DETECTION OF EDGES USING LOCAL GEOMETRY

J. A. Gualtieri [1]
Computer Systems Research Facility
Code 635
NASA Goddard Space Flight Center
Greenbelt, MD

M. Manohar [2]
Information Systems Development Facility
Code 636
NASA Goddard Space Flight Center
Greenbelt, MD

1 Introduction

We seek a computational framework for detecting boundaries or edges present in gray level images. We are guided by two notions from psychophysics espoused by Koenderink and van Doorn [1] [2]:

- Primate visual function can be modeled by the activities of locally oriented receptive fields which are the second, third, and possibly fourth order derivatives of the Gaussian of scale t, $\phi_0(t) = \frac{exp(-\frac{(\xi^2+\eta^2)}{4t})}{4\pi t}$.
- In the visual system the natural coordinates on the retina, ξ, η are locally oriented by the direction of the gradient of the image smoothed by $\phi_0(t)$.

The activities of the receptive fields are given by convolving the image with the set of receptive fields at each point in the image, and from this collection of activities at each image point the local geometry is computed. These activities give a convenient representation of the image irradiance and are used to formulate an edge detector.

With the observation that the activities can be related to the Taylor series expansion of the image irradiance, I, about any point in the image we can then give mathematical forms – locally oriented derivatives –for local properties used to model edges and other features.

2 Mathematical Formalism

The receptive fields are denoted ϕ_1, ϕ_{11},ϕ_{12}, ϕ_{22}, ϕ_{111}, ϕ_{112}, ..., ϕ_{222} and computed by $\phi_{1\cdots12\cdots2}(\xi,\eta) = \frac{\partial^m}{\partial\xi^m}\frac{\partial^n}{\partial\eta^n}\phi_0(\xi,\eta;t)$, where there are m 1's and n 2's in the subscript of ϕ.

[1] National Research Council/Senior Resident Research Associate
[2] National Research Council/Resident Research Associate

While the receptive fields are defined over the infinite plane in actual use we choose a finite support size, $W \times W$, large enough so that the receptive field is very small at the window edge. In Fig. 1 the receptive fields shown are for $t = 2$ with $W = 21$ which give a value of 10^{-4} or smaller for the ratio of the value at the window edge to the maximum value over the entire window.

The receptive fields can be used to compute a finite Taylor series expansion of the smoothed image at each retinal cell, called the jet. The subscripts 1 and 2 denote directions in the local coordinate system along a level contour and along the gradient respectively where the contour direction is given by the convention $\hat{e}_\xi = \hat{e}_\eta \times \hat{e}_I$ where $\hat{e}_\xi, \hat{e}_\eta, \hat{e}_I$ are unit vectors in the direction of increasing ξ, η, I and the cross product follows the right hand rule convention. Fig. 2 shows the coordinate convention. The various receptive fields are named by their subscripts with the convention that the number of 1's is the order of the derivative in the contour direction and the number of 2's the order of the derivative in the gradient direction. The total number of subscripts is the order of the receptive field.

In real images noise and quantitization errors may prevent the Taylor series from being well defined of the image at point (x, y). However, if we use the *derived* image , $I \otimes \phi_0$, given by smoothing the Image with ϕ_0, then we may expand about (x, y) to third order to get

$$
\begin{aligned}
&[I \otimes \phi_0](\xi, \eta) = \\
&\quad [I \otimes \phi_0] + \xi \frac{\partial}{\partial \xi}[I \otimes \phi_0] + \eta \frac{\partial}{\partial \eta}[I \otimes \phi_0] \\
&\quad + \frac{\xi^2}{2!}\frac{\partial^2}{\partial \xi^2}[I \otimes \phi_0] + \xi\eta \frac{\partial^2}{\partial \xi \partial \eta}[I \otimes \phi_0] + \frac{\eta^2}{2!}\frac{\partial^2}{\partial \eta^2}[I \otimes \phi_0] \\
&\quad + \frac{\xi^3}{3!}\frac{\partial^3}{\partial \xi^3}[I \otimes \phi_0] + \frac{\xi^2}{2!}\eta \frac{\partial^3}{\partial \xi^2 \partial \eta}[I \otimes \phi_0] \\
&\qquad + \xi \frac{\eta^2}{2!}\frac{\partial^3}{\partial \xi \partial \eta^2}[I \otimes \phi_0] + \frac{\eta^3}{3!}\frac{\partial^3}{\partial \eta^3}[I \otimes \phi_0] \\
&\quad + O(\xi^4, \xi^3\eta, \xi^2\eta^2, \xi\eta^3, \eta^4)
\end{aligned}
$$

where all the derivatives and $[I \otimes \phi_0]$ are evaluated at the point (x, y).

