MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo 833

May 1985

A REGULARIZED SOLUTION TO EDGE DETECTION

T. Poggio, H. Voorhees and A. Yuille

Abstract. We consider edge detection as the problem of measuring and localizing changes of light intensity in the image. As discussed by Torre and Poggio (1984), edge detection, when defined in this way, is an ill-posed problem in the sense of Hadamard.

Using standard regularization theory, we regularize the problem with a stabilizing functional that is a specific form of a Tikhonov stabilizer, following Reinsch (1967) and Schoenberg (1964). The regularized solution that arises is then the solution to a variational principle. In the case of exact data, one of the standard regularization methods (see Poggio and Torre, 1984) leads to cubic spline interpolation before differentiation. We show that in the case of regularly-spaced data this solution corresponds to a convolution filter—to be applied to the signal before differentiation—which is a cubic spline. In the case of non-exact data, which is the most interesting situation, we use another regularization method that leads to a different variational principle. We prove (1) that this variational principle leads to a convolution filter for the problem of one-dimensional edge detection, (2) that the form of this filter is very similar to the gaussian filter, and (3) that the regularizing parameter λ in the variational principle effectively controls the scale of the filter.

Finally, we outline several issues arising from our solution to the edge detection problem: (1) the use of methods from regularizing theories for finding the optimal value of the regularizing parameter λ; (2) the connection between these methods and the scale-space method for edge detection; (3) the relationship between our edge detector and other detectors, especially the Marr/Hildreth edge detector; (4) the extension of our one-dimensional solution to two-dimensional edge detection; and (5) the extension of our method to deal with differentiation of surface data (though the physical constraint underlying the form of the regularizer is not valid in general for depth data); this issue is connected to the problem of interpolating and approximating depth data.

© Massachusetts Institute of Technology, 1985

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the A. I. Laboratory is provided, in part, by the Advanced Research Projects Agency of the Department of Defense. This research has also been partially supported by a gift from the Artificial Intelligence Center of the Hughes Aircraft Company.

1. Introduction

Edge detection does not have a precisely defined goal. The word "edge" itself, which refers to physical properties of objects, is somewhat of a misnomer. Several years of experience have shown that the ideal goal of detecting and locating physical edges in the surfaces being imaged is very difficult and still out of reach (for a review see Brady, 1982). Edge detection has come to be defined as the first step in this goal of detecting physical changes such as object boundaries—the operation of detecting and locating changes in intensity in the image. Other processes which operate on these measurements of intensity changes will then group boundaries and label and characterize them in terms of the properties of the 3-D surfaces.

Intended in this narrow sense, edge detection—this first step in processing the image—is mainly the process that measures, detects and localizes changes of intensity. Derivatives must be estimated correctly to label the critical points in the image intensity array, characterize their local properties (are they minima or maxima or saddle points?) and thus relate them to the underlying physical process (are they shadow edges or depth discontinuities?). As a consequence, several different derivatives of the image, possibly at different scales, may have to be estimated.[1]

In this sense, Torre and Poggio (1984) considered edge detection as a problem of numerical differentiation of images. The problem is not straightforward, and attempts over many years have proven its difficulties. Considered as a problem of numerical differentiation, edge detection turns out to be an ill-posed problem. As explained by Poggio and Torre (1984), mathematically ill-posed problems are problems where the *solution* either *does not exist* or is *not unique* or *does not depend continuously on the data*.

Numerical differentiation is a (mildly) ill-posed problem because its solution does not depend continuously on the data. It is therefore natural to try to solve this problem by using regularization techniques developed in recent years for dealing with mathematically ill-posed problems. The problem can be regularized by the use of a wide class of filters (Torre and Poggio, 1984, section 2.4; see also Duda and Hart, 1973). In the following section we consider two specific regularizing operators, that in some sense are very natural.

2. Regularizing Edge Detection

To regularize an ill-posed problem and make it well-posed, one has to introduce generic constraints on the problem. In this way, one attempts to force the solution to lie in a subspace of the solution space, where it is well defined. The basic idea of regularization techniques is to restrict the space of acceptable solutions by choosing the function that minimizes an appropriate functional. Poggio and Torre consider in particular *standard regularization theory* based on quadratic variational principles. They list three main techniques for regularizing the ill-posed problem of finding z from the data y such that $Az = y$. They involve the choice of norms $\|\cdot\|$ (usually quadratic) and of a *stabilizing functional* $\|Pz\|$. The choice is dictated by mathematical considerations, and, most importantly, by a physical analysis of the generic constraints on the problem. Three main methods can then be applied (see Bertero, 1982):

(1) Among z that satisfy $\|Pz\| \leq C$, where C is a constant, find z that minimizes

$$\|Az - y\|, \tag{1}$$

[1] A very similar problem arises in the characterization of surface properties—in particular their differential properties—from depth data.

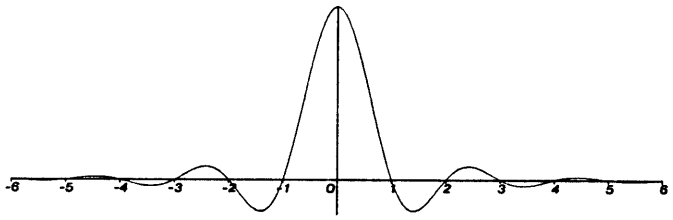


Figure 1 Cubic Interpolating Spline Filter (L_4 function of Schoenberg)

(2) Among z that satisfy $||Az - y|| \leq C$, find z that minimizes

$$||Pz||, \tag{2}$$

(3) Find z that minimizes

$$||Az - y||^2 + \lambda ||Pz||^2, \tag{3}$$

where λ is a regularization parameter.

The first method consists of finding the function z that satisfies the constraint $||Pz|| \leq C$ and best approximates the data. The second method computes the function z that is sufficiently close to the data (C depends on the estimated errors and is zero if the data are noiseless) and is most "regular". In the third method, the regularization parameter λ controls the compromise between the degree of regularization of the solution and its closeness to the data.

In the case of edge detection considered as numerical differentiation, we want an approximation f to the intensity data y_i at sample points x_i that is well behaved under differentiation. Thus we consider an operator A which samples the function f on the lattice such that $Af|_{x_i} = f|_{x_i}$ for $i = 1, ..., N$.

The problem is then to find a suitable norm and a suitable stabilizing functional $||Pf||$. It is natural to chose for P the simplest form of Tikhonov's stabilizing functionals (Tikhonov and Arsenin, 1976) with $P = \frac{d^k}{dx^k}$ and the usual L_2 norm. For $k = 2$ this choice corresponds to a constraint of smoothness on the approximated intensity profile z, with $||Pf|| = \frac{d^2 f}{dx^2}$. Its physical justification is that the noiseless image has to be smooth in the sense that its derivatives must be bounded because the image is band-limited by the optics. Band-limited functions have bounded derivatives because $f' \leq \Omega M$, where $M = sup\, F(\omega)$, Ω is the cut-off frequency, and $F(\omega)$ is the Fourier transform of $f(x)$. Physically, the constraint of smoothness allows us to effectively eliminate the noise that creeps in, after, or during the sampling and transduction process and makes the operation of differentiation unstable. We stress that this is not the only stabilizing functional possible for this problem, although it is probably the simplest one.

