

Using Canny's Criteria to Derive a Recursively Implemented Optimal Edge Detector

RACHID DERICHE

INRIA, Domaine de Voluceau, Rocquencourt BP 105, F-78153 Le Chesnay Cedex, France

Abstract

A highly efficient recursive algorithm for edge detection is presented. Using Canny's design [1], we show that a solution to his precise formulation of detection and localization for an infinite extent filter leads to an optimal operator in one dimension, which can be efficiently implemented by two recursive filters moving in opposite directions. In addition to the noise truncature immunity which results, the recursive nature of the filtering operations leads, with sequential machines, to a substantial saving in computational effort (five multiplications and five additions for one pixel, independent of the size of the neighborhood). The extension to the two-dimensional case is considered and the resulting filtering structures are implemented as two-dimensional recursive filters. Hence, the filter size can be varied by simply changing the value of one parameter without affecting the time execution of the algorithm. Performance measures of this new edge detector are given and compared to Canny's filters. Various experimental results are shown.

1 Introduction

One of the most important tasks in any vision system is the detection of edges in digitized images. The requirements of a good edge detector are that it responds only to true edge structure and is relatively insensitive to noise. Computationally efficient realization is also required. Several methods have been proposed and most of them suggest considering local operations on the elements of the input image to extract edges. Examples are the gradient operators and second-derivative operators generally followed respectively by peak and zero-crossing detections. Thresholding and thinning operations are applied to localize boundaries and remove edge segments that arise from noise in the imaging process. For a large survey of these techniques, one can refer to Davis [2], Grimson and Hildreth [3], Haralick [4], Hildreth [5], and Rosenfeld and Kak [6].

With blurred and noisy edges, the performance of local gradient operators deteriorates rapidly. This has led to the development of some particular methods more specialized for detecting edges in noisy pictures. Modestino and Fries [7] intro-

duced first an approach to edge detection in noisy images using two-dimensional recursive digital filtering. A stochastic model of edge structure was proposed and the edge detection problem formulated was one of least mean-square spatial filtering with an a priori assumption that a spatial frequency-weighted version of the Laplacian operator is the optimum filter. This led to a transfer function having the form:

$$H(\omega) = k \cdot \omega^2 \cdot e^{-\alpha^2 \omega^2} \cdot H_1(\omega) \quad (1)$$

where $H_1(\omega)$ is the Wiener filter which yields a minimum mean-squared error estimate of the edges in the input image.

The results indicated substantial advantages over conventional edge detectors in the presence of noise. In a later article, Shanmugan et al. [8] derived an optimal edge detection filter operating on a global basis and optimizing a given performance measure. The quantity whose maximum value was sought represents the portion of the output signal energy contained in an interval I (descriptive of the resolution of the system) centered in the vicinity of the edge. Their approach

led to an optimum edge detection filter specified in terms of the prolate spheroidal wave functions and the edge function. For the special case of a step edge input, the optimal frequency domain filter corresponds to the filter having the transfer function:

$$H(\omega) = k \cdot \omega^2 \cdot e^{-k_1 \cdot \omega^2} \quad (2)$$

Shanmugan et al. [8] follow their filtering stage with magnitude and thresholding operations.

More recently, in his MS thesis, Canny [1] made noise a central concern in his investigation of edge finding. Using the calculus of variations, he derived an operator from a precise formulation of detection and localization. The contents of our article are organized to first give a description of Canny's design and to describe and evaluate with Canny's performance measure a new edge detector which presents optimal performances and additional advantages in implementation. The large 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. A limitation of many of these schemes has been the use of extremely small operators very sensitive to noise. With our new edge detector recursively implemented, we can easily sacrifice localization for improving the output signal-to-noise ratio and vice versa. This trade-off is accomplished by changing the value of a single parameter, playing the same role as the inverse of the σ parameter of a Gaussian, filter without affecting time execution of the algorithm. This property is very useful in a multi-scaled description of a shape as described by Witkin [9] since it allows multi-scale edge detection to be performed with the same computational complexity for all the scales.

Presented in this article is a comparative study of this edge detector with Canny's optimal edge detector and its approximate filter, the first derivative of a Gaussian; this new edge detector is shown to perform better. An extension to the two-dimensional (2D) case, with the resulting filtering structures implemented as 2D infinite impulse response (IIR) digital filters, is presented. Experimental results with excellent sensitivity to edge detail and remarkable noise immunity are shown.

2 Canny's Design

In his MS thesis, Canny [1] considered first the one-dimensional case edge detection problem with the traditional model of a step in white Gaussian noise.

Let the amplitude of the step be A , and let the variance of the input white noise be n_0^2 . The input signal $I(x)$ can be represented by the step

$$I(x) = A \cdot u_{-1}(x) + n(x) \quad (3)$$

with $n_0^2 = \langle n^2(x) \rangle$ for all x .

He assumed that detection was performed by convolving the noisy edge with a spatial antisymmetric function $f(x)$ and making edges at the maxima in the output $\theta(x_0)$ of this convolution:

$$\theta(x_0) = \int_{-\infty}^{+\infty} I(x) \cdot f(x_0 - x) dx \quad (4)$$

Trying to formulate precisely the criteria for effective edge detection, he set the following goals:

Good detection. There should be a low probability of failing to detect a real edge point and low probability of falsely marking non-edge points. This criterion corresponds to maximizing signal-to-noise ratio (SNR), which is defined as the quotient of the response to the step only and the square root of the mean-squared noise response:

$$SNR = \frac{A}{n_0} \cdot \frac{\int_{-\infty}^0 f(x) dx}{\left[\int_{-\infty}^{+\infty} f^2(x) dx \right]^{\frac{1}{2}}} = \frac{A}{n_0} \cdot \sum \quad (5)$$

Finding the impulse response $f(x)$ which maximizes \sum corresponds to finding the best operator for detection only.

Good localization. The points marked as edges by the operator should be as close as possible to the center of the true edge. This criterion corresponds to minimizing the variance σ^2 of the zero-crossing position or maximizing the localization criterion L defined as the reciprocal of σ :

$$L = \frac{A}{n_0} \cdot \frac{|f'(0)|}{\left[\int_{-\infty}^{+\infty} f^2(x) dx \right]^{\frac{1}{2}}} = \frac{A}{n_0} \cdot \lambda \quad (6)$$

Finding the impulse response $f(x)$ which maximizes λ corresponds to finding the best operator for localization only.

One response to one edge. The detector should not produce multiple outputs in response to a single edge. There is a need to limit the number of peaks in the response so there will be a low probability of declaring more than one edge. The distance between peaks in the noise response of f , denoted x_{\max} , is set to be some fraction k of the operator width W :

$$x_{\max} = k \cdot W = 2 \cdot \pi \cdot \left[\frac{\int_{-\infty}^{+\infty} f'^2(x) dx}{\int_{-\infty}^{+\infty} f''^2(x) dx} \right]^{\frac{1}{2}} \quad (7)$$

Having developed criteria for detection, localization, and limitation of the number of peaks, Canny combined them in a meaningful way: Maximize the product $\sum \lambda$ (invariant under changes of scale or amplitude) under the constraint of the third criterion.

By expressing the criterion as a composite functional he found that this leads to the solution $f(x)$ such that

$$2 \cdot f(x) - 2 \cdot \lambda_1 \cdot f''(x) + 2 \cdot \lambda_2 \cdot f'''(x) + \lambda_3 = 0 \quad (8)$$

with

$$\Delta = \lambda_2 - \lambda_1^2/4 > 0 \quad (9)$$

and where α and ω are real, such that:

$$\alpha^2 - \omega^2 = \frac{\lambda_1}{2 \cdot \lambda_2} \quad 4 \cdot \alpha^2 \cdot \omega^2 = \frac{-\lambda_1^2 + 4 \cdot \lambda_2}{4 \cdot \lambda_2^2} \quad (10)$$

The general solution in the range $[0, W]$ may be written:

$$f(x) = a_1 \cdot e^{\alpha \cdot x} \cdot \sin \omega \cdot x + a_2 \cdot e^{\alpha \cdot x} \cdot \cos \omega \cdot x + a_3 \cdot e^{-\alpha \cdot x} \cdot \sin \omega \cdot x + a_4 \cdot e^{-\alpha \cdot x} \cdot \cos \omega \cdot x + C \quad (11)$$

subject to the boundary conditions:

$$f(0) = 0 \quad f(W) = 0 \quad f'(0) = S \quad f'(W) = 0 \quad (12)$$

where S is an unknown constant equal to the slope of the function $f(x)$ at the origin. Since $f(x)$ is antisymmetric, the above solution is extended to the range $[-W, +W]$ using $f(x) = -f(-x)$. The four boundary conditions enable the quantities a_1 through a_4 to be determined.

Using constrained numerical optimization, Canny found that the largest value of k that could

be obtained was about 0.58 and its performance was given by $\sum \lambda = 1.12$.

On inspection of the shape of this optimal operator, he observed that it is approximately the first derivative of a Gaussian:

$$f(x) = -\frac{x}{\sigma^2} \cdot e^{-x^2/2 \cdot \sigma^2} \quad (13)$$

The reason for doing this was that there are efficient ways to compute the 2D extension.

Because the overall performance index for the first derivative of a Gaussian is

$$\sum \lambda = 0.92 \text{ and } k = 0.51 \quad (14)$$

it is worse than the optimal operator by about 20% and was used exclusively in experiments (however, it should be pointed out that this approximation step has not been made clear by Canny; no analysis has been done concerning this approximation which has been proposed only on the observation of the shape). In two dimensions, the simplest form of the detector used by Canny is created by convolving the image with a symmetric two-dimensional Gaussian and then differentiating in two directions. After the image has been convolved with a symmetric Gaussian, the edge direction is estimated from the gradient of the smoothed image intensity surface. The gradient magnitude is then nonmaximum suppressed in that direction.