For the convolution operator we have

$$
\begin{aligned}
\frac{\partial^{m+n}}{\partial \xi^m \partial \eta^n}[I \otimes \phi_0] &= [\frac{\partial^{m+n}}{\partial \xi^m \partial \eta^n} I \otimes \phi_0] \\
&= [I \otimes \frac{\partial^{m+n}}{\partial \xi^m \partial \eta^n}\phi_0] \\
&= [I \otimes \phi\ \mathbf{w}]
\end{aligned}
$$

where $\phi\ \mathbf{w}$ is shorthand for $\phi_{1\cdots12\cdots2}(\xi, \eta)$. This shows that we may compute activities of the receptive fields and thereby the Taylor series by performing convolutions with the

receptive fields. At this point we have made no commitment to the orientation of the local coordinates ξ, η.

3 Choosing the Local Coordinates

Suppose then as it is proposed by Koenderink [2] that the biological visual system makes no commitment as to the coordinate system it will use, but rather chooses to use all coordinate systems. This it does by measuring the receptive fields activities over many directions $-\pi < \theta \leq \pi$. Thus the approximate continuum of quantities (as a function of θ)

$$a^{rot}_{\mathbf{w}}(\theta) = I \otimes R_\theta(\phi\ \mathbf{w})$$

is available to the low level vision system. Here $\mathbf{w}$ is the receptive field name and $R_\theta(\cdot)$ is an operator that rotates its argument through an angle θ. Assuming the image varies smoothly, the maxima and zeroes of these activities define locally meaningful directions. For example, at a particular image point, the angle θ for which $a^{rot}_2(\theta)$ is a maximum defines a local direction with respect to the image x direction that is along the contour, the 1 direction, at that image point. Note for this angle that $a^{rot}_1(\theta) = 0$. Similarly, at a particular image point, the angle θ' for which $a^{rot}_1(\theta')$ is a maximum defines a local direction with respect to the image x direction that is along the gradient, the 2 direction, at that point.

The activities of the receptive fields incorporate the local geometry and provide a useful representation in terms of which to formulate edge detectors. Unlike biological vision, machine vision is typically presented with a much sparser set of activities – those activities of receptive fields defined by derivatives along the image x and y directions. We make contact with biological vision by defining the local coordinate 2 to be in the direction $\theta_{loc} = tan^{-1}(I \otimes \frac{\partial}{\partial y}\phi_0 / I \otimes \frac{\partial}{\partial x}\phi_0)$ and we then compute the activities of the receptive fields in this particular choice of local coordinates:

$$a^{rot}_{\mathbf{w}} = I \otimes R_{\theta_{loc}}(\phi \mathbf{w}).$$

4 Edge Detection and Local Features

A simple edge detector finds candidate edges points as those points where the gradient is a local maximum. The Canny edge detector [3], [4] in doing this pays particular attention to accurately finding the direction of the gradient at each pixel and then to doing a careful interpolation of the change in gradient along this direction so as to find its local maximum.

If the image varies smoothly this is equivalent to locating the zero crossing in the change of the gradient along the gradient direction. In local geometry this is $a^{rot}_1 = 0$, $a^{rot}_2 > 0$, $a^{rot}_{22} = 0$, which means respectively: align 1 along the contour direction, consider only points with nonzero gradient, locate the edge at the zero crossing of a^{rot}_{22}.

For comparison, the Marr-Hildreth edge detector seeks zero crossings of the Laplacian of a Gaussian [5] and is given by $a_{11}^{rot} + a_{22}^{rot} = 0$. This is rotationally invariant indicating that it contains less local geometry than the Canny detector and thereby can be expected not to perform as well as the Canny detector.