2.1. The second regularization method and interpolating cubic splines

With this choice of P, the second regularization method is: among f such that $f(x_i) = y_i$ find f that minimize $\int f''^2\, dx$. A theorem by Schoenberg (see Greville, 1969) shows that the solution to this problem is a cubic spline interpolating between the data points. The following result is a reformulation of results due to Schoenberg (1946, 1964):

Theorem 1: The cubic spline function f interpolating evenly-spaced data points and minimizing $\int f''^2\, dx$ can be obtained by convolving the data points with a cubic spline filter which corresponds to the L_4 function of Schoenberg.

A plot of L_4 is given in Figure 1. Note that the size of the filter is fixed with respect to the sampling lattice. The filter is 0 at every pixel but the central one, where it is 1. Thus L_4 is an interpolating filter that does not perform any smoothing.

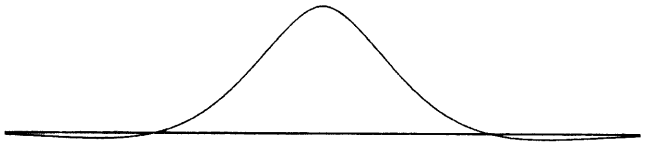


Figure 2 Filter R derived by regularization principles.

2.2. The third regularization method, approximating splines and the edge detection filter

The third regularization method leads to the following problem: Find the f that minimizes

$$\sum_i (y_i - f(x_i))^2 + \lambda \int (f''(x))^2 dx \tag{4}$$

where λ is the regularization parameter which can be found as described later. This problem was considered originally by Reinsch (1967) in the case of numerical differentiation and by Schoenberg (1964) for the problem of graduation. Both Schoenberg and Reinsch gave the solution in terms of approximating cubic splines. In addition, we prove that for most practical purposes, the approximating spline function can be obtained by convolving the data point y_i with the cubic spline *convolution* filter R shown in Figure 2 (see also Appendix 1). We then have the following theorem:

> *Theorem 2: The solution to equation (4)—the regularized solution to the problem of numerical differentiation—in the case of inexact data, can be obtained by convolving the data with a convolution filter which is (a) a cubic spline, and (b) very similar to a gaussian.*

Although this result is especially significant in the context of edge detection where the search for an optimal filter and its justification has been a longstanding preoccupation, it is somewhat surprising that this result does not seem to have been widely appreciated in the numerical analysis literature. The exact assumptions under which Theorem 2 is valid are discussed in Appendix 1. First, the data must be given on a regular grid (as is the case for an image). Second, the image data must either go to zero at infinity or be periodic. Under these conditions, the filtering operation is space-invariant and linear (the Euler-Lagrange equations corresponding to the quadratic variational problem are linear). Thus the approximating spline can be obtained by a convolution operation. Note that the result that the regularizing operator corresponding to a quadratic variational principle is a convolution filter—for data on a regular grid and toroidal boundary conditions—is valid, in general, beyond the case of numerical differentiation.

> *Theorem 3: Quadratic, Tikhonov type regularization principles are equivalent to convolving the data with a generalized spline filter, if the data are given on an regular lattice and the boundary conditions are appropriate.*

A generally interesting question is the physical correctness of the regularized solution. In the case considered in this paper the answer is simple and not very insightful: *a necessary and sufficient condition for the regularized solution to be correct is that the true intensity distribution is a polynomial of order less than 4 between sampling points.* This property can be derived directly from equation (3) of Appendix 1.2.

3. Regularization parameter and comparison with the gaussian filter

Figure 2 shows the filter R obtained by solving the variational principle equation (1) in Appendix 1.1. Its shape and size depends on the regularizing parameter λ. Figure 3(a)

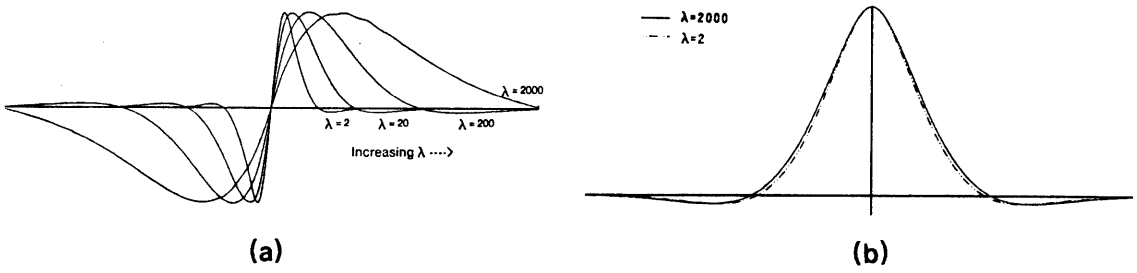


Figure 3 Filter R and first derivative for different values of regularizing parameter λ. **(a)** λ affects size of (the first derivative of) R but **(b)** does not appear to affect shape of R. (Amplitudes are normalized in (a); both amplitudes and widths are scaled linearly and independently in (b) for comparison.)

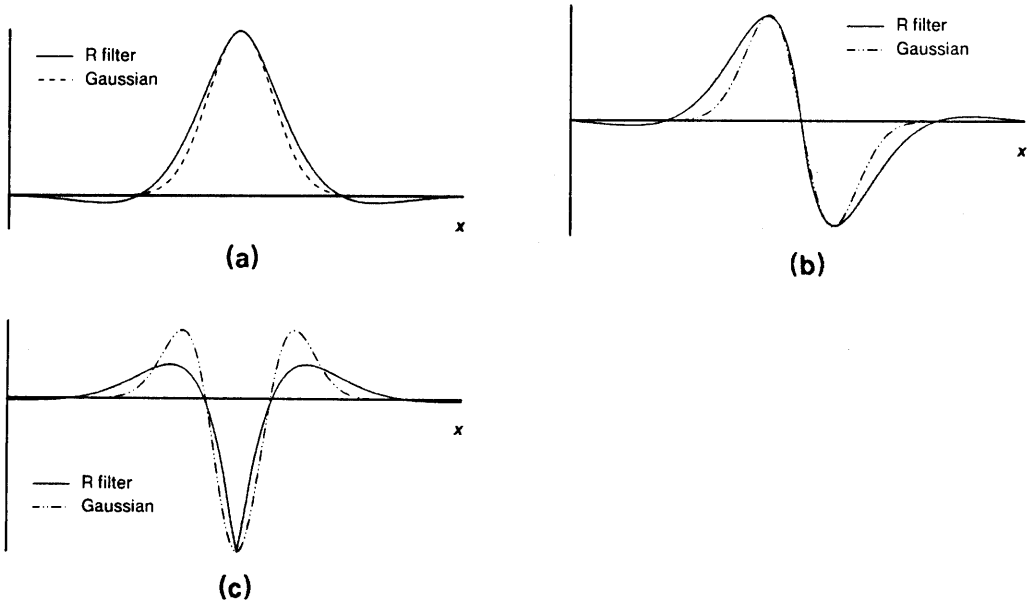


Figure 4 Comparison of one-dimensional regularizing filter R with Gaussian: **(a)** zeroth, **(b)** first, and **(c)** second derivatives.

shows the first derivative of the filter for different values of λ. The continuous version of the filter, derived in Appendix 2, is practically indistinguishable from the discrete filter, as shown by numerical comparisons.