3 An Optimal Edge Detector

It should be pointed out that Canny's design was developed for a finite extent antisymmetric filter. Dealing with an infinite extent antisymmetric filter leads to the same differential equation and therefore the same general solution given by equation (11). However, the boundary conditions, given by equation (12), for finite extent filter defined in the range $[0, W]$ change as follows:

$$f(0) = 0 \quad f(+\infty) = 0 \quad f'(0) = S \quad f'(+\infty) = 0 \quad (15)$$

Applying these conditions to equation (11) leads to the following solution:

$$f(x) = -c \cdot e^{-\alpha \cdot x} \cdot \sin \omega \cdot x \quad (16)$$

with α , ω , and c positive reals.

Since $f(x)$ is antisymmetric and $\sin[\omega \cdot (-x)] = -\sin[\omega \cdot x]$, the above definition will be given for any x by:

$$f(x) = -c \cdot e^{-\alpha \cdot |x|} \cdot \sin \omega \cdot x \quad (17)$$

The Fourier transform of this filter is:

$$F(u) = \frac{i \cdot 4 \cdot c \cdot \alpha \cdot \omega \cdot u}{(\alpha^2 + \omega^2)^2 + u^4 + 2 \cdot u^2 \cdot (\alpha^2 - \omega^2)} \quad (18)$$

with $i^2 = -1$, and the variable u is spatial frequency. Its transfer function is therefore a band-pass.

Evaluating the integral in the performance criteria leads to the following results:

$$|f'(0)| = \omega \cdot c \quad (19)$$

$$\int_{-\infty}^0 f(x) dx = \frac{\omega \cdot c}{\alpha^2 + \omega^2} \quad (20)$$

$$\int_{-\infty}^{+\infty} f^2(x) dx = \frac{\omega^2 \cdot c^2}{2 \cdot \alpha \cdot (\omega^2 + \alpha^2)} \quad (21)$$

$$\int_{-\infty}^{+\infty} f'^2(x) dx = \frac{\omega^2 \cdot c^2}{2 \cdot \alpha} \quad (22)$$

$$\int_{-\infty}^{+\infty} f''^2(x) dx = \frac{\omega^2 \cdot c^2 \cdot (5 \cdot \alpha^4 + \omega^4 + 6 \cdot \alpha^2 \cdot \omega^2)}{2 \cdot \alpha \cdot (\omega^2 + \alpha^2)} \quad (23)$$

It follows that the overall performance index for our operator is:

$$\lambda = (2 \cdot \alpha)^{\frac{1}{2}} \quad \Sigma = \left[\frac{2 \cdot \alpha}{\alpha^2 + \omega^2} \right]^{\frac{1}{2}} \quad (24)$$

$$\Sigma \cdot \lambda = \frac{2 \cdot \alpha}{(\omega^2 + \alpha^2)^{\frac{1}{2}}} \quad k = \left[\frac{\alpha^2 + \omega^2}{5 \cdot \alpha^2 + \omega^2} \right]^{\frac{1}{2}} \quad (25)$$

If we set $\alpha = m \cdot \omega$, then we deduce four interesting cases:

$$\text{a) } m \gg 1 \quad \lambda = (2 \cdot \alpha)^{\frac{1}{2}} \quad \Sigma \approx \left(\frac{2}{\alpha} \right)^{\frac{1}{2}}$$

$$\Sigma \cdot \lambda \approx 2 \quad k \approx .44$$

$$\text{b) } m \ll 1 \quad \lambda = (2 \cdot \alpha)^{\frac{1}{2}} \quad \Sigma \approx m \cdot \left(\frac{2}{\alpha} \right)^{\frac{1}{2}}$$

$$\Sigma \cdot \lambda \approx 2 \cdot m \quad k \approx 1$$

$$\text{c) } m = 1 \quad \lambda = (2 \cdot \alpha)^{\frac{1}{2}} \quad \Sigma \approx \left(\frac{1}{\alpha} \right)^{\frac{1}{2}}$$

$$\Sigma \cdot \lambda \approx 1.414 \quad k \approx .58$$

$$\text{d) } m = (3)^{\frac{1}{2}} \quad \lambda = (2 \cdot \alpha)^{\frac{1}{2}} \quad \Sigma \approx \left(\frac{3}{2 \cdot \alpha} \right)^{\frac{1}{2}}$$

$$\Sigma \cdot \lambda \approx 1.732 \quad k \approx .5$$

Case d) shows that for the same value of k , the performance of the first derivative of a Gaussian is worse than the operator presented here, by more than 90%.

Case c) shows that the final form used by Canny for his optimal operator ($\Sigma \cdot \lambda \approx 1.12$ $k = .58$) is worse than the operator we presented by more than 25%.

Case b) presents an ideal response measure for k , but the problem is that the product $\Sigma \cdot \lambda$ becomes much less than unity.

The first case presents the best trade-off. Since the product $\Sigma \cdot \lambda$ is maximum for $\omega = 0$, and remembering that for $\omega \cdot x$ is very small we have $\sin(\omega \cdot x) \approx \omega \cdot x$. This leads us to rewrite our optimal filter as:

$$g(x) = -c \cdot x \cdot e^{-\alpha \cdot |x|} \quad (26)$$

The Fourier transform of this filter is:

$$F(u) = \frac{i \cdot 4 \cdot \alpha \cdot u \cdot c}{(u^2 + \alpha^2)^2} \quad (27)$$

This is the transfer function of a bandpass filter.

In fact, this function is the solution of equation (8) when Δ , given by equation (9), is null. Using the integral in the performance criteria, it can be easily verified that the overall performance index for this operator is given by:

$$\lambda = (2 \cdot \lambda)^{\frac{1}{2}} \quad \Sigma = \left(\frac{2}{\alpha} \right)^{\frac{1}{2}} \quad \Sigma \cdot \lambda = 2$$

$$k = \frac{1}{(5)^{\frac{1}{2}}} \approx .44 \quad (28)$$

It follows that this operator performs better than Canny's operator. It is very simple and presents only one parameter α . This parameter is adjusted to yield the desired localization or signal-to-noise ratio. It should be pointed out that the maximum of $g(x)$ is situated at $x = 1/\alpha$, thus α plays exactly as the inverse of the parameter σ in the first derivative of Gaussian. Decreasing α will lower the edge localization, but yield better

signal-to-noise ratio and vice versa.

For comparison purposes, figure 1a gives the shape of the operator $g(x)$ (extern curve) and compares it with the first derivative of a Gaussian (intern curve); figure 1b shows the same operator compared with Spacek's operator [10] (intern curve), and figure 1c allows one to compare the same operator $g(x)$ with $f(x)$ (intern curve) when the ratio α/ω is much more than the unity and one can see that no difference can be detected in this case. For the sake of this comparison, all the operators have been scaled to have the same maximum amplitude.

From case a) we see that $f(x)$ presents the same overall performance index as $g(x)$, and from figure 1c that it has the same shape than $g(x)$ when the ratio α/ω is much more than the unity. Because $f(x)$ can be efficiently implemented in a recursive way (see the next section), it is the final form of optimal operator that we will use here. In a later article we will evaluate performances of the operator $g(x)$ and give an efficient implementation in one dimension and two dimensions.

4 One-dimensional Digital Recursive Implementation

In this section, we introduce first the problem of recursive filtering; develop, in a second part, one procedure to implement our optimal operator recursively; and analyze, in the last part, the problem of errors in recursive filtering.

4.1 Introduction

Given an impulse response sequence $\{h(n)\}$ of length N , whose Z -transform is:

$$H_d(Z) = \sum h(n) \cdot Z^{-n} \text{ for } n = 0, \dots, N-1 \quad (29)$$

A causal nonrecursive system is characterized by the following equation:

$$y(i) = \sum h(k) \cdot x(i-k) \text{ for } k = 0, \dots, N-1 \text{ for } i = 1, \dots, M-1 \quad (30)$$

where $\{x(n)\}$ and $\{y(n)\}$ denote the nonrecursive system's input and output of length M points, respectively.

The number of operations required to calculate each output element $y(i)$ can be excessive if we deal with filters having large length N . Thus for large values of N , a direct implementation of the nonrecursive system described by equation (30) can be inefficient since it is very time-consuming. For example, a smoothing operation using a half-Gaussian filter will need about $4 \cdot \sigma$ arithmetic operations per input element, if we decide to truncate the Gaussian filter at about 0.1% of its peak value. Thus for large values of σ , a direct implementation will be very time-consuming.

The problem of recursive filtering design deals with the determination of the coefficients a_k 's and b_k 's of a rational transfer function of the form:

$$H_a(Z) = \frac{a_0 + a_1 \cdot Z^{-1} + \dots + a_{n-1} \cdot Z^{-(n-1)}}{1 + b_1 \cdot Z^{-1} + \dots + b_n \cdot Z^{-n}} \quad (31)$$

which characterizes the causal recursive system of order n :

$$y(i) = \sum a_l \cdot x(i-l) - \sum b_k \cdot y(i-k) \text{ for } l = 0, \dots, n-1 \text{ and } k = 0, \dots, n \quad (32)$$