Besides exploiting properties in the gradient direction we can compute properties along the contour direction. Any contour of the Image irradiance satisfies implicitly the equation $I(x,y) \otimes \phi_0 = I_0$, where I_0 is value of the smoothed irradiance that defines the contour. Along the contour in image coordinates we have

$$\frac{d}{dx}\Big|_c = \frac{\partial}{\partial x} + \frac{dy}{dx}\Big|_c \frac{\partial}{\partial y}.$$

Since $I \otimes \phi_0$ is constant on the contour we have $(\frac{d}{dx}|_c)^n I \otimes \phi_0 = 0$, for all n . In particular we have for $n = 1$ and $n = 2$, respectively

$$\begin{aligned} a_1 + \frac{dy}{dx}\Big|_c a_2 &= 0 \\ a_{11} + 2a_{12}\frac{dy}{dx}\Big|_c + a_{22}(\frac{dy}{dx}\Big|_c)^2 + a_2\frac{d^2y}{dx^2}\Big|_c &= 0 \end{aligned}$$

If we take the local coordinate system so that the contour lies along the 1 direction, then $\frac{d\eta}{d\xi}|_c = 0$, $a_1^{rot} = 0$, and $\frac{d^2\eta}{d\xi^2}|_c = -a_{11}^{rot}/a_2^{rot}$ which is the curvature along the contour.

Another useful result is obtained by setting $\frac{d\eta}{d\xi}|_c = 0$ in $\frac{d^3}{d\xi^3}|_c I \otimes \phi_0 = 0$ which leads to $a_{111}^{rot}a_2^{rot} - a_{12}^{rot}a_{11}^{rot} = 0$, a result given by Koenderink and van Doorn [1]. Points that satisfy this criterion are called *ridges* and can be identified with corner points along edge directions.

5 Accuracy of the Representation

In accordance with the notion that the local geometry accurately represents the local image structure we expect it should be possible to locate edges to sub-pixel resolution. This approach uses the local geometry to *model* a smoothed representation of the image. Thus the gradient directions are not quantitized by the original pixel lattice (angles quantitized to fall into multiples of $0 < \theta \leq \pi/4$), but are accurately given to a fraction of a degree. Similarly we expect that the location of zero crossings to be given to a finer resolution than an individual pixel.

We have tested this hypothesis by locating a zero crossing along a given direction using an interpolation given by Canny [4]. In Fig. 3 a discretely sampled function h, with $0 < \theta < \pi/4$ is given by values $h(i,j), h(i+1,j)$, and $h(i+1,j+1)$ with $h(i,j) < 0$, and $h(i+1,j), h(i+1,j+1) > 0$. The value $h^{int} = (1 - tan\theta)h(i+1,j) + tan\theta h(i+1,j+1)$

is the linear interpolation of h in the direction θ, and the zero crossing is located at a distance $d = -h(i,j)(1 + tan^2\theta)^{\frac{1}{2}}/(h^{int} - h(i,j))$. Note in this case the zero crossing falls outside the pixel (i,j). Similar interpolation formulas hold for other directions of θ.

In what follows we calculate for each candidate edge pixel the values θ and d. To display the sub-pixel location of the edges we have found, the pixel in which the zero crossing point falls is dilated by a factor of 5 so that each original pixel is equivalent to 25 sub-pixels. The edge is then drawn as a *digital line* so as to pass through the zero crossing at an angle perpendicular to θ. The digital line marks only those sub-pixels that contain the line. We do not extend the line outside the original pixel which contains the zero crossing. This we consider a crude approximation, but the results below bear out the claims of sub-pixel accuracy.

6 Results

We have constructed the local geometry of a simple synthetic image of a rectangle, of step edge with additive Gaussian noise, and of a SAR image (substantially subsampled to remove speckle) taken from SEASAT of ice floes.