It is rather intuitive that the smoothing parameter λ controls the effective size of the filter. From our numerical work, it seems that λ does not significantly affect the shape of the filter, but only its size, as shown in Figure 3(b). Changing λ amounts to scaling the size of the filter up or down. If λ is small, smoothness is unimportant, and the filter will tend to be an interpolating filter and therefore be similar to a δ function. On the other hand, with a very large λ, the main weight is on smoothness, and the filter will tend to be very large. The continuous form of the filter suggests that the role of λ is indeed equivalent to the role of σ for the gaussian ($\lambda \simeq \sigma^4$, as shown in Appendix 4).

The regularization filter derived here appears to be quite similar to the Gaussian distribution. Graphs of the filter R, its first and second derivatives are shown with those of the Gaussian in Figure 4. Marr and Hildreth (1980; Hildreth, 1980) have argued that the Gaussian is an optimal smoothing filter for detection due to its localization properties in both the spatial and frequency domains. The fact that the Gaussian is quite similar to

the optimal filter derived here using regularization principles provides further mathematical justification for the use of a Gaussian-like filter in edge detection.

If the boundary conditions are not periodic or natural, then the derivation of Reinsch (see Appendix 1.1) provides the correct Green function for those boundary conditions. In numerical experiments we have found that the Green function obtained in this way is space-invariant for all points but those close to the boundaries.

4. Discussion

Several questions and extensions suggest themselves in a natural way. Here we list some of them.

4.1. Finding the Optimal λ

Regularization theory can give the optimal value of λ if the errors on the smoothing criteria and the error of the approximations are known in advance. If the integral of f''^2 is less than E, and the sum of $(y_i - f(x_i))^2$ is less than ϵ, then $\lambda = \epsilon/E$ (see Bertero, 1981; Tikhonov, 1963; Tikhonov and Arsenin, 1977). Normally, however, errors on the data or on the smoothness conditions are not known in advance. Regularization theory provides several methods for finding the optimal smoothing parameter λ under this circumstance. We want to indicate here two main methods: (1) Tikhonov's method, for convolution type problems, as is the case here, and (2) the cross-validation method and the generalized cross-validation method (Wahba, 1980).

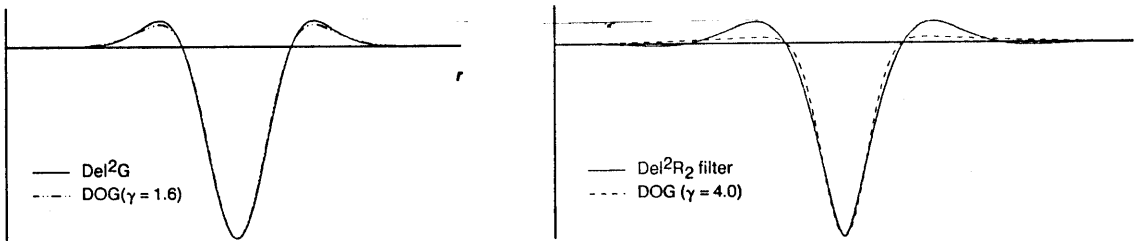
We plan to evaluate these methods for finding the optimal λ for edge detection and to test them on real images. An interesting issue that we are also planning to explore is the following: The basic idea of the generalized cross-validation method is to check the goodness of approximation for each value of λ. In order to do that, one computes the approximation by using not all but only some of the data points. Thus, the data points that are not used for computing the approximations serve as the control points for the goodness of the approximation. If one computes the goodness of the fit at different λ, one can then choose the optimal λ. This idea has obvious connections with the use of fingerprints (Yuille and Poggio, 1983) for finding the natural scale of the filter (Witkin, 1983); this point is discussed next.

4.2. Optimal λ and natural scale

The size of the filter with which to perform edge detection has always been an unresolved issue in computer vision. Our approach makes it clear that one expects, indeed, an optimal size of the filter associated with the optimal value of the smoothing parameter λ. In more recent years, several scales of filtering have been used, partly as a way around this problem. Rosenfeld and Kak (1982), Marr and Poggio (1977), Marr and Hildreth (1980; Hildreth, 1980) have used several sizes of filters in order to perform edge detection.

More recently, Witkin (1983) has suggested the use of scale-space filtering, essentially filtering across a continuum of scales, as the method by which to choose the optimal scale. Witkin suggested some heuristics for picking the natural scale of filtering. We believe that cross-validation-type methods may make more rigorous the idea of selecting an optimal filtering scale by selecting the optimal λ value. We are planning to use fingerprints—the zero-crossings across scales of the Laplacian of the convolved image—to find the optimal λ.

For scale-space filtering to be maximally effective, the shape of the filter should be a gaussian (Yuille and Poggio, 1983a, 1983b; Witkin, 1983). The effect of changing λ is essentially equivalent to changing the size of the filter, and furthermore, the underlying filter is very similar to a gaussian. Therefore, increasing the value of λ is equivalent to filtering



(a) (b)

Figure 5 Difference of two-dimensional Gaussians as approximations to (a) Laplacian of gaussian, (b) Laplacian of two-dimensional regularizing filter R_2. The ratio between the scales (σ) of the two Gaussians is γ.

the signal with a larger and larger gaussian as required in scale-space filtering. Because of the many nice properties of gaussian convolution, this also suggests an efficient way of obtaining approximations at increasing λ.

4.3. Relation with other edge detectors

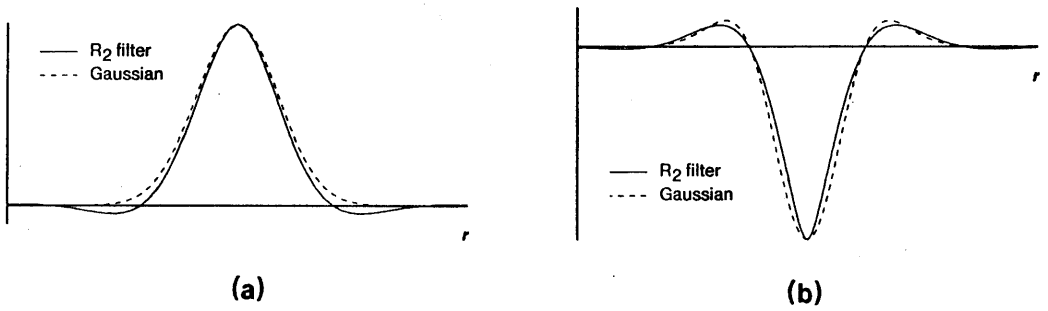
Because of the close similarity of our cubic spline filter to a gaussian, the edge detector that we derive in this paper is very similar to edge detectors proposed previously. Marr and Poggio (1977) proposed the difference of two gaussians as an approximation to the second derivative of a gaussian. Marr and Hildreth (1980; Hildreth, 1980) have shown that the second derivative of a gaussian is, indeed, very close to the difference of gaussians. J. Canny's filter (Canny, 1983) is very close to the derivative of a gaussian, and Haralick's cubic polynomial interpolant (Haralick, 1982) is again similar to Canny's filter.