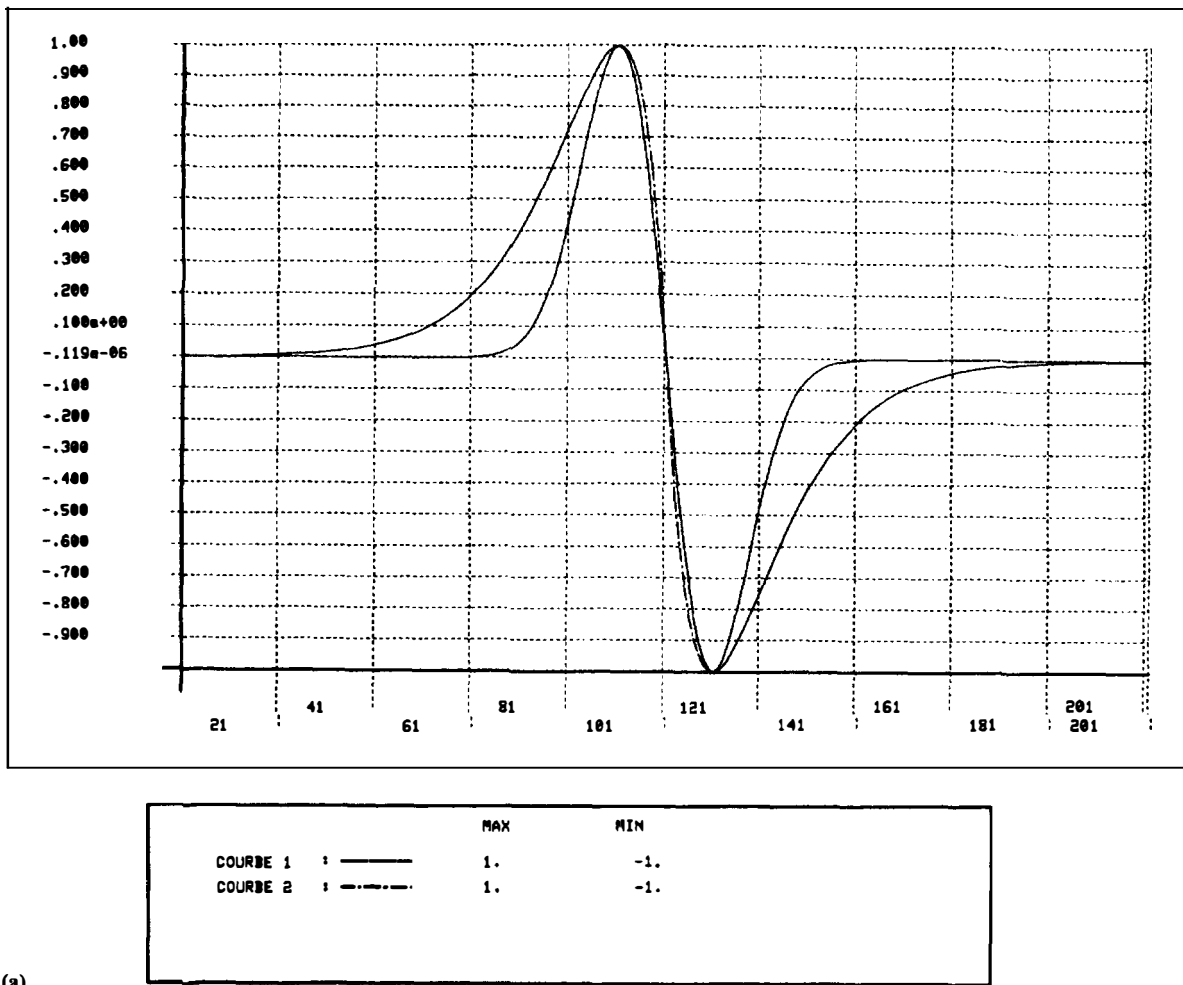
so that the rational transfer function $H_a(Z)$ is exactly, or best approximates in accordance with certain error criterion, $H_d(Z)$ the transfer function of the nonrecursive system described by equation (29).

The most used criterion is that in which the corresponding impulse responses are compared by a least-squares criterion, that is, minimizing E such that:

$$E = \sum (h(k) - h_a(k))^2 \text{ for } k = 0, \dots, +\infty \quad (33)$$

with $\{h_a(k)\}$ denoting the impulse response of recursive system given by equation (32).

We present in reference [11] a procedure to determine the b_k 's and the a_k 's coefficients using a least-squares criterion and an efficient separable and recursive implementation for the most commonly and successfully used filters in edge detection: the 2D Gaussian filter, its first directional derivative, and its 2D Laplacian. We show in particular that these filters can be efficiently implemented in a recursive way with only third-order recursive filters. However, for our optimal operator $f(x)$, given by equation (16), the approx-



(a)

Fig. 1. Comparison between operators. (a) Comparison between the first derivative of a Gaussian (intern curve) and $g(x)$ (extern curve). (b) Comparison between Spacek's operator (intern curve) and $g(x)$ (extern curve). (c) Comparison between the operator $f(x)$ (intern curve) and $g(x)$ (extern curve).

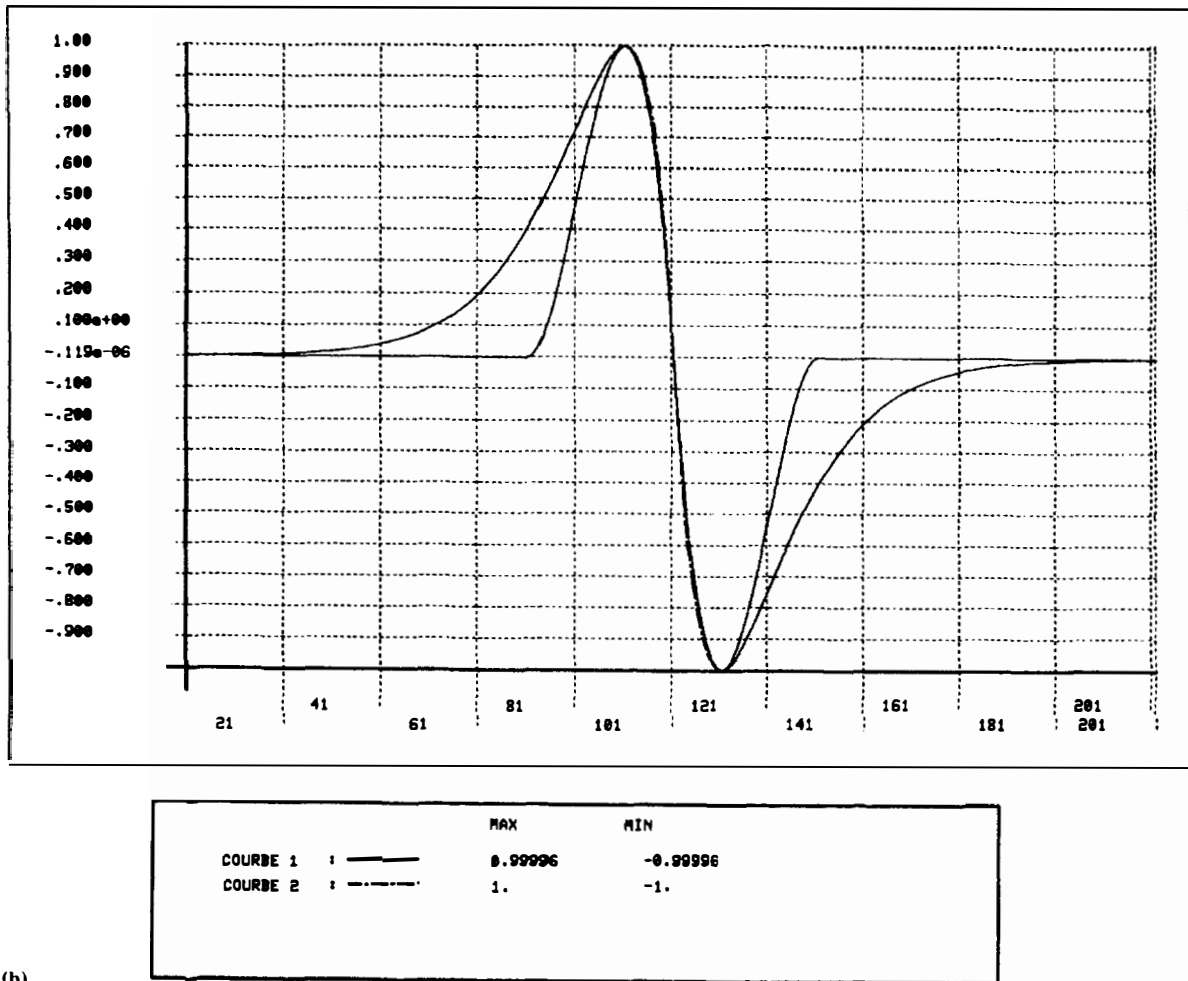
imation step is not necessary. We show in the following subsection that the coefficients a_k 's and b_k 's of equation (31) for $f(x)$ can be exactly calculated from its transfer function and that the value obtained for n is 2.

When we deal with the causal recursive system of order n , given by equation (32), the number of operations per input element is reduced to $2 \cdot n$. Therefore, our implementation reduces the amount of computation from N operations per input element for a direct nonrecursive implementation to only $2 \cdot n$. This renders the use of a

large filter to be very practical. However, it should be pointed out first that the structure of this filter is highly sequential and so its implementation assumes only sequential machines and, second, a recursive filter is useless unless it is stable; therefore stability checks must accompany all design and, to preserve properties of interest like stability and filter frequency response, it is necessary to have sufficient accuracy in the representation of the filter coefficients. This will be discussed in section 4.3.

