MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo 722 June, 1983

SCALING THEOREMS FOR ZERO-CROSSINGS

A.L. Yuille and T. Poggio

Abstract. We characterize some properties of the zero-crossings of the laplacian of signals - in particular images - filtered with linear filters, as a function of the scale of the filter (following recent work by A. Witkin, 1983). We prove that in any dimension the only filter that does not create zero-crossings as the scale increases is the gaussian. This result can be generalized to apply to level-crossings of any linear differential operator: it applies in particular to ridges and ravines in the image intensity. In the case of the second derivative along the gradient we prove that there is no filter that avoids creation of zero-crossings.

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.

1. Introduction

In most physical phenomena, changes in spatial or temporal structure occur over a wide range of scales. Images are no exception: changes in light intensity reflect the many spatial scales at which visible surfaces are organized. It seems intuitive that a great deal of information can be gained by an analysis of the changes in a signal at different scales. For instance, graphs of one-dimensional functions are a very effective tool for describing complex systems. An important reason is that they allow direct visual access to important properties of the data, chiefly to their changes over different scales.

The idea of scale is critical for a symbolic description of the significant changes in images or other types of signals. Changes must be detected at different levels of detail and over different extents. In general different physical processes may be associated with a characteristic behaviour across different scales. In an image, changes of intensity take place at many spatial scales depending on their physical origin. A multiscale analysis, tracing the behaviour of some feature of the signal across scales, can reveal precious information about the nature of the underlying physical process. In images, for instance, spatial coincidence at all scales of zero-crossings in the Laplacian of the intensity values filtered with a gaussian mask, signals a physical "edge", distinct from surface markings or shadows. Not only is it necessary to detect and describe changes in a signal at different scales, but in addition, much useful information can be obtained by combining descriptions across scales.

The importance of this idea has been clearly realized in the field of vision. One of the main contributions of visual psychophysics in the last 10 years was indeed to show that visual information is processed in parallel by a number (perhaps a continuum) of spatial-frequency-tuned channels (Campbell & Robson, 1968). The bulk of the data demonstrates that the visual system analyses the image at different resolutions. Physiological experiments are consistent with the psychophysics. They suggest that in the visual pathway spatial filters of different size operate at the same location. Furthermore, psychophysics, physiology and anatomy all show that the spatial grain of analysis continuously changes from foveal to peripheral locations. Receptive and dendritic field sizes of both retinal and cortical neurons increases monotonically with eccentricity, in agreement with the dependency on eccentricity of the psychophysical channels.

In the field of computer vision, Rosenfeld was one of the first to propose explicitly an edge detection scheme-based on multiscale analysis performed with filters of different sizes (Rosenfeld and Thurston, 1971). A similar algorithm was suggested by Marr (1976) though with different goals and motivations. More recently, he has strongly advocated the use of derivatives of gaussian-shaped filters of different sizes with the goal of detecting changes in intensity at different scales (Marr, 1982). The idea was first proposed in the context of a theory of stereomatching (Marr and Poggio, 1979). In that scheme, analysis at the different scales was effectively kept separate. Later, Marr and Hildreth (1980) proposed some heuristical rules to combine information from the different channels. However, the important problem of how to combine effectively the different scales of analysis at thisn early level has remained open , although recent work by D. Terzopoulos (1982) has successfully applied multi-level algorithms to the problem of reconstructing visual surfaces (see also the work by Richards et al., 1982 and by Canny, 1983 on edge detection). In a recent conference (Cold Spring Harbour, April 1983) we learned from A. Witkin a new way of describing zero-crossings across scale.[1]

A 1-dimensional signal is smoothed by convolution with a small (large) gaussian filter and the zeros of the second derivative are localized and followed as the size of the filter increases (decreases). This procedure originates a plot of the zero contours in the $x - \sigma$ plane (where σ measures the size of the gaussian filter).[2] In this way, Witkin was able to classify and label zero-crossings achieving an effective description of a signal for purposes of recognition and registration. This is possible mainly because the geometry of the zero contours is surprisingly simple. Zero-contours are either lines from small to very large scale or closed, bowl-like shapes. Zero-crossings are never created as the scale increases. Witkin mentioned the striking result (obtained by J. Babaud) that the gaussian filter is the only filter with this remarkable property in 1-D (at the same conference J. Koenderink told us that he has obtained similar results exploiting properties of the diffusion equation).[3]

We have now succeded in obtaining a proof of this result in 2D (and in fact any number of dimensions). We have also obtained related results for zero- and level-crossings of other differential operators, in particular for ridges and ravines in the image intensity.

The 2-D result seems important because it:

(a) lays the necessary mathematical foundation for using multiresolution labels for classifying zero-crossings for a symbolic description of intensity changes.

(b) justifies the use of gaussian filters and an associated linear derivative because of their "nice" properties under changes in scale.

In this paper, we will first state and prove the one-dimensional result. We will then show that only a specific 2-D extension is valid. Zero-crossing of linear derivatives have the "nice scaling behaviour" if and only if the image is filtered by a 2-D rotationally symmetric gaussian. In particular, the laplacian-of- a-gaussian filter suggested by Marr and Hildreth has nice scaling behaviour. The second directional derivative along the gradient, however, does not: no filter exists that can ensure a nice scaling behaviour of the zeros of this derivative. We have then, the following results:

[1] Witkin's prize-winning paper will appear in the 1983 IJCAI Proceedings (Witkin, 1983). We received a preprint after this memo went to press.

[2] J. Stansfield first described — for analysing commodities trends (Stansfield, 1980) — the idea of plotting zero-crossings over scale, but did not develop it.

[3] After completion of this memo we were informed that a technical report containing the 1D proof is now ready, with the title "Uniqueness of the gaussian kernel for scale-space filtering," by J. Babaud, A. Witkin and R. Duda, Fairchild TR 645, Flair 22).

(a) for linear derivative operations—in particular, for the laplacian—the gaussian is the only filter with a nice scaling behaviour.

(b) for the nonlinear directional derivative, no filter will give nice scaling behaviour.

2. Assumptions and results

We will consider filtering the image I with a suitable filter F and then consider the behaviour of the zero crossings as we change the scale of the filter. We make five assumptions about the filter, and impose them as boundary conditions.

(1) Filtering is shift-invariant and, hence, a convolution. We write this as

$$F * I(\underline{x}) = \int F(\underline{x} - \underline{\varsigma}) I(\underline{\varsigma}) d\underline{\varsigma}.$$

(2) The filter has no preferred scale length. In two dimensions standard results of dimensional analysis (Bridgman, 1922) give $F(\underline{x}, \sigma) = \frac{1}{\sigma^2} f(\frac{\underline{x}}{\sigma})$, where σ is the scale of the filter. The factor $\frac{1}{\sigma^2}$ ensures that the filter is properly normalized at all scales.

(3) The filter recovers the whole image at sufficiently small scales. This is expressed by $Lim_{\sigma \mapsto 0}, F(\underline{x}, \sigma) = \delta(\underline{x})$, where $\delta(\underline{x})$ denotes the Dirac delta function.

(4) The position of the centre of the filter is independent of σ. Otherwise, zero crossings of a step edge would change their position with change of scale.

(5) The filter goes to zero as $|\underline{x}| \mapsto \infty$ and as $\sigma \mapsto \infty$.

As will become apparent, our results are independent of scaling the x axis. We usually require that we scale this axis so that the filter is radially symmetric, and state theorems with respect to such axes. However, we can relax this requirement by rescaling the axes.

