dimensional projection of edge contours, the way in which edges intersect in the image, and variations in surface texture. These latter cues are essential in the interpretation of structure from a single, static photograph. Many algorithms that analyze these sources are *feature-based*, in that the initial inferences regarding three-dimensional structure are made at the locations of features such as significant intensity changes in the image. Discussion of some of these processes for recovering three-dimensional structure can be found, for example, in (1, 5, 7, 10, 13, 27, 30, 31, 38–40).

Another important property of edges is the type of physical change from which they arise. For example, edges might be the consequence of object boundaries, changes in surface orientation, shadows, highlights or light sources, surface markings, changes in surface reflectance or material composition, and so on. Ultimately, it is necessary to determine the physical source of each edge in the scene. While some interesting work has been done in these areas, there remain many open problems (examples can be found in (1, 5, 7, 13, 30, 31, 38–41)). The recovery of these physical properties of edges is likely to be a main focus of future research on edge detection.

Acknowledgements: The author wishes to thank Tomaso Poggio for valuable comments on a draft of this article.

REFERENCES

1. D. Marr, *Vision*, W. H. Freeman Co., San Francisco, 1982.
2. A. Rosenfeld, and M. Thurston, "Edge and Curve Detection for Visual Scene Analysis", *IEEE Transactions on Computers* C-20, 562–569, 1971.
3. L. Davis, "A Survey of Edge Detection Techniques", *Computer Graphics and Image Processing* 4, 248–270, 1975.
4. E. Persoon, "A New Edge Detection Algorithm and its Applications", *Computer Graphics and Image Processing* 5, 425–446, 1976.
5. A. Rosenfeld, and A. Kak, *Digital Picture Processing*, Academic Press, New York, 1976.
6. M. J. Brooks, "Rationalizing Edge Detectors", *Computer Graphics and Image Processing* 8, 277–285, 1978.
7. W. Pratt, *Digital Image Processing*, John Wiley and Sons, New York, 1978.
8. H. Weschler, and K. S. Fu, "Image Processing Algorithms Applied to Rib Boundary Detection in Chest Radiographs", *Computer Graphics and Image Processing* 7, 375–390, 1978.
9. I. D. G. Macleod, "Comments on Techniques for Edge Detection", *Proceedings of the IEEE* 60, 344, 1972.
10. D. Marr, and T. Poggio, "A Theory of Human Stereo Vision", *Proceedings of the Royal Society of London B.* 204, 301–328, 1979.
11. D. Marr, and E. C. Hildreth, "Theory of Edge Detection", *Proceedings of the Royal Society of London B.* 207, 187–217, 1980.
12. E. C. Hildreth, "The Detection of Intensity Changes by Computer and Biological Vision Systems", *Computer Vision, Graphics, and Image Processing* 22, 1–27, 1983.
13. B. K. P. Horn, *Robot Vision*, MIT Press, Cambridge, 1985.
14. J. F. Canny, "Finding Edges and Lines in Images", *MIT Artificial Intelligence Laboratory Technical Report* 720, 1983.
15. K. S. Shanmugam, F. M. Dickey, and J. A. Green, "An Optimal Frequency Domain Filter for Edge Detection in Digital Pictures", *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-1, 37–49, 1979.
16. T. Poggio, H. Voorhees, and A. L. Yuille, "A Regularized Solution to Edge Detection", *MIT Artificial Intelligence Laboratory Memo* 773, 1984.
17. V. Torre, and T. Poggio, "On Edge Detection", *MIT Artificial Intelligence Laboratory Memo* 768, 1984.
18. J. Hadamard, *Lectures on the Cauchy Problem in Linear Partial Differential Equations*, Yale University Press, New Havens, 1923.
19. R. M. Haralick, "Edge and Region Analysis for Digital Image Data", *Computer Graphics and Image Processing* 12, 60–73, 1980.

20. J. M. S. Prewitt, "Object Enhancement and Extraction", in: *Picture Processing and Psychopictorics*, B. Lipkin, and A. Rosenfeld, eds., Academic Press, New York, 75–149, 1970.

21. R. M. Haralick, L. T. Watson, and T. J. Laffey, "The Topographical Primal Sketch", *International Journal of Robotics Research* 2, 50–72, 1983.

22. R. M. Haralick, "Digital Step Edges from Zero Crossing of Second Directional Derivatives", *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI–6, 58–68, 1984.

23. B. F. Logan, "Information in the Zero–crossings of Bandpass Signals", *Bell Systems Technical Journal* 56, 487–510, 1977.

24. A. L. Yuille, and T. Poggio, "Fingerprints Theorems for Zero–Crossings", *Proceedings of the Americal Association for Artificial Intelligence Conference*, 1984. Also appears as *MIT Artificial Intelligence Laboratory Memo* 730, 1984.

25. A. P. Witkin, "Scale–Space Filtering", *Proceedings of the International Joint Conference on Artificial Intelligence*, Karlsruhe, 1019–1022, 1983.

26. J. Babaud, A. P. Witkin, M. Baudin, and R. O. Duda, "Uniqueness of the Gaussian Kernel for Scale–Space Filtering", *Fairchild Technical Report* TR–645, 1984.

27. J. J. Koenderink, "The Structure of Images", *Biological Cybernetics* 50, 363–370, 1984.

28. A. L. Yuille, and T. Poggio, "Scaling Theorems for Zero–Crossings", *Proceedings of the IEEE Workshop on Computer Vision and Control*, Computer Society Press, 3–7, 1984. Also appears as *MIT Artificial Intelligence Laboratory Memo* 772, 1984.

29. H. R. Wilson, and J. R. Bergen, "A Four Mechanism Model for Threshold Spatial Vision", *Vision Research* 19, 19–32, 1979.