4.2 Design Algorithm

A recursive realization of the optimal filter $f(x)$, given by equation (17), may be derived by the fol-



(b)

lowing way. Let $f(n)$ be the samples of $f(x)$ and $F(Z)$ its Z-transform:

$$F(Z) = \sum f(n) \cdot Z^{-n} \quad \text{for } n = -\infty, \dots, +\infty \quad (34)$$

To deal with causal sequence, we have to perform some operations to transform the impulse response $f(n)$ as a sum of causal sequences. To this end, we split the impulse response $f(n)$ into two halves $f_-(n)$ and $f_+(n)$ such that:

$$f(n) = f_-(n) + f_+(n) \quad \text{for } n = -\infty, \dots, +\infty \quad (35)$$

$$f_-(n) = \begin{cases} 0 & n \geq 0 \\ \alpha_1(p_2)^n + \alpha_1^*(p_2^*)^n & n < 0 \end{cases} \quad (36)$$

$$f_+(n) = \begin{cases} \alpha_1(p_1)^n + \alpha_1^*(p_1^*)^n & n \geq 0 \\ 0 & n < 0 \end{cases} \quad (37)$$

and

$$\alpha_1 = \frac{-c}{2 \cdot i} \quad p_1 = e^{-\alpha + i \cdot \omega} \quad p_2 = e^{\alpha + i \cdot \omega} \quad (38)$$

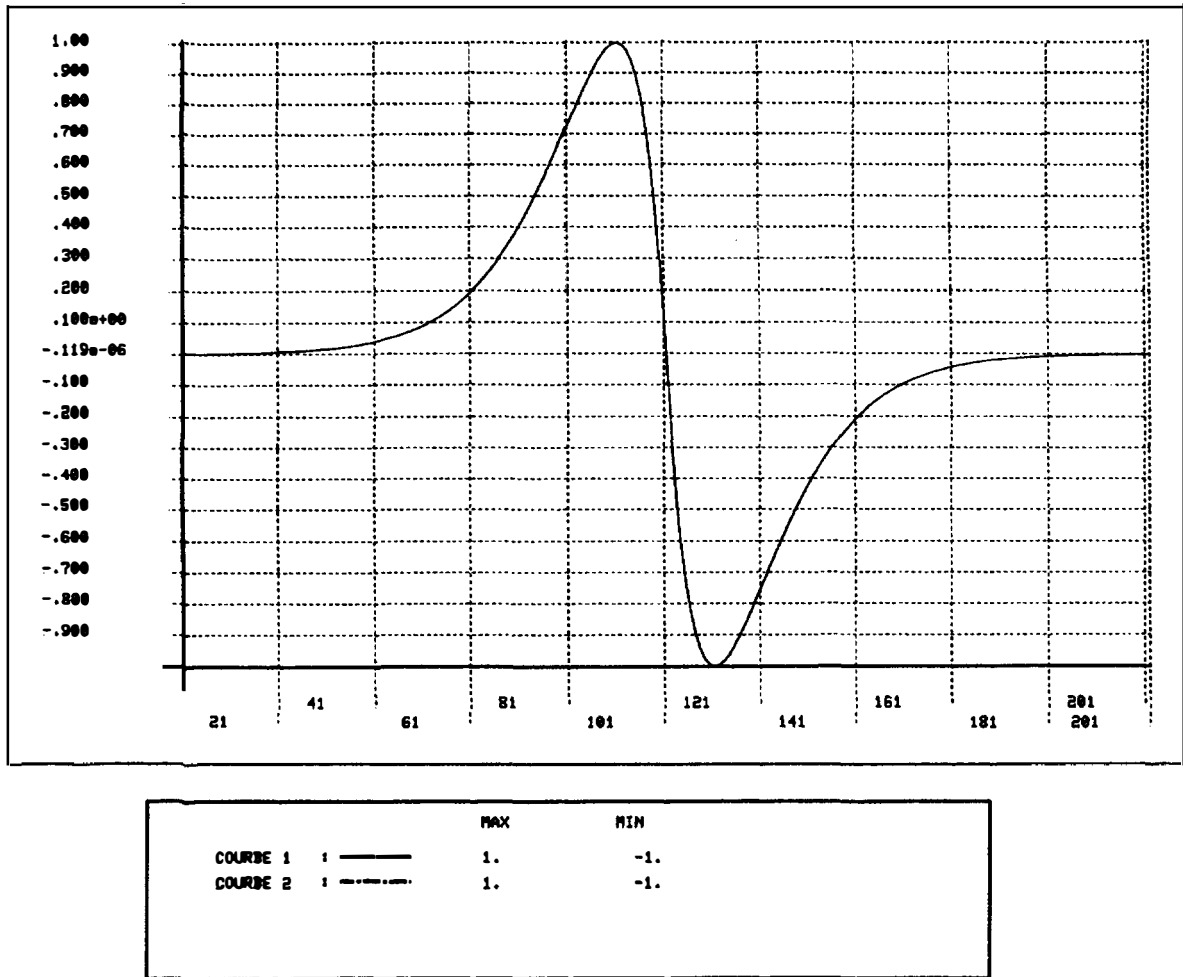
* denotes the complex conjugate.

Using the Z-transform, recognizing that each factor $\alpha_i \cdot (p_i)^n$ has Z-transform $\alpha_i / (1 - p_i \cdot Z^{-1})$ and simplifying, we obtain:

$$F(Z) = F_-(Z) + F_+(Z^{-1})$$

where

$$F_+(Z^{-1}) = \frac{a \cdot Z^{-1}}{1 + b_1 \cdot Z^{-1} + b_2 \cdot Z^{-2}}$$



$$F_{-}(Z) = \frac{-a \cdot Z}{1 + b_1 \cdot Z + b_2 \cdot Z^2} \quad (39)$$

with

$$\begin{aligned} a &= -c \cdot e^{-\alpha} \cdot \sin \omega & b_1 &= -2 \cdot e^{-\alpha} \cdot \cos \omega \\ b_2 &= e^{-2 \cdot \alpha} \end{aligned} \quad (40)$$

$F_{+}(Z^{-1})$ (resp. $F_{-}(Z)$) converges for $|p_1 \cdot Z^{-1}| < 1$ and $|p_1^* \cdot Z^{-1}| < 1$ (resp. $|p_2^{-1} \cdot Z| > 1$ and $|p_2^{-1*} \cdot Z| > 1$). Since all singularities of $F_{+}(Z^{-1})$ (resp. $F_{-}(Z)$) are inside (resp. outside) the unit circle for α positive real, these two Z -transforms correspond to two rational system transfer functions of stable second-order filters recursing from the left to the right (F_{+}) and the right to the left (F_{-}).

In particular, the output sequence $\{y(m)\}$ in response to $\{x(m)\}$ as input to a system, with im-

pulse response $\{f(n)\}$, can be obtained recursively according to:

$$\begin{aligned} y^{+}(m) &= x(m-1) - b_1 \cdot y^{+}(m-1) - \\ &\quad b_2 \cdot y^{+}(m-2) \end{aligned} \quad (41)$$

for $m = 1, \dots, M$

$$\begin{aligned} y^{-}(m) &= x(m+1) - b_1 \cdot y^{-}(m+1) - \\ &\quad b_2 \cdot y^{-}(m+2) \end{aligned} \quad (42)$$

for $m = M, \dots, 1$

$$y(m) = a \cdot (y^{+}(m) - y^{-}(m)) \quad (43)$$

for $m = 1, \dots, M$

with a , b_1 , and b_2 given by equation (33) and M denoting the length of $\{(y(m))\}$. The constant c may be fixed by the following normalization requirement:

$$\begin{aligned} & \{\sum f(n) \text{ for } n = 0, \dots, +\infty\} \\ &= \{-\sum f(n) \text{ for } n = -\infty, \dots, 0\} \\ &= -1 \end{aligned} \quad (44)$$

Straightforward manipulations lead to a value for c such that:

$$c = \frac{1 - 2 \cdot e^{-\alpha} \cdot \cos \omega + e^{-2 \cdot \alpha}}{e^{-\alpha} \cdot \sin \omega} \quad (45)$$

The importance of the equations (41) to (43) as a practical tool for determining the output of the optimal edge detector without having to perform a convolution should not be neglected. The shape of the filter, determined by the parameters α and ω , can be varied by simply changing the values of α and ω , without affecting time execution. Small values for α give large filters. The algorithm requires only five multiplications and five additions per point independently of the size of the operator specified by the parameters α and ω . For comparison purposes, an FIR (finite impulse response) implementation of $\{f(n)\}$, truncated to have $2 \cdot N + 1$ coefficients, will require N operations (multiplication + addition) per point using the antisymmetry of $\{f(n)\}$ when the recursive algorithm requires only five operations without any truncature operation. Thus, we have a gain ratio of $N/5$, which can be important for large N . However, even if equations (41) and (42) can be performed in separate processors, the structure of this filter is highly sequential and assumes only sequential machines.

Applying the same technique, we can show that samples $\{h(n)\}$ of the integral $h(x)$ of $f(x)$ can also be implemented recursively. We will apply this to the extension of the two-dimensional case. We have:

$$h(n) = (c_1 \cdot \sin \omega \cdot |n| + c_2 \cdot \cos \omega \cdot |n|) \cdot e^{-\alpha \cdot |n|} \quad (46)$$

with

$$c_1 = \frac{k \cdot \alpha}{\alpha^2 + \omega^2} \quad c_2 = \frac{k \cdot \omega}{\alpha^2 + \omega^2} \quad (47)$$

A careful observation of this smoothing function reveals that for the case where α and ω are such that:

$$\alpha = \omega = \frac{1}{\sqrt{2 \cdot \lambda}} \quad (48)$$

$h(x)$ is equal, up to a multiplicative constant (k/α^2), to the optimal smoothing operator $R(x, \lambda)$ derived by Poggio et al. [12] using the tools of regularization theory to transform the ill-posed differentiation problem into a well-posed problem.

Let λ be the regularizing parameter determining the scale of the optimal filter; the one-dimensional regularizing filter in Poggio et al. [12] is then given by:

$$R(x, \lambda) = [\alpha/(\sqrt{2})] \cdot e^{-\alpha \cdot |x|} \cdot \cos(\alpha \cdot |x| - \pi/4) \quad (49)$$

with α given by equation (48).

This point is very interesting since the approach developed here is completely different from Poggio's work.

Let us see how $h(n)$ can be implemented recursively. Calculating the Z -transform $H(Z)$ of $h(n)$, we obtain:

$$\begin{aligned} H(Z) &= H_+(Z^{-1}) + H_-(Z) \\ H_+(Z^{-1}) &= \frac{a_0 + a_1 \cdot Z^{-1}}{1 + b_1 \cdot Z^{-1} + b_2 \cdot Z^{-2}} \quad \text{and} \\ H_-(Z) &= \frac{a_2 \cdot Z + a_3 \cdot Z^2}{1 + b_1 \cdot Z + b_2 \cdot Z^2} \end{aligned} \quad (50)$$

with

$$\begin{aligned} a_0 &= c_2 \quad a_1 = (-c_2 \cdot \cos \omega + c_1 \cdot \sin \omega) e^{-\alpha} \\ a_2 &= a_1 - c_2 \cdot b_1 \quad a_3 = -c_2 \cdot b_2 \\ b_1 &= -2e^{-\alpha} \cdot \cos \omega \quad b_2 = e^{-2 \cdot \alpha} \end{aligned} \quad (51)$$

$H_+(Z^{-1})$ (resp. $H_-(Z)$) converges for $|p_1 \cdot Z^{-1}| < 1$ and $|p_1^* \cdot Z^{-1}| < 1$ (resp. $|p_2^{-1} \cdot Z| > 1$ and $|p_2^{-1*} \cdot Z| > 1$). Since all singularities of $H_+(Z^{-1})$ (resp. $H_-(Z)$) are inside (resp. outside) the unit circle for α positive real, these two Z -transforms correspond to two rational system transfer functions of stable second-order filters recursing from left to right (H_+) and right to left (H_-).