Figure (1) shows the typical scaling behaviour of zero crossings in one dimension observed by Witkin. Figure (2) shows possible behaviour of zero crossings which is never empirically observed when the filter is a Gaussian. The generic properties of the zero-crossings curves in the x, σ plane can be derived from the Implicit Function Theorem. To yield a C^r curve the theorem requires that the Laplacian of the filtered image is C^r. Therefore the filter must be reasonably smooth. Observe that filtering with a gaussian will ensure a C^∞ output for all images, because of the equivalence with the Cauchy problem for the diffusion equation. The Implicit Function Theorem may break down at degenerate critical points when all first derivatives of the filtered image vanish together with the Hessian.[4] These points are non-generic in the sense that a small perturbation will destroy them. Observe that "true" zero-crossings can only disappear in pairs in the x, σ plane. Only trivial zeros that do not cross zero can disappear by themselves. They are, however, non generic.

In one-dimension, the zero crossings obey

[4] Zeros of the Hessian correspond to zeros of the gaussian curvature.

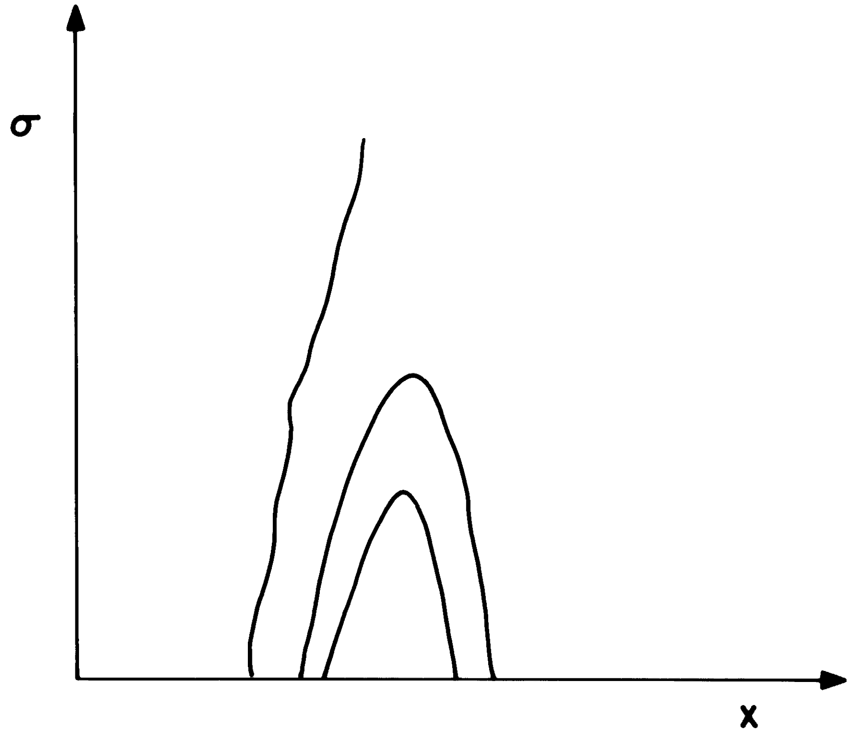

Figure 1 See text.

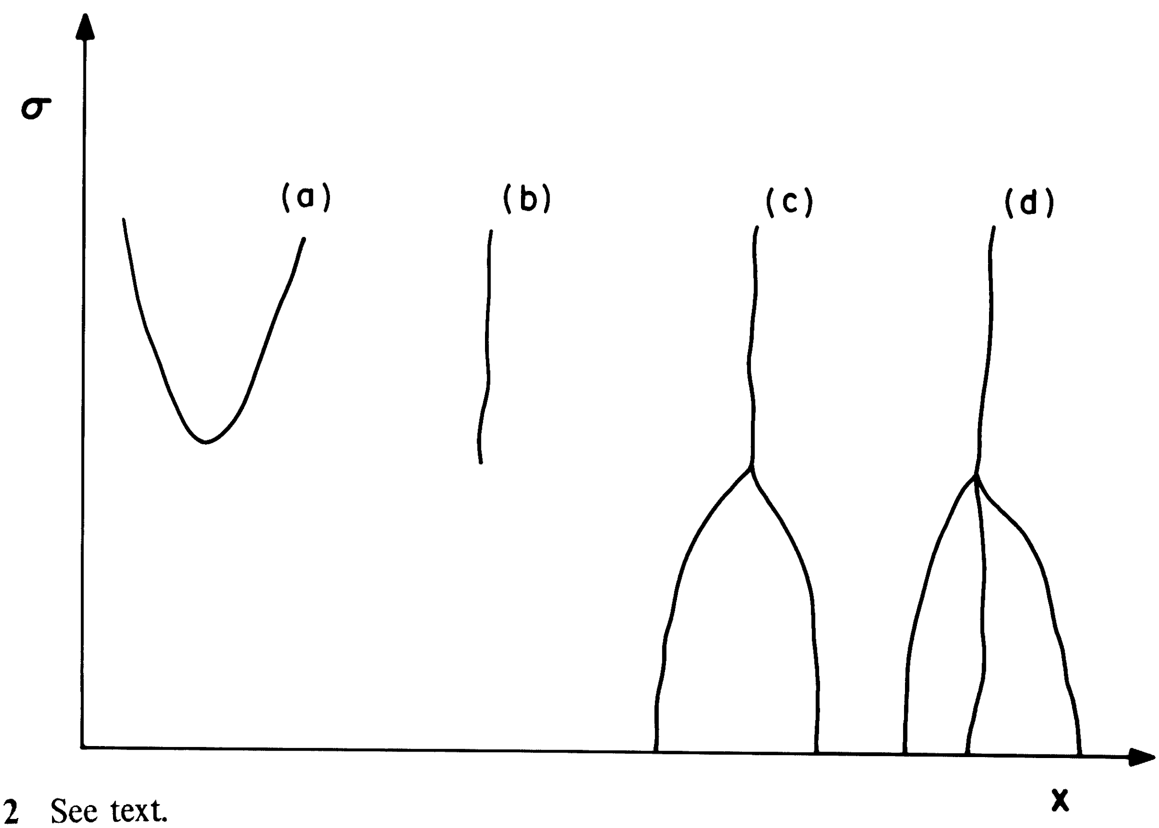

Figure 2 See text.

$$0 = \int_{-\infty}^{\infty} f''\Big(\frac{x-\varsigma}{\sigma}\Big) I(\varsigma)d\varsigma. \tag{2.1}$$

This equation gives x as an implicit function of σ, i.e., $x = x(\sigma)$. If we vary x and σ so that (2.1) is still satisfied, we obtain

$$\frac{dx}{d\sigma} = \frac{\int_{-\infty}^{\infty} \big(\frac{x-\varsigma}{\sigma}\big) f'''\big(\frac{x-\varsigma}{\sigma}\big) I(\varsigma)d\varsigma}{\int_{-\infty}^{\infty} f'''\big(\frac{x-\varsigma}{\sigma}\big) I(\varsigma)d\varsigma} \tag{2.2}$$

So the tangent to the curve is uniquely defined at a point, as are all the higher order derivatives. This prevents the behaviour shown in Figures 2b, 2c with the possible exception of the nongeneric cases, when the Implicit Function Theorem breaks down.

The curve in Figure 2(a) is more interesting because it corresponds to a pair of zero crossings being "created" as the scale increases. The Implicit Function Theorem does not rule out this case. It therefore seems natural to require a filter such that this never occurs. In the following three sections, we will prove some theorems showing that such a filter can only be a gaussian and, moreover, that not all differential zero-crossings operators can obey this property. More precisely, we prove:

Theorem 1. In one-dimension, with the second derivative, the gaussian is the only filter—obeying our five boundary conditions—which never creates zero crossings as the scale increases.