Synthetic Image of a Rectangle

For the synthetic image in Fig. 4 the various activities of the receptive fields locate local properties in the image. The zero crossings of the activities of the receptive fields a_{22}^{rot} where $a_2^{rot} > 0$ can be seen to locate the edges of the object while the zero crossings of $a_{111}^{rot}a_2^{rot} - a_{12}^{rot}a_{11}^{rot}$, the *ridge detector* locate the corners. With so little structure in the image it is difficult to give further meaning to the other receptive field activities. The edges found lie to sub-pixel accuracy exactly along the rectangle sides while at the corners they are rounded reflecting the effect of the smoothing by convolution with ϕ_0.

Noisy Step Edges

We created synthetic noisy step edges by adding Gaussian random noise of zero mean and variance one to step edges of varying height. Images were then scaled to the grey scale range of 0 to 255. Defining the signal to noise ratio (SNR) [6] as the ratio of the square step height to the variance of the Gaussian noise we have found, see Fig. 5, that for SNR ≤ 1 the edge becomes broken while for SNR > 1 the edge is continuous. Here the window size and receptive field sizes are, $W = 21$, and $t = 2$. In addition zero crossings of a_{22} are considered only if the values of a_2 exceed 0.1 $\max(a_2)$.

As can be seen the edge wanders about the *true* edge but remains smooth and continuous. Roughly then , when the SNR excceeds 2, we expect this edge detector to perform reasonably.

Ice Image

The ice image in Fig. 6 is a 256×256 grey scale image dilated in size from a 128×128 original image. This was processed using receptive fields of size $W = 21$ and $t = 2$. The dilation was done to reduce numerical errors that would have resulted from using receptive fields of size $W = 11$ and $t = 1$ on the original 128×128 image. As for the noisy step edges, zero crossings of a_{22} are considered only if the values of a_2 exceed $0.1 \max(a_2)$.

Fig. 6 shows I, $I \otimes \phi_0$, a_2^{rot}, a_{22}^{rot} in the top four images and edges located to one pixel accuracy in the lower left. In the lower right are shown the values of d, the distance of the zero crossing from the center of the found edge pixels, coded by intensity with high intensity corresponding to larger d. The large variation in d suggests that the local geometry contains more information that can be used to locate the edge, indeed to sub-pixel accuracy.

To examine the sub-pixel accuracy we have enlarged the upper quadrant of the ice image in Fig 7. As can be seen, the edges found form a smooth almost continuous boundary to the butterfly shaped island and other regions in the image. We take this as evidence that the local geometry contains sufficient information to locate edges to a sub-pixel accuracy that increases resolution by a factor of five.

7 Summary

We have described a new representation, the local geometry, for early visual processing which is motivated by results from biological vision. This representation is richer than is often used in image processing. It extracts more of the local structure available at each pixel in the image by using receptive fields that can be continuously rotated and that go to third order in spatial variation. Early visual processing algorithms such as edge detectors and ridge detectors can be written in terms of various local geomtries and are computationally tractable. For example, Canny's edge detector has been implemented in terms of a local geometry of order two, and a ridge detector in terms of a local geometry of order three.

The edge detector in local geometry was applied to synthetic and real images and it was shown using simple interpolation schemes that sufficient information is available to locate edges with sub-pixel accuracy (to a resolution increase of at least a factor of five). This is reasonable even for noisy images because the local geometry fits a smooth surface –the Taylor series – to the discrete image data.

Only local processing was used in the implementation so it can readily be implemented on parallel mesh machines such as the MPP [7]. We expect that other early visual algorithms, such as region growing, inflection point detection, and segmentation can also be implemented in terms of the local geomtry and will provide sufficiently rich and robust representations for subsequent visual processing.

References

[1] J. J. Koenderink and A. J. van Doorn, *Rperesentation of Local Geometry in the Visual System*, Biological Cybernetics, **55**, 367-375 (1987).

[2] J. J. Koenderink, *Operational Significance of Receptive Field Assemblies*, Biological Cybernetics, **58**, 163-171 (1987).

[3] J. Canny, *A Computational Approach to Edge Detection*, PAMI-8,(1986).

[4] J. Canny, *Finding Edges and Lines in Images*, MIT AI Lab, TR No. 720 (1983).

[5] D. C. Marr and E. Hildreth, *Theory of Edge Detection*, Proc. Roy. Soc. Lond. **B 207**, 187-217 (1980).

[6] W. K. Pratt, *Digital Image Processing*, John Wiley and Sons, 495-500 (1978).