Our derivation justifies the use of a gaussian or a filter very close to a gaussian as the best filter for edge detection. Regularization theory yields derivative-of-gaussian filters as the optimal filter in a simpler, more general, and, we believe, more rigorous way, than previous derivations. In particular, our result makes clear that the quasi-gaussian filter regularizes the ill-posed problem of numerical differentiation. The regularizing constraint here is that the norm of the derivatives in the noise-free image is small.

It is interesting that we derive a filter very similar to Canny's, based on simpler and more general principles that are not restricted to the optimal detection of *step* edges. It is also interesting to note that the second derivative of the regularization filter, like the second derivative of the Gaussian, can be approximated by a difference of Gaussians (although not as well). While the second derivative of the Gaussian is best approximated by a space constant ratio (the ratio of scales of the two Gaussians) $\gamma \approx 1.6$ (Marr and Hildreth 1980; Hildreth, 1980), increasing the ratio to $\gamma \approx 4$ results in a function which better fits the main (excitatory) lobe of the regularizing filter, as shown in Figure 5.

4.4. Extension to two dimensions

Our approach can be extended to two dimensions in several ways. The most straightforward method involves the use of directional derivatives. First order directional derivatives are taken along several directions. Each one of them is one-dimensional and can be performed according to our one-dimensional edge detector scheme. Depending on the specific goal, one may then choose the direction that gives the maximum value of the derivative. This corresponds to the non-maximum supression scheme used by Canny (1983). It is also equivalent to taking the second directional derivative along the gradient and looking for its zeros (Torre and Poggio, 1984).



(a) (b)

Figure 6 Cross section of **(a)** two-dimensional regularization filter R_2 and **(b)** its Laplacian.

A second method requires directly formulating the regularization principle in two dimensions. Instead of equation (2), one would then have the problem of minimizing

$$\sum_{i,j}(y_{ij} - f(x_{ij}))^2 + \lambda \int\int (P\, f)^2 dx\, dy \tag{5}$$

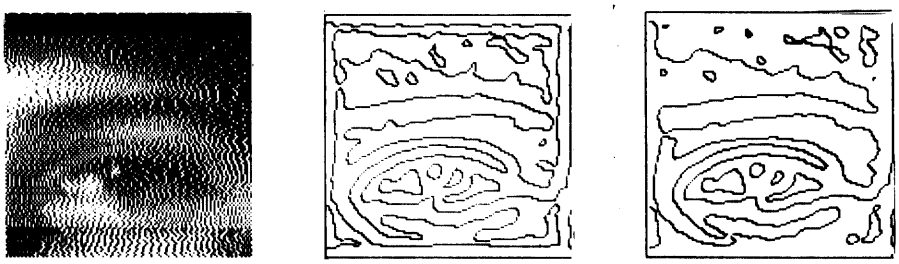
The main problem is the choice of the operator P. If we consider a Tikhonov stabilizer (see Poggio and Torre, 1984), then a choice for P that is smooth enough to allow the use of second derivatives of the regularized image is $P = \nabla^2\nabla$. This choice is used in Appendix 3 to derive the filter for the two-dimensional continuous case. The filter is shown in Figure 6. The choice of the derivative to be used on the filter is a *separate*, important issue that we do *not* address in this paper. Torre and Poggio (1984) discuss the properties of several two-dimensional differential operators, including the second directional derivative along the gradient.

If P is chosen to be the quadratic variation or the square Laplacian, the resulting approximations, known as as thin plate splines (Wahba, 1980; Terzopoulos, 1984a), are not smooth enough for finding zeros of second derivatives of a function[2], as implied by Terzopoulos (1984b). It may also be interesting to explore the filters resulting from a non-linear functional P. In the case of a non-linear P, one cannot use, in general, the standard results about uniqueness and other properties of the solution that are available for the quadratic case of Tikhonov's stabilizers, because the functional is no longer convex.

Clearly, formulations of this type are also relevant for the problem of surface interpolation and approximation in the sense of Grimson (1982) and Terzopoulos (1984a). In the case of sparse data, which they considered, the variational principle does not lead to a convolution filter, although it does lead to a standard Green function. On a regular grid it leads to a convolution filter similar to the gaussian. As a practical implication, evenly-spaced surface data (for example, laser range data) may be interpolated or approximated effectively by gaussian convolution. Hence, tasks which involve differentiating surface data, such as computing lines of curvature (Brady, Ponce, Yuille and Asada, 1984), could use the simpler convolution method to smooth the data. Since Reinsch's method (see Appendix 1.1) can deal with boundary conditions different from periodic ones, the corresponding Green function can be used to prevent smoothing across depth discontinuities.

The results of applying the Laplacian of the two-dimensional regularization filter and the Laplacian of a gaussian to an image are shown in Figure 7. As expected, due to the similarilty of the two filters, both edge detection operators yield similar results.

[2]We are indebted to Demetri Terzopoulos for this remark



(a) (b) (c)

Figure 7 Comparison of Laplacian of two-dimensional regularization filter and Laplacian of Gaussian as Edge Detectors: **(a)** image I, **(b)** zero-crossings of $\nabla^2 R_2 * I$, **(c)** zero-crossings of $\nabla^2 G_2 * I$.

5. Acknowledgements

Several people contributed valuable comments to this paper, especially Eric Grimson, Ellen Hildreth and Demetri Terzopoulos. Also, Carol Bonomo should be acknowledged, if only because she typed the acknowledgement section. H. Voorhees was supported by a Fellowship in Artificial Intelligence from NL Industries.

Appendix 1: The regularization filter for discrete data in one dimension

Below we present two methods for proving that the third regularization method described in the text yields a convolution filter in the case of evenly-spaced discrete data in one dimension with appropriate boundary conditions. Each derivation includes a method for computing the filter, which has the form of a cubic spline.

1.1. The convolution version of Reinsch's optimal filter

Reinsch (1967) considered the general problem of interpolating and smoothing a sequence of (not necessarily evenly-spaced) data points with uncertainties. Given a sequence of points $\langle y_i \rangle$ with uncertainties $\langle \delta y_i \rangle$, representing values of a function at points $\langle x_i \rangle$, where $i = 1, 2, \ldots, n$ and $x_1 < x_2 < \cdots < x_n$, the problem of finding a smooth function which passes near each point (x_i, y_i) can be formulated as finding the function $f(x)$ that minimizes the functional (to be compared with equation 4 in the text)

$$\sum_{i=1}^{n} \left(\frac{f(x_i) - y_i}{\delta y_i} \right)^2 + \lambda \int_{x_1}^{x_n} (f''(x))^2 dx. \tag{1}$$

Using calculus of variations, a cubic spline function is shown to be the optimal function satisfying (1),

$$f(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3, \qquad x_i \leq x < x_{i+1}, \tag{2}$$

for $i = 1, 2, \ldots, n-1$, and formulas are presented for calculating from the data $\langle x_i \rangle$ and $\langle y_i \rangle$ the coefficients $\langle a_i \rangle$, $\langle b_i \rangle$, $\langle c_i \rangle$, and $\langle d_i \rangle$.

We specialize Reinsch's derivation as follows: First, the uncertainty associated with each data point is assumed to be the same so that all $\delta y_i = 1$. Equation (1) above thus reduces to equation (4) in the text. Second, the data are assumed to be uniformly spaced, so that $x_{i+1} - x_i = h$ for $i = 1, 2, \ldots, n-1$. Finally, unlike Reinsch's derivation, we assume periodic boundary conditions, *i.e.*, that the solution f is a periodic function over $\Re$, with $f(x) = f(x \pm kn)$ for integer k.