Theorem 2. In two-dimensions, with the laplacian operator, the gaussian is the only filter obeying the boundary conditions which never create zero crossings as the scale increases.

Theorem 3. In two-dimensions, with the directional derivative along the gradient, there is no filter obeying the boundary conditions which never creates zero crossing as the scale increases.

In section (5), we show that results similar to Theorems 1 and 2 can be extended to all linear differential operators (in particular, directional derivatives) and therefore to other features of the image, such as ravines and ridges (but not peaks) in the image intensity. These theorems can be extended to any dimension, but we will not give these extensions here.

It should be emphasized that, although zero crossings can only annihilate themselves in pairs, the intensity change corresponding to a zero crossing could become arbitrarily smaller as sigma increases. The zero crossing would then become so weak that for practical purposes the curve may terminate.

3. The 1-D case

Let the image be I and the filter be F. We consider the zero crossings in the filtered image.

$$F * I(x) = \int_{-\infty}^{\infty} F(x - \varsigma) I(\varsigma) d\varsigma \tag{3.1}$$

Denote $\frac{d^2}{dx^2}(F * I)$ by E. Hence the zero crossings are the solutions of

$$E(x) = 0. \tag{3.2}$$

These form curves in the $x - \sigma$ plane. The condition that zero crossings are not created at larger scales is that for all such curves $\sigma(x)$ the extrema of $\sigma(x)$ are not minima. Hence, for all points x_o such that $\sigma'(x_o) = 0$, we require that $\sigma''(x_o) < 0$. Let t be a parameter along a curve in $\sigma - x$ space. Then

$$\frac{dE}{dt} = \frac{\partial E}{\partial x}\frac{dx}{dt} + \frac{\partial E}{\partial \sigma}\frac{d\sigma}{dt}. \tag{3.3}$$

On a curve of zero crossings, $E = 0$, and so $\frac{dE}{dt} = 0$ on such a curve. We can choose the parameter t to be x. Then, using the Implicit Function Theorem, we obtain:

$$\frac{d\sigma}{dx} = \frac{-E_x}{E_\sigma}. \tag{3.4}$$

This result vanishes at x_o if and only if

$$E_x(x_o) = 0, \tag{3.5}$$

and we calculate

$$\frac{d^2\sigma(x_o)}{dx^2} = \frac{-E_{xx}(x_o)}{E_\sigma(x_o)}. \tag{3.6}$$

Thus, our filter must be such that if

$$E(x_o) = E_x(x_o) = 0 \tag{3.7}$$

then

$$\frac{E_{xx}(x_o)}{E_\sigma(x_o)} > 0. \tag{3.8}$$

The Diffusion Equation can be written as

$$\frac{\partial^2 E}{\partial x^2} = \frac{1}{\sigma}\frac{\partial E}{\partial \sigma}. \tag{3.9}$$

Note that by the substitution $t = \frac{\sigma^2}{2}$, we obtain the standard diffusion equation. If the filter F is a gaussian,

$$F(x) = \frac{1}{\sigma} exp\{\frac{-x^2}{2\sigma^2}\}, \tag{3.10}$$

then it will obey the diffusion equation of which it is the Green function and hence $E(x)$ will also obey the equation. Thus, $\frac{E_{xx}}{E_\sigma} = \frac{1}{\sigma}$ and so a gaussian filter will always satify conditions (3.7) and (3.8).

We now show that the gaussian is the *only* filter which satisfies the conditions and obeys the boundary conditions specified in section (1).

Consider an image which is the sum of delta functions:

$$I(\varsigma) = \sum_{i=1}^{n} A_i \delta(\varsigma - \varsigma_i). \tag{3.11}$$

It is possible to generate any image in this way by taking the limit as $n \mapsto \infty$.

Set

$$T(x) = F_{xx}(x). \tag{3.12}$$

Equations (3.7) and (3.8) yield

$$\sum_{i=1}^{n} A_i T(x_o - \varsigma_i) = 0 \tag{3.13}$$

$$\sum_{i=1}^{n} A_i T_x(x_o - \varsigma_i) = 0 \tag{3.14}$$

and

$$\frac{\sum_{i=1}^{n} A_i T_{xx}(x_o - \varsigma_i)}{\sum_{i=1}^{n} A_i T_\sigma(x_o - \varsigma_i)} > 0. \tag{3.15}$$

We can construct a counter example if we can solve the simultaneous equations for any $x_o, \varsigma_1 .. \varsigma_n$ and any positive ℓ^2:

$$\sum_{i=1}^{n} A_i T(x_o - \varsigma_i) = 0. \tag{3.16}$$

$$\sum_{i=1}^{n} A_i T_x(x_o - \varsigma_i) = 0. \tag{3.17}$$

$$\sum_{i=1}^{n} A_i T_{xx}(x_o - \varsigma_i) = -\ell^2. \tag{3.18}$$

$$\sum_{i=1}^{n} A_i T_\sigma(x_o - \varsigma_i) = 1. \tag{3.19}$$

We can write these as a matrix equation:

$$\begin{pmatrix} T(x_o - \varsigma_1) & \dots & T(x_o - \varsigma_n) \\ T_x(x_o - \varsigma_1) & \dots & T_x(x_o - \varsigma_n) \\ T_{xx}(x_o - \varsigma_1) & \dots & T_{xx}(x_o - \varsigma_n) \\ T_\sigma(x_o - \varsigma_1) & \dots & T_\sigma(x_o - \varsigma_n) \end{pmatrix} \begin{pmatrix} A_1 \\ \cdot \\ \cdot \\ A_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -\ell^2 \\ 1 \end{pmatrix} \tag{3.20}$$

Using Appendix (1) a necessary and sufficient condition for it to be impossible to solve these equations for any values of $x_o, \varsigma_1 ... \varsigma_n$ is that there exists a vector $\underline{\lambda} = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ independent of x such that

$$\lambda_1 T(x) + \lambda_2 T_x(x) + \lambda_3 T_{xx}(x) + \lambda_4 T_\sigma(x) = 0 \tag{3.21}$$

and

$$-\lambda_3 \ell^2 + \lambda_4 \neq 0 \tag{3.22}$$

Equation (3.22) will be satisfied for all positive ℓ^2 if and only if

$$\lambda_3 \lambda_4 < 0 \tag{3.23}$$

Our boundary condition (2) means that $F(x)$, and hence $T(x)$, cannot depend on any scale length. The $\underline{\lambda}$'s are independent of x and so to make (3.21) dimensionally correct (Bridgman, 1922) we set

$$\lambda_1 = \frac{a}{\sigma^2}, \quad \lambda_2 = \frac{b}{\sigma}, \quad \lambda_3 = c, \quad \lambda_4 = \frac{-d}{\sigma} \tag{3.24}$$

and rewrite it as

$$\frac{aT}{\sigma^2} + \frac{bT_x}{\sigma} + cT_{xx} = \frac{d}{\sigma} T_\sigma \tag{3.25}$$

Condition (3.23) implies that $\frac{d}{c}$ is positive.

Now $T = \frac{d^2 F}{dx^2}$ so F will also satisfy (3.25) although it is possible to add a term ϕ to F where $\frac{d^2\phi}{dx^2} = 0$. However, this term will not satisfy the boundary condition (5) as $x \mapsto \infty$ and so we discard it.