The output $\{y(m)\}$ of the convolution of an input sequence $\{x(m)\}$ with the impulse response $\{h(n)\}$ is obtained as follows:

$$\begin{aligned} y^+(m) &= a_0 \cdot x(m) + a_1 \cdot x(m-1) - \\ &\quad b_1 \cdot y^+(m-1) - b_2 \cdot y^+(m-2) \end{aligned} \quad (52)$$

for $m = 1, \dots, M$

$$\begin{aligned} y^-(m) &= a_2 \cdot x(m+1) + a_3 \cdot x(m+2) - \\ &\quad b_1 \cdot y^-(m+1) - b_2 \cdot y^-(m+2) \end{aligned} \quad (53)$$

for $m = M, \dots, 1$

$$y(m) = y^+(m) + y^-(m) \quad \text{for } m = 1, \dots, M \quad (54)$$

with a_0, a_1, a_2, a_3, b_1 , and b_2 given by equation (51). The constant k may be fixed by the following normalization requirement:

$$\{\sum h(n) \text{ for } n = -\infty, \dots, +\infty\} = 1 \quad (55)$$

This leads to a value for k such that:

$$k = \frac{[1 - 2 \cdot e^{-\alpha} \cos \omega + e^{-2\alpha}] \cdot [\alpha^2 + \omega^2]}{[2 \cdot \alpha \cdot e^{-\alpha} \sin \omega + \omega - \omega \cdot e^{-2\alpha}]} \quad (56)$$

This smoothing step requires only eight multiplications and seven additions per point. This has to be compared to the $(N+1)$ operations per point required in a FIR implementation of $\{h(n)\}$, truncated to have $(2 \cdot N + 1)$ coefficients nonnull, using the symmetry of $\{h(n)\}$. In addition to the substantial saving in computational effort that results in sequential machines when N is large enough, it should be pointed out that this implementation leads to a noise truncature immunity.

4.3 Effect of Finite Word Length on the Accuracy of Filters

We have discussed the computations to be done by the filters $\{f(n)\}$ and $\{h(n)\}$ as if the arithmetic would be handled accurately. In practice, it is well known [13, 14] that because of finite word length used in their implementation, the performance of the actual filter will differ from the proposed filter. This question becomes increasingly important when the filter is to be built out of definite hardware (integrated circuit chips) or implemented in a computer with fixed-point arithmetic assuming few bits. For efficiency, the chip is apt to have a rather short word length. In addition to the word length, the accuracy of a digital filter depends on two factors: the form of realization and the type of arithmetic used. This section attempts to discuss the effects of a finite number of bits on the accuracy of the filters $f(n)$ and $h(n)$, when their implementation is realized in the direct form characterized by equations (41), (42), (52), (53), and (54) using a fixed-point implementation. The effects analyzed are the roundoff errors

and the filter coefficient quantization errors.

Roundoff errors occur every time a multiplication is performed. Multiplying an n bits (including sign bit) number by an m bits number will result in an $n+m-1$ bits numbers. Generally only p bits ($p < n+m-1$) will be available to store the result of multiplication. Rounding/truncation operation will be performed introducing the so-called roundoff error.

Quantization errors occur because the filter coefficients a_k 's and b_k 's represented by only a finite number of bits will not take the exact value they were assigned by the proposed filter.

4.3.1 Modeling of the Multiplication Roundoff Error. In a fixed-point representation, the error introduced in rounding a number off can be modeled as a random variable with uniform probability density function, having zero mean and a variance of $(2^{-2 \cdot b})/12$, when $(b+1)$ is the total number of bits (including the sign bit).

In order to model the effects of rounding due to multiplication in a digital filter, the following assumptions are generally made:

1. Any two different samples from the same noise source are uncorrelated.
2. Any two different noise sources are uncorrelated.
3. Each noise source is uncorrelated with the input sequence.

The above set of assumptions make it possible to calculate the total noise variance at the output of the filter as a superposition of the variances due to each noise source with the other noise sources and input deactivated.

Consider the direct form realization of the transfer function $G(Z)$ given as:

$$\begin{aligned} G(Z) &= \frac{N(Z)}{D(Z)} \\ &= \frac{a_0 + a_1 \cdot Z^{-1} + \dots + a_{n-1} \cdot Z^{-(n-1)}}{1 + b_1 \cdot Z^{-1} + \dots + b_n \cdot Z^{-n}} \end{aligned} \quad (57)$$

Since the total number of noise source is $2 \cdot n$, the total output noise variance can be readily obtained as:

$$\sigma^2 = 2 \cdot n((2^{-2b})/12) \cdot \sum h(n)^2 \quad (58)$$

for $n = 0, \dots, +\infty$

where $\{h(n)\}$ is the impulse response corresponding to the following transfer function:

$$H(Z) = \frac{1}{D(Z)} = \frac{1}{1 + b_1 \cdot Z^{-1} + \dots + b_n \cdot Z^{-n}} \quad (59)$$

It is worth noting that the infinite sum in equation (58) can be computed by making use of the Parseval's relation [13]:

$$\sum h(n)^2 = (1/(2 \cdot \pi \cdot j)) \cdot \int H(Z) \cdot H(Z^{-1}) \cdot Z^{-1} dZ \quad (60)$$

for $n = 0, \dots, +\infty$

Where the contour of integration C is the unit circle, traversed in the counterclockwise direction.

It is then easy to determine σ^2 for the specified transfer function given in equations (39) and (50). For both cases, $H(Z)$ is given by:

$$H(Z) = \frac{1}{D(Z)} = \frac{1}{1 - 2 \cdot \cos(\omega) \cdot e^{-\alpha} \cdot Z^{-1} + e^{-2 \cdot \alpha} \cdot Z^{-2}} \quad (61)$$

Therefore, applying the Parseval's relation and taking into account N noise sources, we obtain the following interesting results:

$$\sigma^2 = N \cdot \frac{2^{-2 \cdot b}}{12} \cdot \frac{(1 + e^{-2 \cdot \alpha})}{(1 - e^{-2 \cdot \alpha})} \cdot \frac{1}{(1 + e^{-4 \cdot \alpha} - 2 \cdot e^{-2 \cdot \alpha} \cdot \cos(2 \cdot \omega))} \quad (62)$$

For the optimal filter $\{f(n)\}$ implemented by equations (41) and (42), since there are two multiplications in each direction, two noise sources are introduced for each direction and therefore $N = 2$. For the smoothing filter $\{h(n)\}$, implemented by equations (52) and (53), N is equal to 4. An important aspect of the statistical model presented above is to show how σ^2 becomes important when α is clustered to 0. Further, it enables us to calculate for different values of α , ω , and b the number of bits, the value of σ^2 , and assess whether a practical realization of the filters $\{h(n)\}$ and $\{f(n)\}$ can be made on a given word-length

computer. As an example of the use of this analysis of the filter $\{h(n)\}$ with $\alpha = 1.0$ and $\omega = 0.0001$, the number of bits required to meet $\sigma \approx 10^{-5}$ can readily be found to be about 16. For $\alpha = 0.1$, $\omega = 0.0001$, to deal with the same error, the number of bits required has been found to be about 19. The same analysis can be made for the smoothing filter $\{h(n)\}$.

4.3.2 Coefficient Accuracy. The frequency response of the actual filter realized with a fixed number of bits deviates somewhat from that which would have been obtained with an infinite-word-length computer. By regarding the filter coefficients errors as statistically independent and uniformly distributed with zero mean, Knowles and Olcayto [15] have calculated the statistically expected mean-square difference between the real frequency responses of the actual and real filters given by:

$$\sigma^2 = E[(1/(2 \cdot \pi)) \cdot \int_{-\pi}^{+\pi} H(j \cdot \omega) - H_{\infty}(j \cdot \omega)|^2 \cdot d\omega] \quad (63)$$

Where $H(j \cdot \omega)$ and $H_{\infty}(j \cdot \omega)$ represent the frequency responses of the actual and ideal (infinite word length) filters and E denotes expectation.

Let the ideal digital filter being realized as a transfer function:

$$H(Z) = \frac{A_{\infty}(Z)}{B_{\infty}(Z)} \quad (64)$$

Then Knowles and Olcayto [15] showed that σ^2 , given by equation (63), can be readily evaluated by the computational technique showed in figure 2 where q is the size of the quantization step ($q = 2^{-b}$) and ν and ν are the number of nonzero, nonunity coefficients in the numerator and denominator, respectively. Experimental measurements verified the analysis in the practical case.