30. T. O. Binford, "Inferring Surfaces from Images", *Artificial Intelligence* 17, 205–244, 1981.

31. T. O. Binford, "Survey of Model–Based Image Analysis Systems", *International Journal of Robotics Research* 1, 18–64, 1982.

32. R. Nevatia, and R. Babu, "Linear Feature Extraction and Descripion", *Computer Graphics and Image Processing* 13, 257–269, 1980.

33. J. W. Modestino, and R. W. Fries, "Edge Detection in Noisy Images Using Recursive Digital Filtering", *Computer Graphics and Image Processing* 6, 409–433, 1977.

34. V. Berzins, "Accuracy of Laplacian Edge Detectors", *Computer Graphics and Image Processing* 27, 195–210, 1984.

35. A. Huertas, and G. Medioni, "Edge Detection with Subpixel Precision", *Proceedings of the IEEE Workshop on Computer Vision: Representation and Control*, October 1985, to appear.

36. R. W. Rodieck, and J. Stone, "Analysis of Receptive Fields of Cat Retinal Ganglion Cells", *Journal of Neurophysiology 28*, 833–849, 1965.

37. F. M. deMonasterio, "Properties of Concentrically Organized X and Y Ganglion Cells of Macaque Retina", *Journal of Neurophysiology* 41, 1394–1417, 1978.

38. M. Brady, editor, *Computer Vision*, North–Holland Co., Amsterdam, 1981.

39. D. H. Ballard, and C. M. Brown, *Computer Vision*, Prentice–Hall Inc., Englewood Cliffs, New Jersey, 1982.

40. K. M. Mutch, and W. B. Thompson, "Analysis of Accretion and Deletion at Boundaries in Dynamic Scenes", *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI–7, 133–138, 1985.

41. S. A. Shafer, *Shadows and Silhouettes in Computer Vision*, Kluwer Academic Publishers, Boston–Dordrecht–Lancaster, 1985.

CS-TR Scanning Project
Document Control Form

Date : 11/9/95

Report # AIM-858

Each of the following should be identified by a checkmark:
Originating Department:

☒ Artificial Intellegence Laboratory (AI)
☐ Laboratory for Computer Science (LCS)

Document Type:

☐ Technical Report (TR) ☒ Technical Memo (TM)
☐ Other:_____

Document Information

Number of pages: 22(28-IMAGES)
Not to include DOD forms, printer intstructions, etc... original pages only.

Originals are:
☒ Single-sided or
☐ Double-sided

Intended to be printed as :
☐ Single-sided or
☒ Double-sided

Print type:
☐ Typewriter ☐ Offset Press ☐ Laser Print
☐ InkJet Printer ☐ Unknown ☐ Other: ?

Check each if included with document:
☒ DOD Form (2) ☐ Funding Agent Form ☐ Cover Page
☐ Spine ☐ Printers Notes ☐ Photo negatives
☐ Other: _____

Page Data:

Blank Pages (by page number): _____

Photographs/Tonal Material (by page number): 3, 13, 15, 17

Other (note description/page number):

Description :	Page Number:
(A) IMAGE MAP: (1-22) PAGES #'ED 1-22	
(23-28) SCANCONTROL, DOD(2), TRGT'S(3)	
(B) ORS. OF PAGES 3-4, 15-16, 17-18, ARE TWO-SIDED.	
(C) CUT & PASTE FIGS ON PAGES 5, 7, 10, 13	

Scanning Agent Signoff:
Date Received: 11/9/95 Date Scanned: 11/29/95 Date Returned: 11/30/95

Scanning Agent Signature: Michael W. Cook

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 858	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER A162662
4. TITLE (and Subtitle) Edge Detection		5. TYPE OF REPORT & PERIOD COVERED A.I. Memo 858
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Ellen C. Hildreth		8. CONTRACT OR GRANT NUMBER(s) N0014-80-C-0505
9. PERFORMING ORGANIZATION NAME AND ADDRESS Artificial Inteligence Laboratory 545 Technology Square Cambridge, MA 02139		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency 1400 Wilson Blvd. Arlington, VA 22209		12. REPORT DATE September, 1985 13. NUMBER OF PAGES 22
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Office of Naval Research Information Systems Arlington, VA 22217		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Edge Detection
Computer Vison
Image filtering
Image Filtering
Intensity Changes
Gaussian Filtering
Multi-Resolution Image Analysis
Zero-Crossings

MAY 19 1986

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

For both biolgical systems and machines, visoin begins with a large and unwieldly array of measurements of the amount of light reflected from surfaces in the environment. The goal of vision is to recover physical properties of objects in the scene such as the location of object boundaries and the structure, color and texture of object surfaces, from the two-dimensional image that is projected ontohe eye or camera. This goal is not achieved in a single step: vision proceeds in stages, with each stage producing increasingly

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

more useful descriptions of the image and then the scene. The first clues about the physical properties of the scene are provided by the changes of intensity in the image. The importance of intensity changes and edges in early visual processing has led to extensive research on their detection, description and use, both in computer and biological vision systems. This article reviews some of the theory that underlies the detection of edges, and the methods used to carry out this analysis.

Scanning Agent Identification Target

Scanning of this document was supported in part by the **Corporation for National Research Initiatives,** using funds from the **Advanced Research Projects Agency** of the **United states Government** under Grant: **MDA972-92-J1029.**

The scanning agent for this project was the **Document Services** department of the **M.I.T Libraries.** Technical support for this project was also provided by the **M.I.T. Laboratory for Computer Sciences.**

Scanned

Date: 11/29/1995

M.I.T. Libraries
Document Services