Thus, we have shown that we can always construct a counter example *unless* our filter F obeys to the equation

$$\frac{aF}{\sigma^2} + \frac{b}{\sigma}F_x + cF_{xx} = \frac{d}{\sigma}F_\sigma \tag{3.26}$$

with $\frac{d}{c}$ positive. It is shown in Appendix (2) that the only solution of this equation obeying the boundary conditions is the gaussian. Hence we obtain Theorem (1).

4. The 2-D Case

We now consider the two-dimensional case when the zero crossing operator is the laplacian ∇^2 and the image depends on $\underline{x} = (x, y)$. Again, we consider the filtered image

$$F * I(\underline{x}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\underline{x} - \underline{\varsigma})I(\underline{x})d\underline{\varsigma} \tag{4.1}$$

We set

$$E(x) = \nabla^2\{F * I(\underline{x})\} \tag{4.2}$$

The zero crossings are solutions of $E(\underline{x}) = 0$ and form surfaces in the three-dimensional $(\underline{x}, \sigma)$ space. Our requirements that zero crossings are not created at larger scales is satisfied if the extrema of these zero crossing surfaces are either maxima or saddle points. Minima are forbidden. Thus, if we have a surface $\sigma(x, y)$ and there is a point (x_o, y_o) with

$$\sigma_x(x_o, y_o) = \sigma_y(x_o, y_o) = 0 \tag{4.3}$$

we cannot have $\sigma_{xy} = 0$ and both

$$\sigma_{xx} > 0, \quad \sigma_{yy} > 0. \tag{4.4}$$

Let t be a parameter of a curve of the surface $E(\underline{x}) = 0$. Then,

$$\frac{dE}{dt} = \frac{\partial E}{\partial x}\frac{dx}{dt} + \frac{\partial E}{\partial y}\frac{dy}{dt} + \frac{\partial E}{\partial \sigma}\frac{d\sigma}{dt} \tag{4.5}$$

Since we are on the zero crossing surface, we have $\frac{dE}{dt} = 0$ and setting $t = x$ and then $t = y$, we obtain

$$\sigma_x = \frac{-E_x}{E_\sigma} \tag{4.6}$$

$$\sigma_y = \frac{-E_y}{E_\sigma} \tag{4.7}$$

Suppose we are at an extremum (x_o, y_o). Choose the x and y axes so that they coincide with the directions of principal curvature at (x_o, y_o). Then we calculate

$$\sigma_{xx}(x_o, y_o) = \frac{-E_{xx}(x_o y_o)}{E_\sigma(x_o, y_o)} \tag{4.8}$$

$$\sigma_{yy}(x_o, y_o) = \frac{-E_{yy}(x_o, y_o)}{E_\sigma(x_o, y_o)} \tag{4.9}$$

It should be emphasized that (4.8) and (4.9) are true only at an extremum of $\sigma(x, y)$ and only if the x and y axes are taken along the directions of the lines of curvature (this ensures $\sigma_{xy} = 0$).

As in the 1-D case, it follows that the conditions (3) and (4) will always be satisfied if E obeys the Diffusion Equation. Since if $\sigma_{xx}(x_o, y_o)$ and $\sigma_{yy}(x_o, y_o)$ are both positive, (4.8) and (4.9) imply that $\frac{E_{xx}(x_o y_o)}{E_\sigma(x_o, y_o)}$ and $\frac{E_{yy}(x_o, y_o)}{E_\sigma(x_o, y_o)}$ are both negative. Thus, a gaussian filter will always obey our condition.

We now show that if the filter is not a gaussian, we can construct a counter-example. The argument is a generalization of the proof of Theorem 1. Let

$$I(\underline{\varsigma}) = \sum_{i=1}^{n} A_i \delta(\underline{\varsigma} - \underline{\varsigma}_i) \tag{4.10}$$

Set

$$T(\underline{x}) = \nabla^2 F(\underline{x}) \tag{4.11}$$

We can construct a counter-example if we can solve the matrix equation for any $x_o, \varsigma_1 \ldots, \varsigma_n$ and any positive ℓ_1^2 and ℓ_2^2:

$$\begin{pmatrix} T(x_o - \varsigma_1) & \ldots & T(x_o - \varsigma_n) \\ T_x(x_o - \varsigma_1) & \ldots & T_x(x_o - \varsigma_n) \\ T_y(x_o - \varsigma_1) & \ldots & T_y(x_o - \varsigma_n) \\ T_{xx}(x_o - \varsigma_1) & \ldots & T_{yxx}(x_o - \varsigma_n) \\ T_{yy}(x_o - \varsigma_1) & \ldots & T_{yy}(x_o - \varsigma_n) \\ T_\sigma(x_o - \varsigma_1) & \ldots & T_\sigma(x_o - \varsigma_n) \end{pmatrix} \begin{pmatrix} A_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ A_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -\ell_1^2 \\ -\ell_2^2 \\ 1 \end{pmatrix} \tag{4.12}$$

Using Appendix (1), a necessary and sufficient condition for no solution to exist for all $x_o, \varsigma_1, \ldots, \varsigma_n$ is that we can find $\underline{\lambda} = (\lambda_1, \ldots, \lambda_5)$ such that

$$\lambda_1(x) + \lambda_2 T_x(x) + \lambda_3 T_y(x) + \lambda_4 T_{xx}(x) + \lambda_5 T_{yy} + \lambda_6 T_\sigma(x) = 0 \tag{4.13}$$

and

$$-\ell_1^2\lambda_4 - \ell_2^2\lambda_5 + \lambda_6 \neq 0 \tag{4.14}$$

Equation (4.14) can be satisfied for all positive ℓ_1^2 and ℓ_2^2 if and only if:

$$\lambda_4\lambda_5 > 0, \quad \lambda_4\lambda_6 < 0. \tag{4.15}$$

Again, boundary condition (2) implies the λ's are of form

$$\lambda_1 = \frac{a}{\sigma^2}, \lambda_2 = \frac{b_1}{\sigma}, \lambda_3 = \frac{b_2}{\sigma}, \lambda_4 = c_1, \lambda_5 = c_2, \lambda_6 = \frac{-d}{\sigma} \tag{4.16}$$

and T satisfies

$$\frac{aT}{\sigma^2} + \frac{b_1}{\sigma}T_x + \frac{b_2}{\sigma}T_y + c_1 T_{xx} + c_2 T_{yy} = \frac{d}{\sigma}T_\sigma \tag{4.17}$$

with $c_1c_2 > 0$ and $c_1 d > 0$.

F will satisfy (4.17) up to a term ψ with $\nabla^2\psi = 0$, which we can discard because of boundary condition (5).

It is shown in Appendix (2) that the only solution of (4.17) which obeys our conditions is the product of two one-dimensional gaussians. If we make the additional assumption of symmetry, we obtain a two-dimensional symmetric gaussian. Hence, the gaussian is the only filter which satisfied our condition, and we have proven Theorem 2. There is an additional property of gaussian filters: allowed zero-crossing surfaces in the x, y, σ space cannot have saddle points with positive mean curvature H. The result of this section forbids the existence of upside-down mountains (in the x, y, σ plane) and also of upside-down volcanos. Sections of the zero-crossings surfaces normal to the x, y plane may appear as suggesting that lines of zero-crossings are created. In fact, because of saddle points of the surface, zeros can be traced *continuously* along the zero-crossing surface to smaller and smaller scales.