Applying the statistical model presented above to the specified transfer functions given in equations (39) and (50), the degeneration in filter performance can be calculated for different values of α , ω , and b , the number of bits used in a fixed point implementation. Following the results obtained for σ^2 , the number of bits can be found to be sufficient or not and changed consequently. For example, for $\alpha = 1$ and $\omega = 0.0001$, σ has

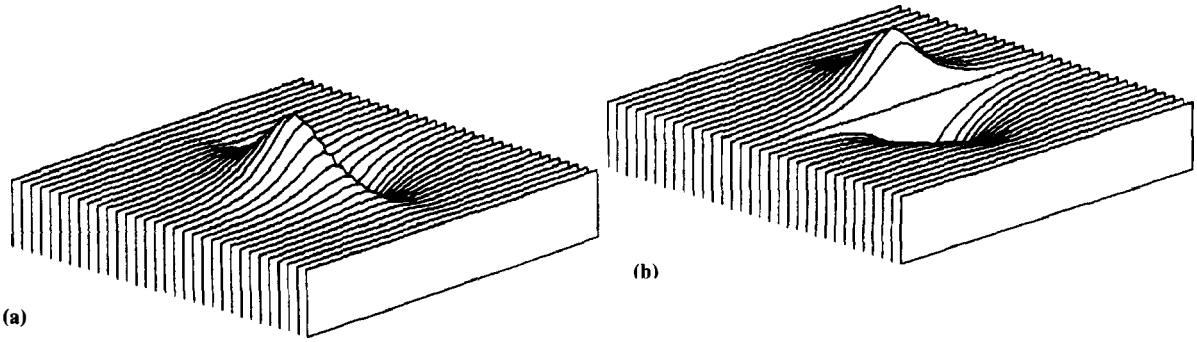


Fig. 2. (a and b) Plots of the masks $\{X(m, n)\}$ and $\{Y(m, n)\}$ with $\alpha = 0.5$ and $\omega = 0.0005$

been found to be about $1.7 \cdot 10^{-5}$ using 15-bit word length. For $\alpha = 0.1$ and $\omega = 0.0001$, to deal roughly with the same error the number of bits required is 19.

5 Two-dimensional Case

The operators in the preceding sections can be extended to the 2D case by making use of the fact that the slope of a surface in any direction can be

determined exactly from its slopes in two directions. We create for the x -direction (y -direction, respectively) a 2D separable mask, product of the edge detector aligned with the x -direction (y -direction, respectively) with a projection function parallel to the y -direction (x -direction, respectively). A highly efficient recursive implementation results if the projection function, which can be considered as the smoothing function, is chosen to be the integral of the edge detector.

Making use of the integral of the operator $f(n)$, given by equation (17), we obtain:

$$X(m, n) = \frac{[-c \cdot e^{-\alpha \cdot |m|} \cdot \sin \omega \cdot m] \cdot [k \cdot (\alpha \cdot \sin \omega \cdot |n| + \omega \cdot \cos \omega \cdot |n|) \cdot e^{-\alpha \cdot |n|}]}{\alpha^2 + \omega^2} \quad (65)$$

$$Y(m, n) = \frac{[k \cdot (\alpha \cdot \sin \omega \cdot |m| + \omega \cdot \cos \omega \cdot |m|) \cdot e^{-\alpha \cdot |m|}] \cdot [-c \cdot e^{-\alpha \cdot |n|} \sin \omega \cdot n]}{\alpha^2 + \omega^2} \quad (66)$$

Figures 2a and b show the plots of $\{X(m, n)\}$ and $\{Y(m, n)\}$ obtained for $\alpha = 0.5$ and $\omega = 0.0005$. For the operator $g(x)$, the two masks are:

$$X(m, n) = \frac{[-c \cdot m \cdot e^{-\alpha \cdot |m|}] \cdot [k \cdot (\alpha \cdot |n| + 1) \cdot e^{-\alpha \cdot |n|}]}{\alpha^2} \quad (67)$$

$$Y(m, n) = \frac{[k \cdot (\alpha \cdot |m| + 1) \cdot e^{-\alpha \cdot |m|}] \cdot [-c \cdot n \cdot e^{-\alpha \cdot |n|}]}{\alpha^2} \quad (68)$$

k and c may be fixed by the normalization requirement as described in equations (44) and (55). This leads to the following values:

$$c = \frac{[1 - e^{-\alpha}]^2}{e^{-\alpha}}$$

$$k = \frac{[1 - e^{-\alpha}]^2 \cdot \alpha^2}{1 + 2 \cdot \alpha \cdot e^{-\alpha} - e^{-2 \cdot \alpha}} \quad (69)$$

In an FIR realization, the size of the two masks $\{X(m, n)\}$ and $\{Y(m, n)\}$ depends on α and the approximation error caused by not extending them to infinity. By making use of a 2D IIR implementation for the two masks of equations (65)

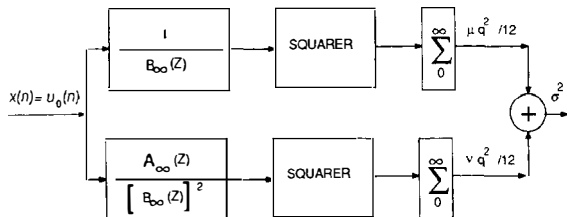


Fig. 3. Technique for measuring variance of error due to coefficient quantization after Knowles and Olcayto [15].

and (66) (see the next section), this error will be zero. After the image has been convolved with these two masks, we obtain two images $\{r(m, n)\}$ and $\{s(m, n)\}$. The edge direction $\{\alpha(m, n)\}$ and the amplitude $\{A(m, n)\}$ are then estimated from $\{r(m, n)\}$ and $\{s(m, n)\}$ by:

$$A(m, n) = (r(m, n)^2 + s(m, n)^2)^{\frac{1}{2}}$$

$$\alpha = \arctg \frac{s(m, n)}{r(m, n)} \quad (70)$$

$\{A(m, n)\}$ is then nonmaximum suppressed in the exact edge direction $\alpha(m, n)$ and thresholded with hysteresis, i.e., if any part of a contour is above a high threshold, that point is immediately output, as is the entire connected segment of the contour that contains the point and that lies above a low threshold (for more details, the reader is referred to Canny [1]).

6 Two-dimensional Recursive Implementation

By making use of their separability, the masks $\{X(m, n)\}$ and $\{Y(m, n)\}$ given by equations (65) and (66) can easily be implemented as a 2D separable recursive filter.

To convolve an $NLG \times NCL$ image $\{I(i, j)\}$ with the mask $\{X(m, n)\}$, we first do a filtering in the horizontal direction with the operator $\{f(n)\}$ and then apply the filter $\{h(n)\}$ in the vertical direction to the image obtained by the horizontal filtering. Using the results obtained in the preceding sections, we obtain:

$$Y^+(i, j) = I(i, j-1) - b_1 \cdot Y^+(i, j-1) - b_2 \cdot Y^+(i, j-2) \quad (71)$$

$$j = 1, \dots, NCL \quad i = 1, \dots, NLG$$

$$Y^-(i, j) = I(i, j+1) - b_1 \cdot Y^-(i, j+1) - b_2 \cdot Y^-(i, j+2) \quad (72)$$

$$j = NCL, \dots, 1 \quad i = 1, \dots, NLG$$

$$S(i, j) = a \cdot (Y^+(i, j) - Y^-(i, j)) \quad (73)$$

$$j = 1, \dots, NCL \quad i = 1, \dots, NLG$$

$$R^+(i, j) = a_0 \cdot S(i, j) + a_1 \cdot S(i-1, j) - b_1 \cdot R^+(i-1, j) - b_2 \cdot R^+(i-2, j) \quad (74)$$

$$i = 1, \dots, NLG \quad j = 1, \dots, NCL$$

$$R^-(i, j) = a_2 \cdot S(i+1, j) + a_3 \cdot S(i+2, j) - b_1 \cdot R^-(i+1, j) - b_2 \cdot R^-(i+2, j) \quad (75)$$

$$i = NLG, \dots, 1 \quad j = 1, \dots, NCL$$

$$R(i, j) = R^-(i, j) + R^+(i, j) \quad (76)$$

$$i = 1, \dots, NLG \quad j = 1, \dots, NCL$$

the values of $a, a_0, a_1, a_2, a_3, b_1$, and b_2 are given by equations (40) and (51).

This implementation allows us to convolve the input image with the mask $\{X(m, n)\}$ with a computational effort of 13 multiplications and 12 additions per point and without any truncation of the mask. The same technique can be applied to the mask $\{Y(m, n)\}$.

7 Experimental Results

This section is concerned with testing the algorithm for different types of images. Various real-world images and two noisy images were selected to show the capabilities of our optimal filter recursively implemented. All the images processed are of size 256×256 except the 512×512 stereo pair representing indoor scenes (figure 6).

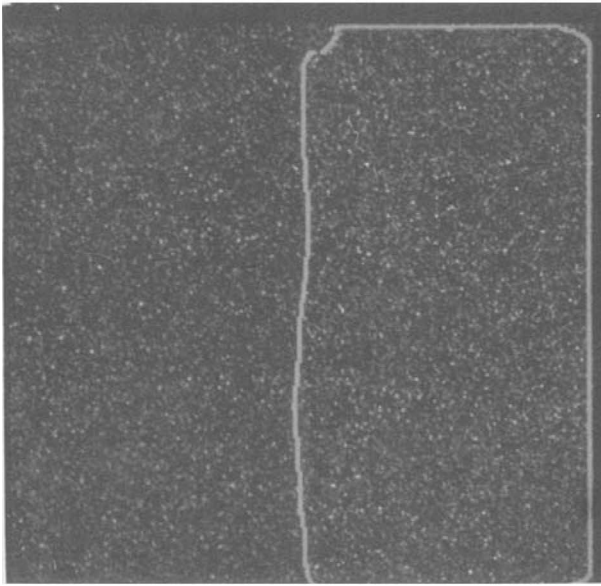
In figures 4a and b, the performance of our filter is illustrated when input images consist of a step edge and a centered circle with a radius of 64 pixels, added in a white Gaussian noise with zero mean. The signal-to-noise ratio (SNR), defined as the ratio of the amplitude of the step to the standard deviation of the noise at each pixel, is $SNR = 0.2$, and the parameters used to detect the contour superimposed in the original images were $(\alpha = 0.08, \omega = 0.00008)$ and $(\alpha = 0.06, \omega = 0.00006)$, respectively. Observe how this algorithm yields excellent results for noisy images.

Figure 5a–d represents various classes of real-world images including aerial images and seismic data (figure 5d).