[7] J. R. Fisher (ed.), *Frontier of Massively Parallel Scientific Computation*, NASA Conference Publication 2478 (1986).

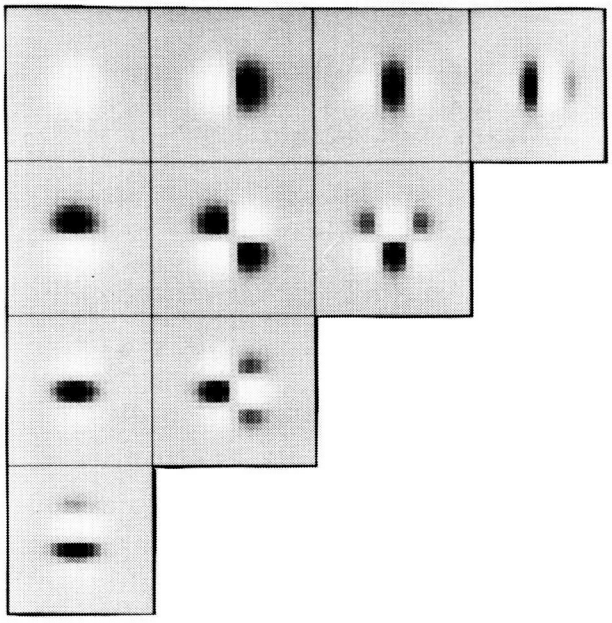

Figure 1: Receptive Fields with $W = 21$ (window size) and $t = 2$ arranged according to:

ϕ_0	ϕ_1	ϕ_{11}	ϕ_{111}
ϕ_2	ϕ_{12}	ϕ_{112}	
ϕ_{22}	ϕ_{122}		
ϕ_{222}			

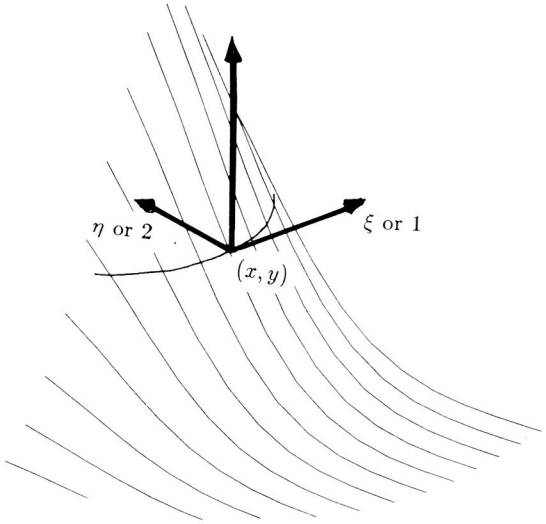

Figure 2: The local coordinates ξ, η (also named 1 2) of the image at a point x, y. The surface shown is the image irradiance versus versus x, y.

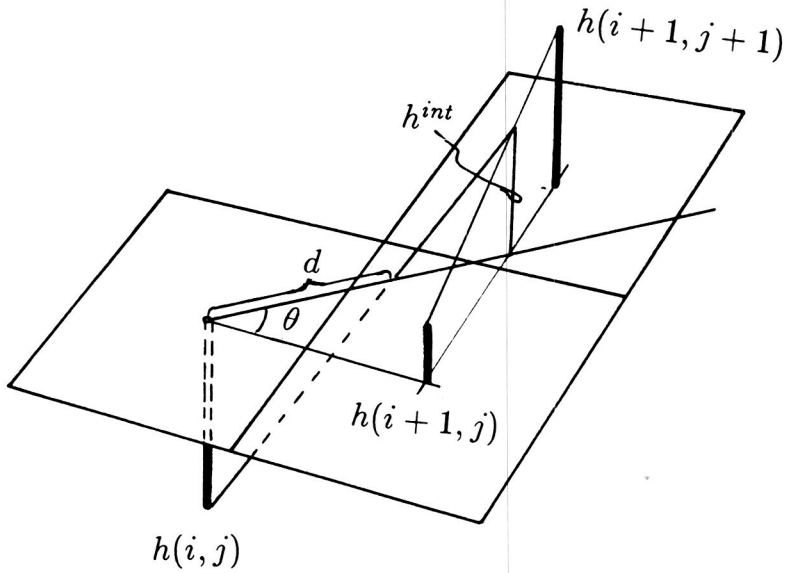

Figure 3: The interpolation scheme used to locate zero crossing.