5. Further results

It is clear that the methods of proof we have developed do not only apply to zero crossings. For example, consider the one-dimensional case and look for solutions of

$$\frac{d}{dx}(F * I) = 0 \tag{5.1}$$

These correspond to maxima and minima of the filtered signal which we call peaks and troughs. If we set $E = \frac{d}{dx}(F * I)$ and duplicate the arguments of section (2), we find that having a gaussian filter is a necessary and sufficient condition for peaks and troughs not to be created.

More generally, if $L(\underline{x})$ is a differential operator in any dimension that commutes with the diffusion equation, then solutions of

$$L(F * I) = const \tag{5.2}$$

will not be created if and only if the filter is gaussian. Zeros of all linear differential operators can be encompassed by Theorem 1.

In particular, in two dimensions, surfaces obeying $\frac{d}{dx}(F * T) = 0$ can only be created by a non-gaussian filter. Thus, ridges and ravines whose creation necessarily involves creation of zeros along some direction, can only be created, as the scale increases, by a non-gaussian filter. The argument, however, does not apply to extremum points (non degenerate critical points, such as peaks and pits, where all derivatives vanish simultaneously).

6. Directional operator

We have considered the two-dimensional case when our operator is the second directional derivative along the direction of the gradient in the filtered image. Let

$$H(\underline{x}) = \int\int F(\underline{x} - \underline{\varsigma}) I(\underline{\varsigma}) d\underline{\varsigma}. \tag{6.1}$$

The directional operator is

$$\frac{d}{dt} = \frac{1}{|\frac{\partial H}{\partial x_i}|} \frac{\partial H}{\partial x_j} \cdot \frac{\partial}{\partial x_j} \tag{6.2}$$

using the summation convention on the j indices. The second directional derivative along the gradient is then

$$\frac{d^2 H}{dt^2} = \frac{H_i H_j H_{ij}}{H_k H_k} \tag{6.3}$$

where $H_i = \frac{\partial H}{\partial x_i}$, $H_{ij} = \frac{\partial^2 H}{\partial x_i \partial x_j}$ and we use the summation convention. We set

$$E(\underline{x}) = H_i(\underline{x}) H_j(\underline{x}) H_{ij}(\underline{x}) \tag{6.4}$$

The zero crossings lie on the surface $\sigma(x, y)$, where $E(\underline{x}) = 0$. Our condition is that if we have a point (x_o, y_o) where

$$\sigma_x(x_o, y_o) = \sigma_y(x_o, y_o) = 0 \tag{6.5}$$

and the x and y axes are along the direction of the lines of curvature of the $\sigma(x, y)$ surface at that point, then it is impossible for both σ_{xx} and σ_{yy} to be negative, i.e.,

$$\sigma_{xx}(x_o, y_o) < 0, \quad \sigma_{yy}(x_o, y_o) < 0 \tag{6.6}$$

We use the Implicit Function Theorem to obtain

$$\sigma_x = \frac{-E_x}{E_\sigma} \tag{6.7}$$

$$\sigma_y = \frac{-E_y}{E_\sigma} \tag{6.8}$$

and we calculate

$$\sigma_{xx}(x_o, y_o) = \frac{-E_{xx}(x_o, y_o)}{E_\sigma(x_o, y_)} \tag{6.9}$$

$$\sigma_{yy}(x_o, y_o) = \frac{-E_{yy}(x_o, y_o)}{E_\sigma(x_o, y_o)} \tag{6.10}$$

Again, note that if E obeys the Diffusion Equation, then the conditions (6.5) and (6.6) cannot be satisfied. However, E is no longer a linear function of the filter, and so we cannot directly obtain a condition the filter must satisfy. Now set

$$I(x) = \sum_{\alpha=1}^{n} A_\alpha \delta(x - \varsigma_\alpha) \tag{6.11}$$

we find

$$H_i H_j H_{ij} = A_\alpha A_\beta A_\gamma F_i(\alpha) F_j(\beta) F_{ij}(\gamma) \tag{6.12}$$

where the summation convention applies to α, β, γ as well as to i, j.
We define

$$\begin{aligned} T(\alpha\beta\gamma) = \frac{1}{6}\{ & F_i(\alpha)F_j(\beta)F_{ij}(\gamma) + F_i(\beta)F_j(\alpha)F_{ij}(\gamma) \\ & + F_i(\alpha)F_j(\gamma)F_{ij}(\beta) + F_i(\beta)F_j(\gamma)F_{ij}(\alpha) \\ & + F_i(\gamma)F_j(\alpha)F_{ij}(\beta) + F_i(\gamma)F_j(\beta)F_{ij}(\alpha)\} \end{aligned} \tag{6.13}$$

and write (6.12) as

$$H_i H_j H_{ij} = T(\alpha\beta\gamma) A_\alpha A_\beta A_\gamma \tag{6.14}$$

We can produce a counter-example if we can satisfy

$$\begin{pmatrix} T(\alpha\beta\gamma) & \cdots \\ T_x(\alpha\beta\gamma) & \cdots \\ T_y(\alpha\beta\gamma) & \cdots \\ T_{xx}(\alpha\beta\gamma) & \cdots \\ T_{xy}(\alpha\beta\gamma) & \cdots \\ T_\sigma(\alpha\beta\gamma) & \cdots \end{pmatrix} \begin{pmatrix} A_\alpha A_\beta A_\gamma \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -\ell_1^2 \\ -\ell_2^2 \\ 1 \end{pmatrix} \tag{6.15}$$

It follows from Appendix (2) that no solution exists if there is a $\underline{\lambda} = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6)$ such that

$$\lambda_1 T(\alpha\beta\gamma) + \lambda_2 T_x(\alpha\beta\gamma) + \lambda_3 T_y(\alpha\beta\gamma) + \lambda_4 T_{xx}(\alpha\beta\gamma) + \lambda_5 T_{yy}(\alpha\beta\gamma) + \lambda_6 T_\sigma(\alpha\beta\gamma) = 0 \tag{6.16}$$

but

$$-\ell_1^2\lambda_4 - \ell_2^2\lambda_5 + \lambda_6 \neq 0. \tag{6.17}$$

As in section (3), we can use dimensional arguments to show this means that $T(\alpha\beta\gamma)$ satisfies the generalized Diffusion Equation.

However, since we require solutions to (6.15) of specific form $A_\alpha A_\beta A_\gamma$ it is possible that there are no solutions of (6.15) even if $T(\alpha\beta\gamma)$ does not obey the generalized Diffusion Equation. To rule this out, we must show that it is possible to find a solution of form $A_\alpha A_\beta A_\gamma$. From Appendix (2) it is possible to get a solution $B_{\alpha\beta\gamma}$ of

$$\begin{pmatrix} T(\alpha\beta\gamma) \\ \cdot \\ \cdot \\ \cdot \\ T_\sigma(\alpha\beta\gamma) \end{pmatrix} \begin{pmatrix} B_{\alpha\beta\gamma} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -\ell_1^2 \\ -\ell_2^2 \\ 1 \end{pmatrix} \tag{6.18}$$

if and only if the vector

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ -\ell_1^2 \\ -\ell_2^2 \\ 1 \end{pmatrix}$$

lies in the spaced spanned by the

$$\begin{pmatrix} T(\alpha\beta\gamma) \\ \cdot \\ \cdot \\ \cdot \\ T_\sigma(\alpha\beta\gamma) \end{pmatrix}$$

as α, β, γ vary. Denote

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ -\ell_1^2 \\ -\ell_2^2 \\ 1 \end{pmatrix}$$

by ℓ^i and

$$\begin{pmatrix} T_{\alpha\beta\gamma} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ T_\sigma(\alpha\beta\gamma) \end{pmatrix}$$

by $T^i_{\alpha\beta\gamma}$, where $i = 1$ to 6.