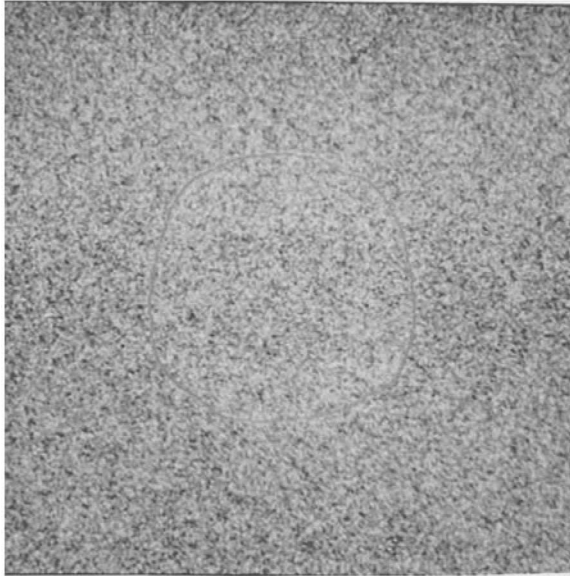
Figure 6a–d represents a stereo pair of classic indoor scenes that our vision group is working with.

Using our optimal operator with $\alpha = 1$ and $\omega = 0.0001$ on the original images shown in figures 5 and 6 yields the results represented in figures 7 and 8, respectively. The choice of α depends on the resolution at what the contours have to be detected. A value of α clustered to the unity tended to produce the best results for this class of images.

Figure 9 illustrates the effect of different values for ω while α is set constant ($\alpha = 1$). To help dis-



(a)



(b)

Fig. 4. Step edges in white Gaussian noise (SNR = 0.2): (a) straight line, and (b) Centered circle with a radius of 64 pixels.

cover what changed when ω decreased, all the maxima detected were set to 1. As expected, the number $N(\omega)$, of points that are marked as local maxima, decreases sensibly as ω becomes smaller. For the 256×256 aerial image of figure 9, we

determined the following experimental results:

$$\begin{aligned} N(2) &= 16672 & N(1) &= 15266 \\ N(0.5) &= 12263 & N(0.1) &= 10885 \\ N(0.01) &= 10811 & N(0.0001) &= 10807 \end{aligned}$$

Also included in figure 9 are the local maxima detected utilizing the first derivative of Gaussian ($\sigma = 1$) and the Canny's operator ($\sum \lambda = 1.12$). The projection function, in implementing Canny's operator, was a Gaussian with $\sigma = 1$. Our operator ($N(\omega) = 10807$) and Canny's operator ($N(\omega) = 11444$) ignore many of the noise-induced contour points that are reported as interesting by the first derivative of Gaussian ($N(\omega) = 13898$) while the width of the operators are the same. The local maxima detected with hysteresis thresholding utilizing the first derivative of Gaussian and the Canny's operator are shown in figure 10a and b so as to be compared to figure 8d. These images are interesting and confirm our earlier results in evaluating the overall performance indexes for each operator. They clearly illustrate the performance of our optimal operator, which has parameters easy to interpret, an overall performances index easy to calculate, and the additional advantage of being efficiently implemented in a recursive way.

8 Conclusion

In this article we presented two optimal edge detector in images. Their impulse responses are given by:

$$\begin{aligned} f(x) &= k \cdot e^{-\alpha \cdot |x|} \cdot \sin \omega \cdot x \quad \text{and} \\ g(x) &= k \cdot x e^{-\alpha \cdot |x|} \end{aligned} \quad (77)$$

We showed that these filters verify Canny's differential equation with its constraints. Their performances can be expressed easily as a function of α and ω , and are much better than the first derivative of a Gaussian (the filter preconized by Canny to approximate its optimal operator). For a particular choice of α and ω , we showed that these two filters present the same overall performance index. We proposed a highly recursive algorithm for implementing the convolution with $f(x)$, which results in a noise truncation immunity and a substantial saving in computational effort. The extension to the 2D case was pre-



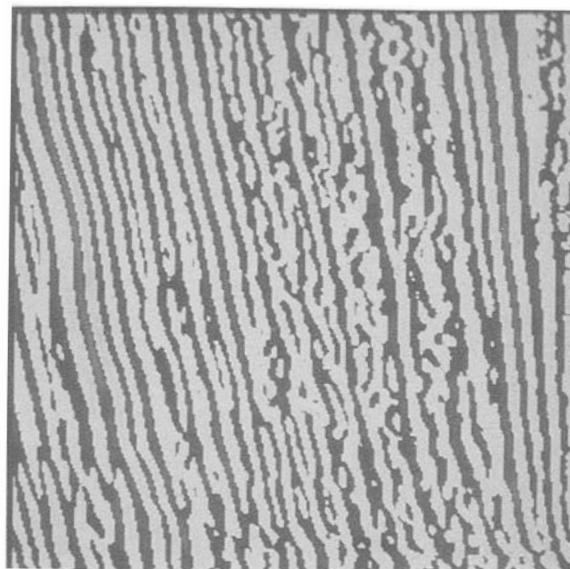
(a)



(b)

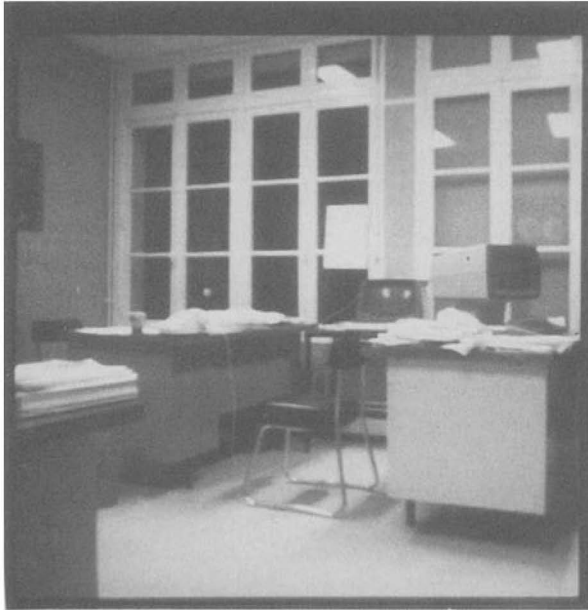


(c)

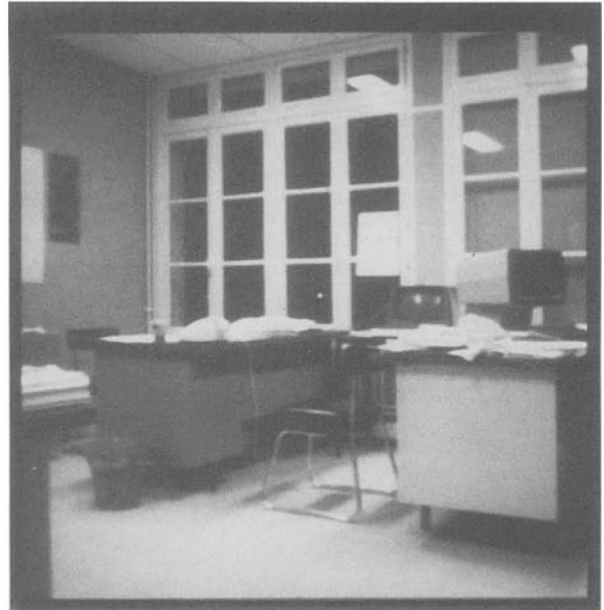


(d)

Fig. 5. (a–d) Various classes of real-world 256×256 image including aerial images and seismic data.



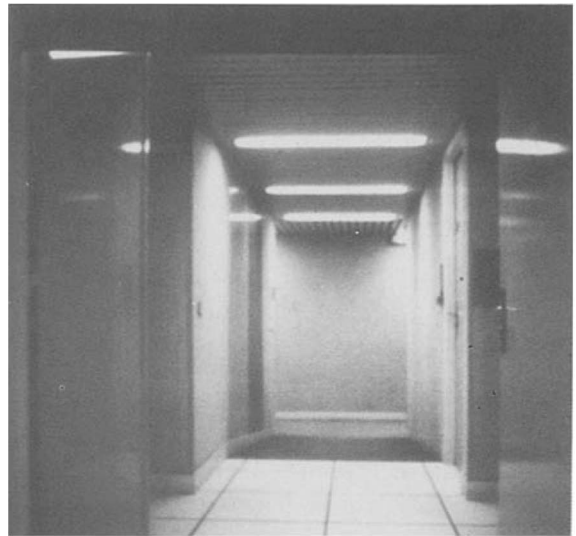
(a)



(b)



(c)



(d)

Fig. 6. (a-d) 512 × 512 stereo pair of classic indoor scenes.



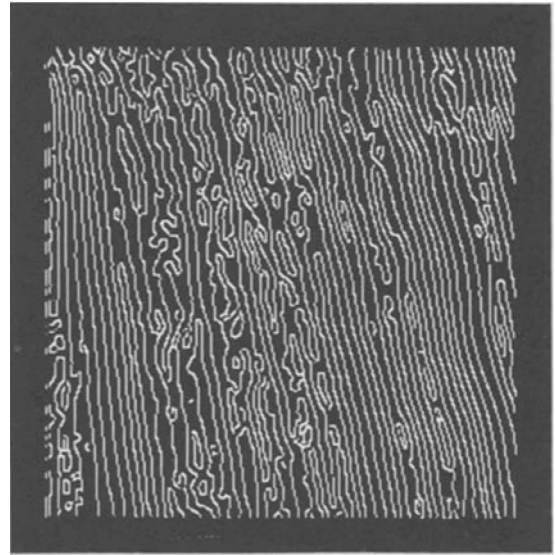
(a)



(b)

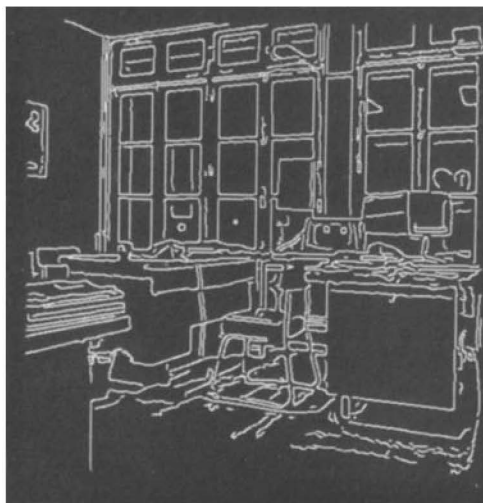


(c)



(d)

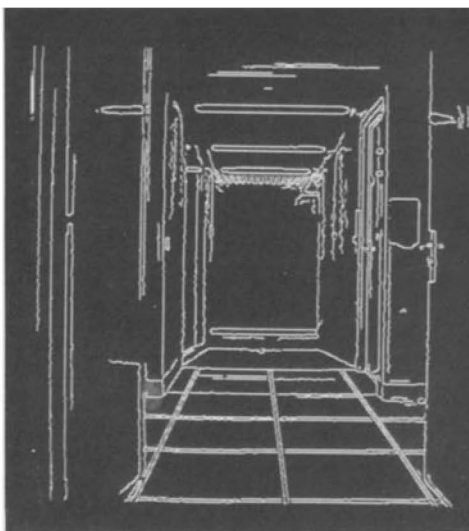
Fig. 7. Edges of figure 5 detected with $\alpha = 1.0$ and $\omega = 0.0001$.