With these specializations we show that each set of coefficients can be calculated by multipying the data points $\underline{y}$ by a constant coefficient matrix:

$$\underline{a} = A\underline{y}, \quad \underline{b} = B\underline{y}, \quad \underline{c} = C\underline{y}, \quad \underline{d} = D\underline{y},$$

where A, B, C, and D are circulant matrices representing the filter and its derivatives. Hence, the data can be optimally smoothed (and differentiated) by a convolution operation.

Using calculus of variations, one obtains from (1) these conditions for the optimal function $f(x)$:

$$f(x_i)_+ - f(x_i)_- = 0 \tag{3a}$$

$$f'(x_i)_+ - f'(x_i)_- = 0 \tag{3b}$$

$$f''(x_i)_+ - f''(x_i)_- = 0 \tag{3c}$$

$$f'''(x_i)_+ - f'''(x_i)_- = -\tfrac{1}{\lambda}(f(x_i) - y_i), \tag{3d}$$

(using $\delta y_i = 1$) and

$$f''''(x) = 0, \qquad x_i \leq x < x_{i+1}, \tag{4}$$

for $i = 1, \ldots, n$ (with $x_1 = x_{n+1}$), where

$$f^{(k)}(x_i)_\pm \doteq \lim_{\epsilon \to 0} f^{(k)}(x_i \pm \epsilon).$$

Equations (3–4) show that $f(x)$ is a cubic spline, having the form of (2); hence,

$$f(x_i)_+ = a_i \tag{5a}$$
$$f'(x_i)_+ = b_i \tag{5b}$$
$$f''(x_i)_+ = 2c_i \tag{5c}$$
$$f'''(x_i)_+ = 6d_i \tag{5d}$$

and, since all $x_i - x_{i-1} = h$,

$$f(x_i)_- = a_{i-1} + b_{i-1}h + c_{i-1}h^2 + d_{i-1}h^3 \tag{6a}$$
$$f'(x_i)_- = b_{i-1} + 2c_{i-1}h + 3d_{i-1}h^2 \tag{6b}$$
$$f''(x_i)_- = 2c_{i-1} + 6d_{i-1}h \tag{6c}$$
$$f'''(x_i)_- = 6d_{i-1}. \tag{6d}$$

Now substitution into (3) yields

$$a_i - a_{i-1} - b_{i-1}h - c_{i-1}h^2 - d_{i-1}h^3 = 0 \tag{7a}$$
$$b_i - b_{i-1} - 2c_{i-1}h - 3d_{i-1}h^2 = 0 \tag{7b}$$
$$2c_i - 2c_{i-1} - 6d_{i-1}h = 0 \tag{7c}$$
$$6d_i - 6d_{i-1} = -\tfrac{1}{\lambda}(f(x_i) - y_i). \tag{7d}$$

Equations (7) can be manipulated to yield

$$b_i h = (a_{i+1} - a_i) - c_i h^2 - d_i h^3 \tag{8a}$$
$$a_{i+1} - 2a_i + a_{i-1} = \tfrac{1}{3}(c_{i-1} + 4c_i + c_{i+1})h^2 \tag{8b}$$
$$d_i h^3 = \tfrac{1}{3}(c_{i+1} - c_i)h^2 \tag{8c}$$
$$(c_{i-1} - 2c_i + c_{i+1})h^2 = \tfrac{1}{2\lambda}(y_i - a_i)h^3 \tag{8d}$$

Using the notation that:

$I =$ the $n \times n$ identity matrix $[i_{j,k}]$ where

$$i_{j,k} = \begin{cases} 1, & \text{if } j = k,\ j = 1, \ldots, n \\ 0, & \text{otherwise} \end{cases}$$

$N =$ the $n \times n$ "next" matrix $[n_{j,k}]$ where

$$n_{j,k} = \begin{cases} 1, & \text{if } j = k - 1 \bmod n,\ j = 1, \ldots, n \\ 0, & \text{otherwise} \end{cases}$$

$P =$ the $n \times n$ "previous" matrix $= N^T$,

we define the matrices

$$T = \tfrac{1}{3}P + \tfrac{4}{3}I + \tfrac{1}{3}N$$
$$Q = P - 2I + N.$$

For example, if $n = 4$,

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad N = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$T = \begin{pmatrix} 4/3 & 1/3 & 0 & 1/3 \\ 1/3 & 4/3 & 1/3 & 0 \\ 0 & 1/3 & 4/3 & 1/3 \\ 1/3 & 0 & 1/3 & 4/3 \end{pmatrix} \qquad Q = \begin{pmatrix} -2 & 1 & 0 & 1 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 1 & 0 & 1 & -2 \end{pmatrix}.$$

Defining the vectors $\underline{a} = (a_i)^T$, $\underline{b} = (b_i h)^T$, $\underline{c} = (c_i h^2)^T$, $\underline{d} = (d_i h^3)^T$, and $\underline{y} = (y_i)^T$, for $i = 1, \ldots, n$, equations (8) can be expressed as

$$\underline{b} = (N - I)\underline{a} - \underline{c} - \underline{d} \tag{9a}$$

$$Q\underline{a} = T\underline{c} \tag{9b}$$

$$\underline{d} = \tfrac{1}{3}(N - I)\underline{c} \tag{9c}$$

$$Q\underline{c} = \tfrac{h^3}{2\lambda}(\underline{y} - \underline{a}). \tag{9d}$$

These can be simplified to

$$\underline{a} = A\underline{y}, \quad \underline{b} = B\underline{y}, \quad \underline{c} = C\underline{y}, \quad \underline{d} = D\underline{y}, \tag{10}$$

where

$$C = \tfrac{h^3}{2\lambda}(Q^2 + \tfrac{h^3}{2\lambda}T)^{-1}Q \tag{11a}$$

$$A = I - \tfrac{2\lambda}{h^3}QC \tag{11b}$$

$$D = \tfrac{1}{3}(N - I)C \tag{11c}$$

$$B = (N - I)A - C - D. \tag{11d}$$

Since I, N, and P are circulant matrices, and since the set of circulant matrices is closed under matrix addition, multiplication, and inversion, the resulting coefficient matrices A, B, C, and D are also circulant, representing filters $R(x)$, $R'(x)$, $R''(x)$, and $R'''(x)$. The fact that the filters are derivatives of each other follow from (5); hence, since differentiation is a linear operator, the filter R'', for example, can be used to both optimally smooth and twice differentiate the data.

1.2 Another derivation of the regularization filter for discrete data in one dimension

In this derivation we show that the third regularization method, described by equation (4) of the text, yields a convolution filter which is a cubic spline. Standard results from the calculus of variations guarantee that our solution has continuous second derivatives.