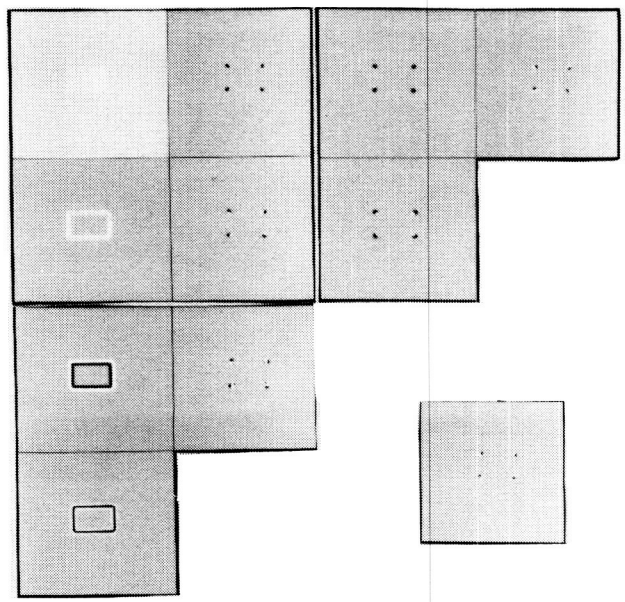

Figure 4: Activities of the receptive fields for the rectangle image, where the size of the rectangle is 64×44 pixels, and window size $W = 21$, $t = 2$ according to:

Original Image	a_1^{rot}	a_{11}^{rot}	a_{111}^{rot}
a_2^{rot}	a_{12}^{rot}	a_{112}^{rot}	
a_{22}^{rot}	a_{122}^{rot}		
a_{222}^{rot}			Ridge Points: $a_{111}^{rot} a_2^{rot} - a_{12}^{rot} a_{11}^{rot} = 0$