Each $T^i(\alpha\beta\gamma)$ is symmetric in all indices α, β and γ and so there are $N = \frac{n(n+1)(n+2)}{6}$ such vectors. They have only six components each and so they are not linearly independent. There will be at least $N-6$ linearly independent vectors $\varsigma^{(p)}_{\alpha\beta\gamma}$ such that

$$\sum_{\alpha\beta\gamma} T^{(i)}_{(\alpha\beta\gamma)} \varsigma^{(p)}_{\alpha\beta\gamma} = 0, \quad p = 1 \quad to \quad N-6. \tag{6.19}$$

If $T_{(\alpha\beta\gamma)}$ does not obey the generalized Diffusion Equation there will be at least one solution $B_{\alpha\beta\gamma}$ to (6.18). The general solution is of form

$$B_{\alpha\beta\gamma} + \sum_{p=1}^{N-6} \mu_p \varsigma^p_{\alpha\beta\gamma} \tag{6.20}$$

where μ is arbitrary.

We now ask under what conditions can we find A_α and μ which satisfy

$$B_{\alpha\beta\gamma} + \sum_{p=1}^{N-6} \mu_p \varsigma^p_{\alpha\beta\gamma} = A_\alpha A_\beta A_\gamma \tag{6.21}$$

From the form of (6.15) it is clear that scaling the A's will not affect the counter-example. Hence, satisfying (6.21) is equivalent to finding an A_α such that $A_\alpha A_\beta A_\gamma$ lies in the N-5 dimensional vector space spanned by $B_{\alpha\beta\gamma}, \varsigma^1_{\alpha\beta\gamma}, \ldots, \varsigma^{N-6}_{\alpha\beta\gamma}$. A necessary and sufficient condition is that $A_\alpha A_\beta A_\gamma$ is perpendicular to the five vectors which span the complement of this N-5 dimensional space in the full N dimensional space.

Let the five vectors be $P_{\alpha\beta\gamma}, Q_{\alpha\beta\gamma}, T_{\alpha\beta\gamma}, X_{\alpha\beta\gamma}$ and $Y_{\alpha\beta\gamma}$. It will be possible to solve (6.21) and hence (6.15) if we can satisfy

$$\begin{aligned} P_{\alpha\beta\gamma}A_\alpha A_\beta A_\gamma &= 0 \\ Q_{\alpha\beta\gamma}A_\alpha A_\beta A_\gamma &= 0 \\ T_{\alpha\beta\gamma}A_\alpha A_\beta A_\gamma &= 0. \\ X_{\alpha\beta\gamma}A_\alpha A_\beta A_\gamma &= 0 \\ Y_{\alpha\beta\gamma}A_\alpha A_\beta A_\gamma &= 0. \end{aligned} \tag{6.22}$$

These are a system of five simultaneous cubic equations in n variables. If we take n sufficiently large, it will always be possible to solve them (Yuille, in preparation).

Thus, unless $T(\alpha\beta\gamma)$ obeys the generalized Diffusion Equation, it will always be possible to construct a counter-example.

We now show that no reasonable filter will satisfy these requirements.

First suppose we have a gaussian filter $G(\underline{x}, \sigma)$

$$G(\underline{x}, \sigma) = \frac{1}{\sigma^m} exp\{\frac{-\underline{x}^2}{2\sigma^2}\} \tag{6.23}$$

where m is an arbitrary number.

Then we find

$$G_i(\alpha) = \frac{-(\underline{x} - \varsigma_\alpha)_i}{\sigma^{m+2}} exp\{\frac{-(\underline{x} - \varsigma_\alpha)^2}{2\sigma^2}\} \tag{6.24}$$

$$G_{ij}(\alpha) = \frac{-\delta_{ij}}{\sigma^{m+2}} exp\{-\frac{(\underline{x} - \varsigma_\alpha)^2}{2\sigma^2}\} + \frac{(\underline{x} - \varsigma_\alpha)_i(\underline{x} - \varsigma_\alpha)_j}{\sigma^{m+4}} exp\{-\frac{(\underline{x} - \varsigma_\alpha)^2}{2\sigma^2}\} \tag{6.25}$$

So we obtain

$$T(\alpha\beta\gamma) = 2exp\left\{-\frac{(\underline{x}-\varsigma_\alpha)^2}{2\sigma^2} - \frac{(\underline{x}-\varsigma_\beta)^2}{2\sigma^2} - \frac{(\underline{x}-\varsigma_\gamma)^2}{2\sigma^2}\right\}$$
$$\times \frac{1}{\sigma^{3m+6}}\left\{-(\underline{x}-\varsigma_\alpha)\cdot(\underline{x}-\varsigma_\beta) - (\underline{x}-\varsigma_\beta)\cdot(\underline{x}-\varsigma_\gamma) - (\underline{x}-\varsigma_\gamma)\cdot(\underline{x}-\varsigma_\alpha)\right.$$
$$+ \frac{(\underline{x}-\varsigma_\alpha)^2(\underline{x}-\varsigma_\beta)}{\sigma^2}\cdot(\underline{x}-\varsigma_\gamma) + \frac{(\underline{x}-\varsigma_\gamma)^2(\underline{x}-\varsigma_\alpha)}{\sigma^2}\cdot(\underline{x}-\varsigma_\beta)$$
$$\left. + \frac{(\underline{x}-\varsigma_\beta)^2(\underline{x}-\varsigma_\alpha)}{\sigma^2}\cdot(\underline{x}-\varsigma_\gamma)\right\} \qquad (6.26)$$

As shown in Appendix (2), the general Diffusion Equation can be written

$$\frac{b_1}{\sigma}T_x + \frac{b_2}{\sigma}T_y + c_1T_{xx} + c_2T_{yy} = \frac{d}{\sigma}T_\sigma \qquad (6.27)$$

If we substitute (6.26) into (6.27) we see that $c_1T_{xx} + c_2T_{yy}$ contains a term

$$Z = -2\frac{(c_1+c_2)}{\sigma^{3m+6}}exp\{-\frac{(\underline{x}-\underline{\varsigma}_\alpha)^2}{2\sigma^2} - \frac{(\underline{x}-\underline{\varsigma}_\beta)^2}{2\sigma^2} - \frac{(\underline{x}-\underline{\varsigma}_\gamma)^2}{2\sigma^2}\} \qquad (6.28)$$

All other terms in (6.27) will be of this form multiplied by powers of $(\underline{x}-\underline{\varsigma}_\gamma), (\underline{x}-\underline{\varsigma}_\beta)$ and $(\underline{x}-\underline{\varsigma}_\gamma)$. From (6.17), c_1 and c_2 have the same sign and so it is impossible for Z to be zero and, hence, (6.27) cannot be satisfied if the filter is a gaussian.