Again, the problem is to minimize

$$\lambda \int (f''(x))^2 dx + \sum_i (f(x_i) - y_i)^2. \tag{1}$$

We find the minimum by sending $f(x) \mapsto f(x) + \delta f(x)$ and setting the first variation of (1) to zero,

$$\lambda \int f''''(x)\, \delta f(x) dx + \sum_i (f(x_i) - y_i)\, \delta f(x_i) = 0. \tag{2}$$

This yields the Euler-Lagrange equation

$$\lambda f''''(x) + f(x) \sum_i \delta(x - x_i) = \sum_i y_i \delta(x - x_i). \tag{3}$$

So far, we have deliberately not specified boundary conditions. For infinite, or toroidal, boundary conditions, the function $f(x)$ can be determined in terms of the $\langle y_i \rangle$ by a convolution filter if and only if the data points $\langle x_i \rangle$ are evenly spaced. This is because the system is then translation-invariant[2]. We will show this explicitly and give a method for constructing the convolution filter.

The function in (1) is convex, and hence has a unique minimum, so there is a unique solution for $f(x)$ in (3). Thus, we only need to see whether a convolution can solve (3). We

[2]Boundary conditions other than infinite or toroidal will destroy translation invariance.

try a solution

$$f(x) = R(x) * y(x) = \int R(x-\varsigma)y(\varsigma)d\varsigma \tag{4}$$

where "$*$" denotes convolution, $R(x)$ is a filter, and $y(x) = \sum_i y_i \delta(x - x_i)$. We substitute (4) into (3) and obtain

$$\lambda \sum_i y_i R''''(x - x_i) + \sum_i \sum_j y_j R(x_i - x_j)\delta(x - x_i) - \sum_i y_i \delta(x - x_i) = 0 \tag{5}$$

We compare coefficients of y_i and obtain

$$\lambda R''''(x - x_i) + \sum_j R(x_j - x_i)\delta(x - x_j) - \delta(x - x_i) = 0 \tag{6}$$

If the $\langle x_i \rangle$ are not evenly-spaced, then these equations are inconsistent, and no convolution filter exists. If the $\langle x_i \rangle$ are evenly spaced, then the set of equations in (6) reduces to a single equation

$$\lambda R''''(x) + \sum_j R(x_j)\delta(x - x_j) - \delta(x) = 0 \tag{7}$$

The solutions to (7) correspond to cubic splines "stitched" together at the points $\{x_i\}$. Let $R_i(x)$ denote the solution in the range $x_i < x < x_{i+1}$. We write

$$R_i(x) = \alpha_i x^3 + \beta_i x^2 + \gamma_i x + \delta_i \tag{8}$$

The splines are stitched together so that $R(x)$, $R'(x)$, and $R''(x)$ are continuous at the points $\{x_i\}$. From (7), we see that R''' has a discontinuity of $-\frac{1}{\lambda}R(x_i)$ at x_i. It is straightforward to find the relations between the $R_i(x)$ and $R_{i-1}(x)$ in terms of the parameters $\alpha_i, \beta_i, \gamma_i$ and δ_i. This gives

$$\begin{aligned}
\alpha_{n+1} &= \alpha_n - \frac{1}{6\lambda}R(nh) \\
\beta_{n+1} &= \beta_n - \frac{nh}{2\lambda}R(nh) \\
\gamma_{n+1} &= \gamma_n - \frac{(nh)^2}{2\lambda}R(nh) \\
\delta_{n+1} &= \delta_n + \frac{(nh)^3}{6\lambda}R(nh)
\end{aligned} \tag{9}$$

where h is the spacing between the lattice points, $h = x_{i+1} - x_i$, and

$$R(nh) = \alpha_n(nh)^3 + \beta_n(nh)^2 + \gamma_n(nh) + \delta_n. \tag{10}$$

Appendix 2: The regularization filter for continuous data in one dimension

We want to find the function $f(x)$ which minimizes the quadratic functional

$$\int (f(x) - y(x))^2 dx + \lambda \int (f''(x))^2 dx. \tag{1}$$

where $y(x)$ are the data, given on the continuum. We will show that $f(x)$ can be expressed as a convolution of the data $y(x)$ with a suitable filter $R(x)$.

We obtain the Euler-Lagrange equations for (1) by setting the first variation to zero as $f(x) \mapsto f(x) + \delta f(x)$. This gives

$$\int \delta f(x)(f(x) - y(x))dx + \lambda \int \delta f(x) f''''(x)dx = 0 \tag{2}$$

for all variations $\delta f(x)$ and hence we have

$$\lambda f''''(x) + f(x) = y(x). \tag{3}$$

The solution to the linear equation (3) is given by

$$f(x) = \int r(x, x')\, y(x')dx \tag{4}$$

where $r(x, x')$ is the Green function of (3) obeying

$$\lambda \frac{d^4}{dx^4} r(x, x') + r(x, x') = \delta(x - x'). \tag{5}$$

Now (5) is a translation invariant equation and our boundary conditions are periodic or at infinity, so $r(x, x')$ is a function of $(x - x')$ only. We write

$$r(x, x') = r(x - x'). \tag{6}$$

For each different value of λ, we have a different equation (5) and hence a different Green's function which we denote by $R(x, \lambda)$. Note from (5) that

$$R(x, 0) = \delta(x). \tag{7}$$

So in the limit as the confidence in the data is complete, the filter becomes the delta function. (It is easy to see all this by taking the Fourier transform of equation 3.)

It is straightforward to find the Green's function of (5) which vanishes at plus and minus infinity. This is given by

$$R(x, \lambda) = \frac{1}{2\lambda^{1/4}} e^{-|x|/\sqrt{2}\lambda^{1/4}} \cos\left(\frac{|x|}{\sqrt{2}\lambda^{1/4}} - \frac{\pi}{4}\right). \tag{8}$$

In Fourier space, the transform of $R(x, \lambda)$ can be obtained directly from (5):

$$\mathcal{F}R(\omega\lambda) = \frac{1}{1 + \lambda\omega^4}, \tag{9}$$

so we can write

$$R(x, \lambda) = \frac{1}{2\pi} \int \frac{1}{1 + \lambda\omega^4} e^{-i\omega x} d\omega. \tag{10}$$

At $\lambda = 0$ the Green function goes to a delta function, as we saw in (7).

Define μ by

$$\lambda\mu^4 = 1. \tag{11}$$

Then, for $x > 0$, we have

$$R(x, \mu) = \frac{\mu}{2\sqrt{2}} e^{-x\mu/\sqrt{2}} \left(\cos\frac{x\mu}{\sqrt{2}} + \sin\frac{x\mu}{\sqrt{2}}\right) \tag{12}$$

and

$$R'(x,\mu) = -\frac{\mu^2}{2} e^{-x\mu/\sqrt{2}} \sin \frac{x\mu}{\sqrt{2}}. \tag{13}$$

Thus the extrema of $R(x,\mu)$ occur at

$$x = \frac{\sqrt{2}}{\mu} n\pi, \quad n = 1, 2 \ldots \tag{14}$$

and at these extrema, R takes the value

$$R\left(\frac{\sqrt{2}}{\mu} n\pi, \mu\right) = \frac{\mu}{2\sqrt{2}} e^{-n\pi} (-1)^n. \tag{15}$$

So if μ is small (λ large), the extrema are at large x and correspond to small values of R. If μ is large (λ small), the nearest extrema occur at small x and correspond to large values of R. The function changes sign many times.