(a)



(b)



(c)



(d)

Fig. 8. Edges of figure 6 detected with $\alpha = 1.0$ and $\omega = 0.0001$.

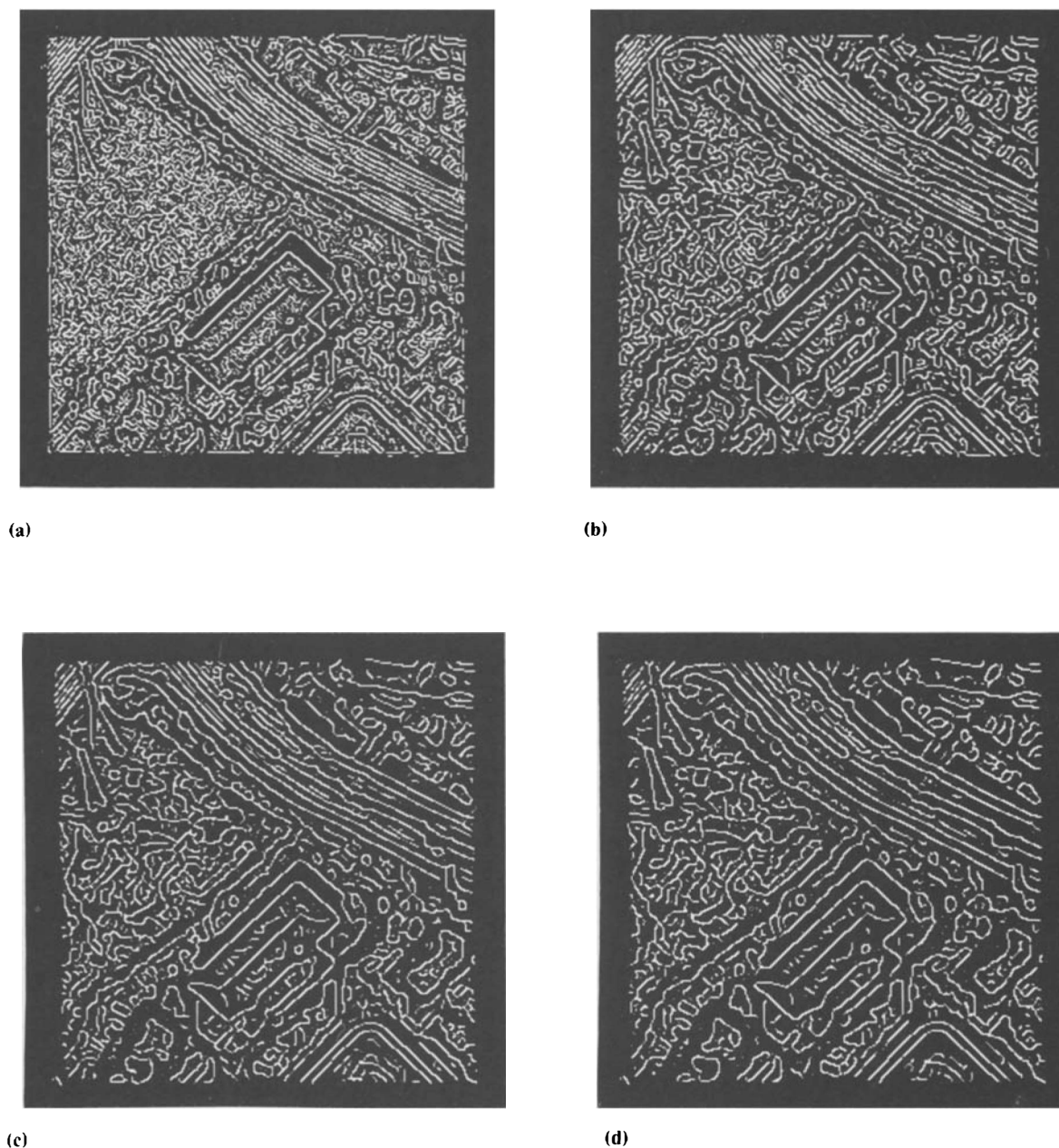


Fig. 9. Local maxima detected and set to 1 for the aerial image of figure 5c with $\alpha = 1.0$ and various values for ω : (a) $\omega = 2.0$, (b) $\omega = 1.0$, (c) $\omega = 0.5$, (d) $\omega = 0.1$, (e) $\omega = 0.01$, (f) $\omega = 0.0001$, (g) first derivative of a Gaussian with $\sigma = 1.0$, and (h) Canny's operator with a Gaussian projection function.



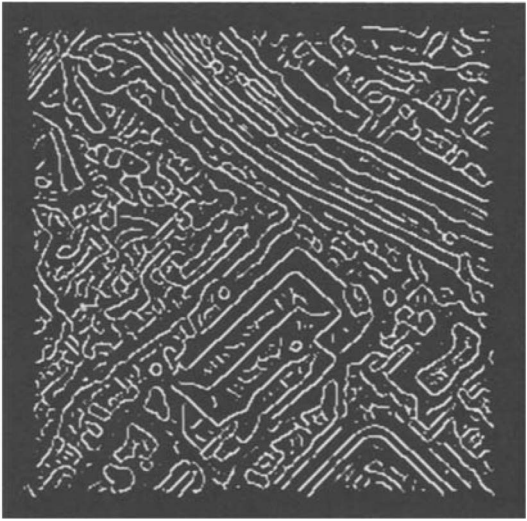
(e)



(f)



(g)



(h)



(a)



(b)

Fig. 10. Local maxima detected and thresholded with hysteresis: (a) utilizing the first derivative of a Gaussian with $\sigma = 1.0$, and (b) Utilizing the Canny's operator ($\sum \lambda = 1.12$).

sented and the results obtained by this algorithm were showed.

Acknowledgment

The author would like to thank Dr. O.D. Fau-

geras for his helpful suggestions and comments throughout this work.

References

1. J.F. Canny, "Finding edges and lines in images," M.I.T. Artif. Intell. Lab., Cambridge, MA, TR-720, 1983.
2. L.S. Davis, "A survey of edge detection techniques," *COMPUT. GRAPHICS IMAGE PROCESSING* vol. 4, pp. 248–270, 1975.
3. W.E.L. Grimson and E.C. Hildreth, Comments on "Digital step edges from zero crossings of second directional derivatives," *IEEE TRANS. PAMI* vol. PAMI-7, pp. 121–129, January 1985.
4. R.M. Haralick, "Digital step edges from zero crossings of second directional derivative," *IEEE TRANS. PAMI* vol. PAMI-6, pp. 58–68, January 1984.
5. Ellen C. Hildreth, "The detection of intensity changes by computer and biological vision systems," *COMPUT. VISION GRAPHICS IMAGE PROCESSING* vol. 22, pp. 1–27, 1983.
6. A. Rosenfeld and A.C. Kak, *DIGITAL PICTURE PROCESSING*, 2nd edn., vol. 2. Academic Press: New York, 1982.
7. J.W. Modestino and R.W. Fries, "Edge detection in noisy images using recursive digital filtering," *COMPUT. GRAPHICS IMAGE PROCESSING* vol. 6, pp. 409–433, 1977.
8. K.S. Shanmugan, F.M. Dickey, and J.A. Green, "An optimal frequency domain filter for edge detection in digital pictures," *IEEE TRANS. PAMI* vol. PAMI-1, pp. 37–49, 1979.
9. A. Witkin, "Scale-space filtering," in *PROC. 7TH INT. JOINT CONF. ARTIF. INTELL.*, Karlsruhe, FRG, 1983, pp. 1019–1021.
10. L.A. Spacek, "The detection of contours and their visual motion," PhD dissertation, University of Essex at Colchester, December 1985.
11. R. Deriche, "Separable recursive filtering for efficient multi-scale edge detection," in *PROC. INT. WORKSHOP INDUSTRIAL APPLICATIONS MACH. VISION INTELL.*, Roppongi, Tokyo, 2–4 February 1987.
12. T. Poggio, H. Voorhees, and A. Yuille, "A regularized solution to edge detection," M.I.T. Artif. Intell. Lab., Cambridge, MA, A.I. Memo. 833, May 1985.
13. L. R. Rabiner and B. Gold, *THEORY AND APPLICATION OF DIGITAL SIGNAL PROCESSING*. Prentice-Hall: Englewood Cliffs, NJ.
14. N.K. Bose, *DIGITAL FILTERS: THEORY AND APPLICATIONS*. Elsevier: New York.
15. J.B. Knowles and E.M. Olcayto, "Coefficient accuracy and digital filter response," *IEEE TRANS. CIRCUIT THEORY* vol. CT-15, no. 1, March 1968.