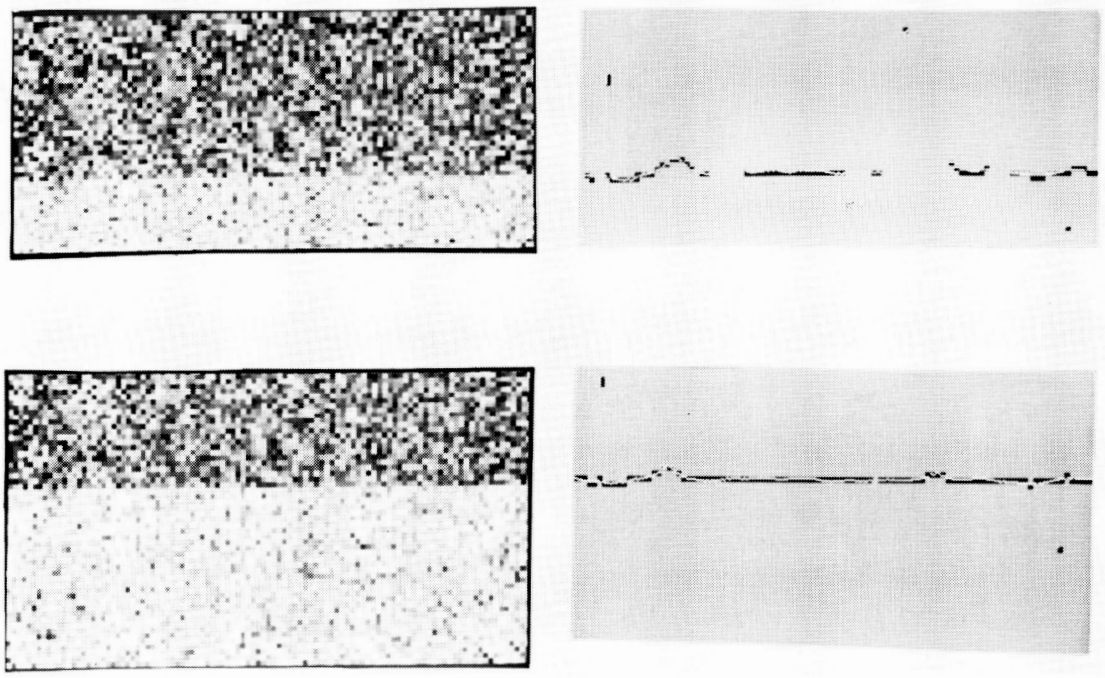

Figure 5: Synthetic noisy step edges on the left and edges found on the right. Upper pair is for SNR = 1 and lower pair for SNR = 2.

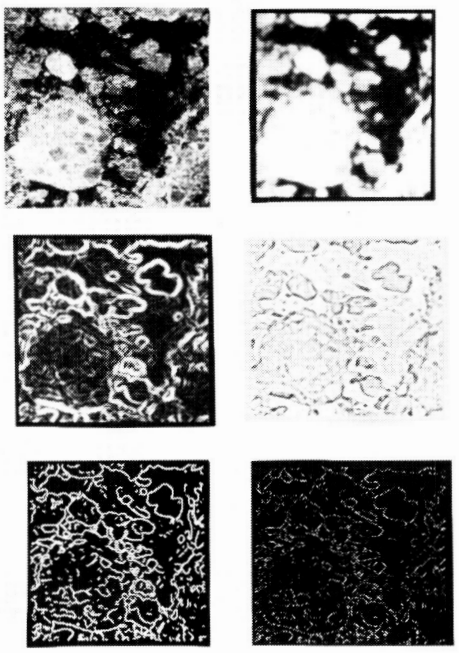

Figure 6: Results for the ice floe image according for $W = 21, t = 2$ according to:

Original Image, I	$I \otimes \phi_0$
a_2^{rot}	a_{22}^{rot}
edge pixels	d for edge pixels

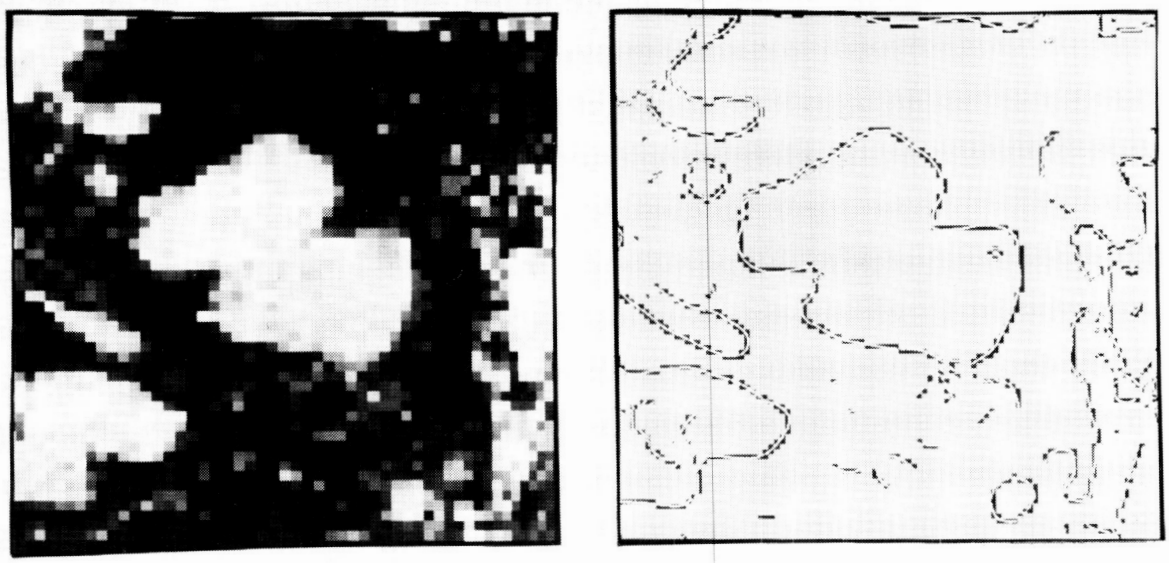

Figure 7: Enlarged view of ice image and edges found.