Appendix 3: The regularization filter for continuous data in two dimensions

We can extend the results of the previous section to two dimensions. Our generalization of the smoothing function $\int (f''(x))^2 dx$ is

$$\int\int \nabla^2 \underline{\nabla} f(\underline{x}) \nabla^2 \underline{\nabla} f(\underline{x}) d\underline{x}, \tag{1}$$

Courant and Hilbert (1953) show that the Euler-Lagrange equation is

$$\nabla^2 \nabla^2 \nabla^2 f = 0. \tag{2}$$

We write the regularization of the two-dimensional smoothing problem in the form

$$\int\int (f(\underline{x}) - y(\underline{x}))^2 d\underline{x} + \lambda \int\int (\nabla^2 \underline{\nabla} f(\underline{x}) \nabla^2 \underline{\nabla} f(\underline{x})) d\underline{x}. \tag{3}$$

The Euler-Lagrange equations of the combined system are

$$\lambda \nabla^2 \nabla^2 \nabla^2 f(\underline{x}) + f(\underline{x}) = y(\underline{x}). \tag{4}$$

This equation is translation invariant and so, for boundary conditions at infinity or periodic, the solution can be written as a convolution

$$f(\underline{x}) = R_2(\underline{x}) * y(\underline{x}) \tag{5}$$

where

$$\lambda \nabla^2 \nabla^2 \nabla^2 R_2(\underline{x}) + R_2(\underline{x}) = \delta(\underline{x}) \tag{6}$$

where $\delta(\underline{x})$ is the Dirac delta function. Again, observe that for small λ the filter $R_2(\underline{x})$ tends to the delta function.

To solve (6) we take its Fourier Transform, and find

$$\mathcal{F} R_2(\underline{\omega}) = \frac{1}{\lambda \underline{\omega}^6 + 1}. \tag{7}$$

This gives a solution

$$R_2(\underline{x}) = \frac{1}{2\pi} \int \frac{e^{+i\underline{\omega}\cdot\underline{x}}}{\lambda \underline{\omega}^6 + 1} d\underline{\omega}. \tag{8}$$

We express $\underline{x}$ and $\underline{\omega}$ in polar coordinates

$$\underline{x} = (r, \phi), \quad \underline{\omega} = (w, \theta). \tag{9}$$

So

$$R_2(\underline{x}) = \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} \frac{e^{+iw\cos(\theta-\phi)}}{\lambda w^6 + 1} \; w \, dw \, d\theta. \tag{10}$$

Now

$$\int_0^{2\pi} e^{iwr\cos(\theta-\phi)} d\theta = \pi J_o(wr) \tag{11}$$

where J_o is the zero order Bessel function. This gives

$$R_2(\underline{x}) = \frac{1}{2} \int_0^\infty \frac{J_o(wr)}{\lambda w^6 + 1} w \, dw, \tag{12}$$

that is, R_2 is the Hankel transform of $1/(\lambda w^6 + 1)$. This integral can be solved numerically.

The square Laplacian stabilizer leads to a similar formula, i.e. to

$$R_2(\underline{x}) = \frac{1}{2} \int_0^\infty \frac{J_o(wr)}{\lambda w^4 + 1} w \, dw, \tag{13}$$

The corresponding filter is not smooth enough to allow second derivatives to be taken, but is sufficient for first derivatives. The integral is found in standard texts (for example,

Gradshteyn and Ryzhik, 1980), yielding

$$R_2(x) = \frac{1}{2\lambda}\left(-\lambda^{1/2}\right) kei\left(\frac{r}{\lambda^{1/4}}\right) \tag{14}$$

where $kei(x)$ is a Thompson function given by

$$kei(x) = \left(ln\tfrac{2}{x} - C\right) bei(x) - \tfrac{\pi}{4} ber(x) + \sum_{k=0}^{\infty} (-1)^k \frac{x^{4k+2}}{2^{4k+2}[(2k+1)!]^2} \sum_{m=1}^{2k+1} \frac{1}{m} \tag{15}$$

where

$$\begin{aligned} bei(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{4k+2}}{2^{4k+2}[(2k+1)!]^2}, \\ ber(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{4k}}{2^{4k}[(2k)!]^2}, \end{aligned} \tag{16}$$

and $C = 0.5772...$ is Euler's constant. The asymptotic behavior is

$$kei(x) \mapsto \sqrt{\frac{\pi}{2x}} e^{-\alpha x} \sin(-\beta x) \tag{17}$$

where α and β are constants.

Appendix 4: The regularization filter as an approximation to the Gaussian

In this appendix we show that the regularization filter approximates the gaussian, both in one and two dimensions.

4.1 Comparison of one-dimensional filters

We show that the one-dimensional regularization filter is an approximate solution to the diffusion equation and therefore approximates a gaussian.

The one-dimensional regularization filter in is given by

$$R(x,\mu) = \frac{\mu}{2} e^{-\mu x/\sqrt{2}} \cos\left(\frac{\mu x}{\sqrt{2}} - \frac{\pi}{4}\right). \tag{1}$$

We expand this in a Taylor series in $\frac{\mu x}{\sqrt{2}}$ to get

$$R(x,\mu) = \frac{\mu}{2\sqrt{2}}\left[1 - \left(\frac{\mu x}{\sqrt{2}}\right)^2 + 0(\mu x)^3\right]. \tag{2}$$

This expansion is valid when μx is small, *i.e.*, when x is small compared to $\lambda^{1/4}$. In this case the first two terms which we denote by $\tilde{R}(x,\mu)$ are a good approximation to the function. We calculate

$$\frac{\partial^2 \tilde{R}}{\partial x^2} = \frac{-\mu^3}{2\sqrt{2}} \tag{3}$$

and

$$\frac{\partial \tilde{R}}{\partial \mu} = \frac{1}{2\sqrt{2}} + O(\mu x)^2, \tag{4}$$

which satisfy, to order $(\mu x)^2$, the equation

$$\frac{\partial^2 \tilde{R}}{\partial x^2} = -\mu^3 \frac{\partial \tilde{R}}{\partial \mu}. \tag{5}$$

Thus this function obeys the diffusion equation,

$$\frac{\partial^2 \tilde{R}}{\partial x^2} = \frac{\partial \tilde{R}}{\partial t} \tag{6}$$

with parameter $t = \frac{1}{2}\mu^{-2} = \frac{1}{2}\lambda^{1/2}$, in the region where μx is small. As μ decreases, this region gets larger and the region in which the function approximates a Gaussian increases.

This theoretical analysis supports the numerical results (for discrete data) which show that R can be approximated by a Gaussian. Furthermore, recalling that the standard deviation σ of the Gaussian is given by $\sigma = \sqrt{2t}$, the analysis shows that the standard deviation of the corresponding Gaussian is $\lambda^{1/4}$.

A comparison of R with the gaussian $G = (e^{-x^2/2\sigma^2})$ can also be done directly in the Fourier domain, where $\mathcal{F}R = 1/(1+\lambda\omega^4)$ and $\mathcal{F}G = e^{-\sigma^2\omega^2}$, as shown in Figure 8.