Now suppose we have a filter which satisfies this requirement. Set $\underline{\varsigma}_\gamma = \underline{\varsigma}_\alpha + \underline{\varsigma}_\beta$ and integrate $T(\alpha\beta\gamma)$ with respect to $\underline{\varsigma}_\alpha$ and $\underline{\varsigma}_\beta$. We find

$$\int\int F_i(\underline{x}-\underline{\varsigma}_\alpha)F_j(\underline{x}-\underline{\varsigma}_\beta)F_{ij}(\underline{x}+(\underline{\varsigma}_\alpha+\underline{\varsigma}_\beta))d\underline{\varsigma}_\alpha d\underline{\varsigma}_\beta = F_i * F_j * F_{ij}(3\underline{x}) \qquad (6.29)$$

Hence, with $\underline{\varsigma}_\gamma = \underline{\varsigma}_\alpha + \underline{\varsigma}_\beta$, we have

$$\int\int T(\alpha\beta\gamma)d\underline{\varsigma}_\alpha d\underline{\varsigma}_\beta = F_i * F_j * F_{ij}(3\underline{x}) \qquad (6.30)$$

This will still satisfy the generalized Diffusion Equation since $T(\alpha\beta\gamma)$ obeys this equation for all values of $\underline{\varsigma}_\alpha, \underline{\varsigma}_\beta$ and $\underline{\varsigma}_\gamma$. From Appendix (2), the solution to the generalized Diffusion Equation is $P * f(\underline{x})$, where f is an arbitrary function and

$$P(\underline{x},\sigma) = \frac{1}{\sigma^2}exp\{-\frac{(x+b_1\sigma)}{2\sigma^2}\frac{d}{c_1}\}exp\{-\frac{(y+b_2\sigma)^2}{2\sigma^2}\frac{d}{c_2}\} \qquad (6.31)$$

We have

$$F_i * F_j * F_{ij}(3\underline{x}) = P * f(\underline{x}) \tag{6.32}$$

The boundary condition (4) means that $b_1 = b_2 = 0$ and we can scale the x and y axes to make P a gaussian. Thus

$$F_i * F_j * F_{ij}(3\underline{x}) = G * f(\underline{x}) \tag{6.33}$$

We Fourier transform this equation denoting the fourier transform of a function $g(\underline{x})$ by $\mathcal{T}g(\underline{\omega})$

$$\mathcal{T}F_i(\underline{\omega})\mathcal{T}F_j(\underline{\omega})\mathcal{T}F_{ij}(\underline{\omega}) = \mathcal{T}G(3\underline{\omega})\mathcal{T}f(3\underline{\omega}) \tag{6.34}$$

But we have

$$\mathcal{T}F_i(\underline{\omega}) = -i\omega_i \mathcal{T}F(\underline{\omega}) \tag{6.35}$$

and

$$\mathcal{T}G(3\underline{\omega}) = \exp\{\frac{-9\omega^2}{2\sigma^2}\} \tag{6.36}$$

Hence,

$$\omega^4\{\mathcal{T}F(\underline{\omega})\}^3 = exp\{\frac{-9\omega^2}{2\sigma^2}\}\mathcal{T}f(3\underline{\omega}) \tag{6.37}$$

and

$$\mathcal{T}F(\underline{\omega}) = \{\frac{\mathcal{T}f(3\underline{\omega})}{\omega^4}\}^{\frac{1}{3}} exp\{\frac{-3\omega^2}{2\sigma^2}\} \tag{6.38}$$

Thus F is the convolution of a function with a gaussian and obeys the Diffusion Equation. But, as shown in Appendix 2, the only such filter which satisfies the boundary conditions is a gaussian.

So a filter which obeys the conditions (6.16) and (6.17) must be a gaussian, and yet a gaussian cannot satisfy these conditions. Therefore, for this directional operator, it is impossible to satisfy our requirement. Notice that if the gradient direction does not change rapidly the second directional derivative along the gradient can be approximated by the second derivative along the x axis, where the x axis is chosen in the direction of the gradient. The arguments of section 5 then show that no zero-crossings are created if, and only if, the filter is gaussian. If these assumptions are satisfied at one scale, they may break down at larger scales because of the influence of other parts of the image. We therefore expect that at large scales zero-crossings may be created even for gaussian filters, unless the image is very simple (for instance an isolated straight step-edge).

7. Conclusions

The behavior of the zero- (or level-) crossings is more complex in two dimensions than in one dimension. In the 2-D case, two zero crossing contours can merge into one closed contour as the scale increases. The zero-crossing surface has a one-dimensional crossection (for given y, say) that corresponds to an allowed one-dimensional case. In 2-D, however, the "complementary" situation can also occur: a closed zero-crossing contour can split into two as the scale increases, just as the trunk of a tree may split into two branches. This occurs at saddle points of the zero-crossing surface. This case would correspond in 1-D to the "creation" of a zero crossing (imagine a one-dimensional section of the zero-crossing surface) which is forbidden. In 2-D, however, no new zero crossing is created, since the corresponding surface is continuous down to zero scale. We have constructed two-dimensional examples of both these two cases, using the gaussian filters. Both examples would also work for all other filters.

Several other functions have been proposed for filtering images. We expect that they only give a nice scaling behavior for values of σ for which they approximate the solution of the diffusion equation. The DOG (difference of gaussians) *does not* satisfy the diffusion equation, but is a good approximation except when σ is very small. One-dimensional real Gabor functions (the product of a gaussian and a sine or a cosine) approximate the solution of the diffusion equation only for large values of σ. Our conditions are violated even more by the sinc function which only satisfies the diffusion equation at best in a weak asymptotic sense. Figure 3 shows an example of the zero-crossings generated by the gaussian and the sinc filter.

It is interesting that our proof implies that the difference equation is the only linear equation that has, with suitable boundary conditions, a nice scaling behavior of its solutions. This may have some implications in physics.

In summary, we have shown that the gaussian is the only filter that guarantees a nice scaling behavior of the zero- and level-crossings of linear differential operators. Notice that the gaussian need not be symmetric: elongated directional filters, obtained by stretching the axes, also have a nice scaling behaviour. We are presently studying the practical use of the scaling diagrams (in 2-D) for a symbolic representation of images, as suggested by Witkin, and, in particular, for solving the correspondance problem in stereo. In this context, the robustness of the "scaling representation" under small perturbations of the image is clearly critical and has to be carefully studied.

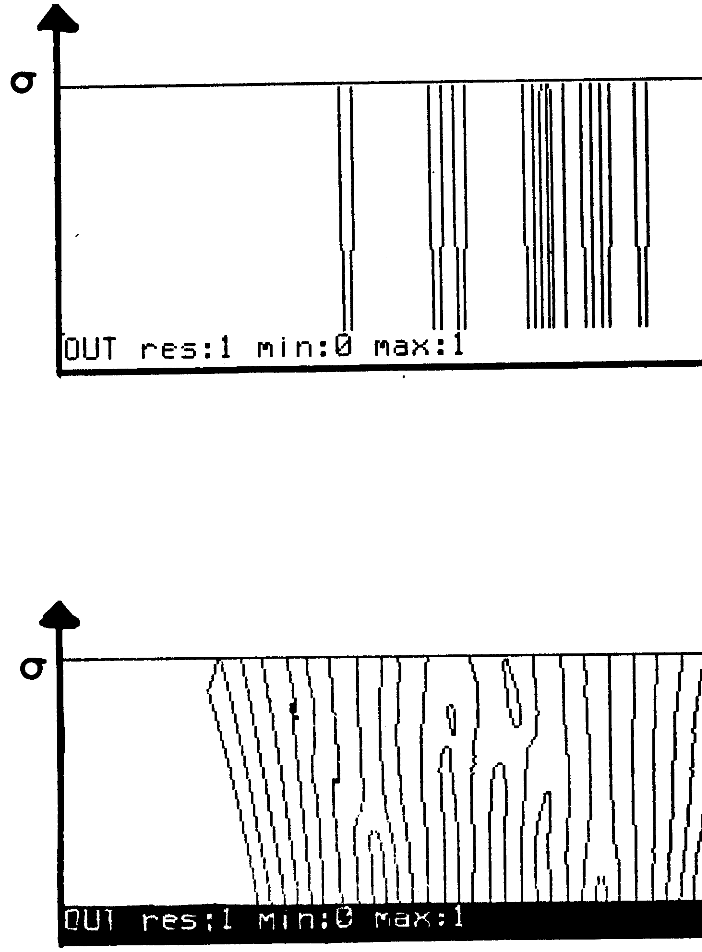

Figure 3 Examples of the zero-crossings of the secon
(a) and of the sinc filter (b) for the same input functi

Acknowledgements We are grateful to C. Koch and M.
and especially to V. Torre and M. Kass for many discu
Thanks also to D. Terzopoulos for useful suggestions
than type the equations.

Appendix 1

If we have a matrix equation

$$\underline{Bx} = \underline{a} \tag{1}$$

the necessary and sufficient condition for a solution is that

$$rank\begin{pmatrix} b_{11} & \dots & b_{1n} \\ . & \dots & . \\ b_{m1} & . & b_{mn} \end{pmatrix} = rank\begin{pmatrix} b_{11} & \dots & b_{1n} & a_1 \\ . & \dots & . & \\ b_{m1} & \dots & b_{mn} & a_m \end{pmatrix} \tag{2}$$

Hence a necessary and sufficient condition for the non-existence of a solution is that we can find a vector $\underline{\lambda} = (\lambda_1, \dots, \lambda_m)$, such that

$$\lambda_1(b_{11}, \dots, b_{1n}) + \dots + \lambda_m(b_{m1}, \dots, b_{mn}) = 0 \tag{3}$$

but that

$$\lambda_1 a_1 + \dots + \lambda_m a_m \neq 0. \tag{4}$$

Appendix 2

Suppose we have a generalized Diffusion Equation of form

$$a\frac{F}{\sigma^2} + \frac{bF_x}{\sigma} + cF_{xx} = \frac{dF_\sigma}{\sigma} \tag{1}$$

We can remove the first term by the scaling $F \mapsto \sigma^{-(a/d)}F$. Consider the remaining terms

$$\frac{bFx}{\sigma} + cF_{xx} = \frac{dF_\sigma}{\sigma} \tag{2}$$

We write

$$F(x,\sigma) = \frac{1}{\sqrt{2\pi}} \int f(\omega,\sigma)e^{-i\omega x}d\omega \tag{3}$$

where $f(\omega,\sigma)$ is the Fourier transform of $F(x,\sigma)$ with respect to x. Combining (3) and (2) we obtain

$$\frac{b(-i\omega)}{\sigma}f + c(-\omega^2)f = \frac{d}{\sigma}\frac{\partial f}{\partial \sigma} \tag{4}$$

We integrate and get

$$f(\omega,\sigma) = g(\omega)\{e^{\frac{-i\omega b\sigma}{d}} e^{\frac{-c\omega^2}{d}\frac{\sigma^2}{2}}\} \tag{5}$$

where $g(\omega)$ is a function of integration independent of σ.
Hence, substituting (5) into (3) gives us

$$F(x,\sigma) = \frac{1}{\sqrt{2\pi}}\int g(\omega)\{e^{\frac{-i\omega b\sigma}{d}} e^{\frac{-c\omega^2}{d}\frac{\sigma^2}{2}}\}e^{-i\omega x}d\omega \tag{6}$$

Note that we are considering equations for which c/d is positive and so the integral is well defined. We now apply the convolution theorem to (6) and get

$$F(x,\sigma) = \frac{1}{\sqrt{2\pi}}\int \lambda(x-\varsigma,\sigma)\mu(\varsigma)d\varsigma \tag{7}$$

where $\mu(\varsigma)$ is the fourier transform of $g(\omega)$ and $\lambda(x,\sigma)$ is the fourier transform of $\{e^{\frac{-i\omega\delta\sigma}{d}} e^{\frac{-i\omega^2\sigma^2}{2d}}\}$. We calculate

$$\lambda(x,\sigma) = \sqrt{\frac{d}{c}\frac{1}{\sigma}}e^{\frac{-d}{2c\sigma^2}(x+b\sigma)^2} \tag{8}$$

Thus the general solution to (1) is of form

$$F(x,\sigma) = \frac{1}{\sqrt{2\pi}}\sigma^{(\frac{a}{d}-1)}\sqrt{\frac{d}{c}}\int e^{\frac{-d}{2c\sigma^2}(x-\varsigma+b\sigma)^2}\mu(\varsigma)d\varsigma \tag{9}$$

We now impose the boundary conditions stated in section (1). First, note that $\lambda(x,\sigma)$ is a gaussian with centre $x=-b\sigma$. The requirement that the centre of the filter does not move implies that $b=0$.
Write

$$F(x,\sigma) = \sigma^{\frac{a}{d}}\int \frac{1}{\sqrt{2\pi}}\sqrt{\frac{d}{c}\frac{1}{\sigma}}e^{\frac{-d}{2c\sigma^2}(x-\varsigma)^2}\mu(\varsigma)d\varsigma \tag{10}$$

and consider the limit as σ tends to 0. Now,

$$Lim_{\sigma\mapsto 0}\frac{1}{\sqrt{2\pi}}\sqrt{\frac{d}{c}\frac{1}{\sigma}}e^{\frac{-d}{2\sigma^2}(x-\varsigma)^2} = \delta(x-\varsigma) \tag{11}$$

where δ denotes the Dirac delta function. If $(\frac{a}{d})$ is non-zero the limits of $F(x,\sigma)$ will either be undefined or zero. Hence our boundary condition (3) forces $a=0$. Morever, substituting into (10) we obtain

$$Lim_{\sigma \mapsto 0} F(x, \sigma) = \mu(x) \tag{12}$$

and condition (3) means that $\mu(x)$ must be the delta function. Hence, on substituting this back into (10) the only solutions of (1) which satisfies our boundary condition is the gaussian

$$G(x, \sigma) = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{d}{c}} \frac{1}{\sigma} e^{\frac{-d}{2c} \frac{x^2}{\sigma^2}} \tag{13}$$

This analysis can be extended to the two dimensional generalized Diffusion Equation

$$\frac{aF}{\sigma^2} + \frac{b_1 F_x}{\sigma} + \frac{b_2 F_y}{\sigma} + c_1 F_{xx} + c_2 F_{yy} = \frac{d}{\sigma} F\sigma \tag{14}$$

A similar argument shows that the only solution obeying the boundary condition in a two-dimensional space is

$$G(x, y, \sigma) = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{d}{c_1}} \sqrt{\frac{d}{c_2}} \frac{1}{\sigma^2} e^{\frac{-d}{2c_1} \frac{x^2}{\sigma^2}} e^{\frac{-d}{2c_2} \frac{y^2}{\sigma^2}} \tag{15}$$

We use the symmetry requirement of section (1) to set $c_1 = c_2$. Then we obtain

$$G(x, y, \sigma) = \frac{1}{2\pi} \frac{d}{c} \frac{1}{\sigma^2} e^{\frac{-d}{2c} \frac{(x^2+y^2)}{\sigma^2}} \tag{16}$$

We can scale the σ-axis by $\sqrt{\frac{c}{d}}$ and write (13) and (16) as

$$G(x, \sigma) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} e^{\frac{-x^2}{2\sigma^2}} \tag{17}$$

and

$$G(x, y, \sigma) = \frac{1}{2\pi} \frac{1}{\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}} \tag{18}$$

respectively. This ensures that σ is the standard deviation of the function.