4.2 Comparison of two-dimensional filters

We now consider the two-dimensional case. The regularized filter can be written in terms of a Fourier integral

$$R_2(\underline{x},\lambda) = \frac{1}{2\pi}\int \frac{e^{-i\underline{\omega}\cdot\underline{x}}}{1+\lambda\underline{\omega}^6} d\underline{\omega}. \tag{7}$$

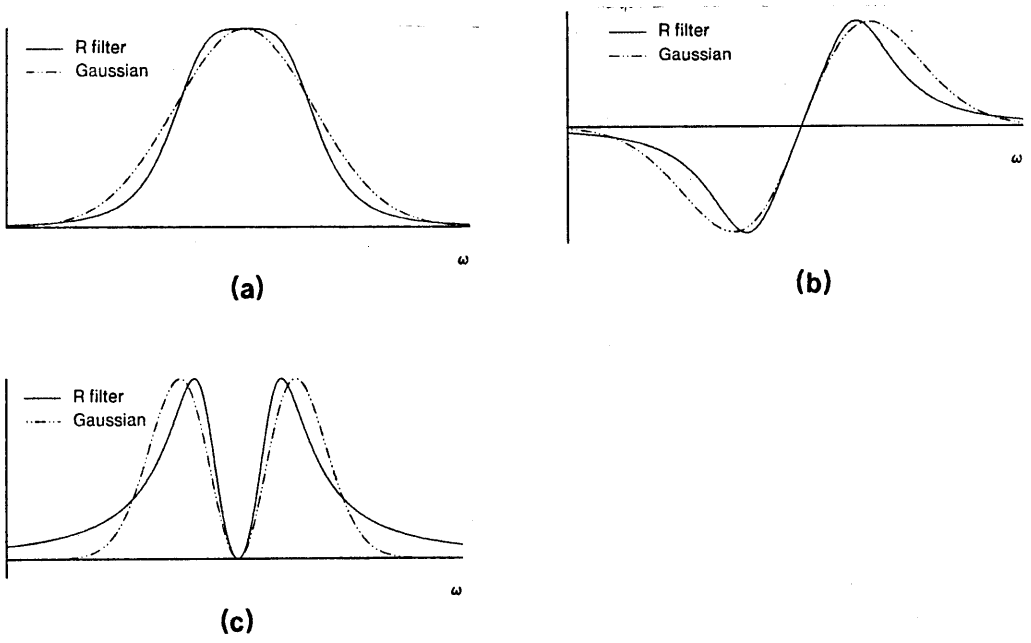


Figure 8 Comparison of one-dimensional regularizing filter R with gaussian in Fourier domain: **(a)** zeroth, **(b)** first, and **(c)** second derivatives.

We perform the transformation $\underline{\omega} \mapsto \lambda^{1/6}\underline{\omega}$ to obtain, with $\underline{\mu} = \frac{1}{\lambda^{1/6}}$,

$$R_2(\underline{x}, \underline{\mu}) = \frac{\underline{\mu}}{2\pi} \int \frac{e^{-i\underline{\omega}\cdot\underline{\mu}\underline{x}}}{1+\underline{\omega}^6} d\underline{\omega}. \tag{8}$$

We expand the exponential in a power series

$$R_2(\underline{x}, \underline{\mu}) = \frac{\underline{\mu}}{2\pi} \int \frac{1}{1+\underline{\omega}^6} (1 + i\underline{\omega}\cdot\underline{\mu}\underline{x} + \tfrac{1}{2}(i\underline{\omega}\cdot\underline{\mu}\underline{x})^2 + 0(\underline{\mu}^3\underline{x}^3)) d\underline{\omega}. \tag{9}$$

Thus we have

$$R_2(\underline{x}, \underline{\mu}) = \frac{\underline{\mu}}{2\pi} \int \frac{1}{1+\underline{\omega}^6} (1 + \tfrac{1}{2}(i\underline{\omega}\cdot\underline{\mu}\underline{x})^2 + 0(\underline{\mu}^3\underline{x}^3)) d\underline{\omega}. \tag{10}$$

Note that the linear term drops out due to asymmetry of the integrand. Keeping the first two terms on the right hand side of (10), and denoting this approximation by $\tilde{R}_2(\underline{x}, \underline{\mu})$, we calculate

$$\nabla^2 \tilde{R} = \frac{-\underline{\mu}^3}{\pi} \int \frac{\underline{\omega}^2 d\underline{\omega}}{1+\underline{\omega}^6} \tag{11}$$

and

$$\frac{\partial \tilde{R}}{\partial \underline{\mu}} = \frac{1}{2\pi} \int \frac{1}{1+\underline{\omega}^6} d\underline{\omega}. \tag{12}$$

Thus as before, the approximation $\tilde{R}_2(\underline{x}, \lambda)$ satisfies the diffusion equation with t proportional to $\lambda^{1/2}$. The exact function of proportionality can be calculated from (12).

Again, a direct comparison of the Gaussian with our regularizing filter R_2 is done easily in the Fourier plane. Both filters are circularly symmetric and therefore depend only on the radial frequency w. A comparison in Fourier space of the two-dimensional gaussian $\mathcal{F}G_2 = e^{-w^2\sigma^2}$ and the regularizing filter $\mathcal{F}R_2 = 1/(\lambda w^6 + 1)$ is shown in Figure 9.

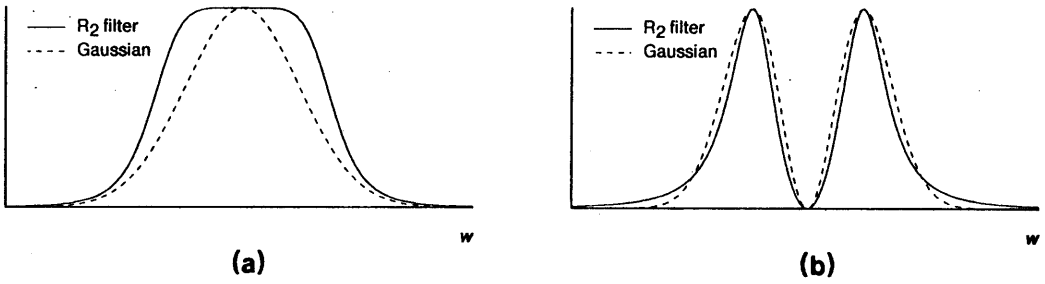


Figure 9 Comparison of two-dimensional filter R_2 with two-dimensional gaussian G_2 in Fourier plane **(a)** filters and **(b)** their Laplacians.

6. References

Bertero, M. "Problemi lineari non ben posti e metodi di regolarizzazione," *Problem non ben posti ed inversi*, Istituto di Analisi Globale, Firenze, 1982.

Brady, J. M. "Computational approaches to image understanding," *Computing Surveys*, **14**, 3–71, 1982.

Brady, J. M., Ponce, J., Yuille, A. L. and Asada, H. "Describing surfaces," *Proc. 2nd International Symposium, Robotics Research*, Harafusa, H. and Inoue, H. (eds), MIT Press, 1985.

Canny, J. F. "Finding edges and lines in images," *AI-TR-720*, MIT AI Lab, June 1983.

Courant, R. and Hilbert, D. *Methods of Mathematical Physics*, vol. 1., Interscience Publishers, Inc., New York, 1953.

Duda, Richard O., and Hart, Peter E. *Pattern Classification and Scene Analysis*, John Wiley and Sons, New York, 1973.

Gradshteyn, I.S. and Ryzhik, I.M. *Table of Integrals, Series and Products*, Academic Press, 1980.

Greville, T. N. E. (ed). "Introduction to spline functions," *Theory and Applications of Spline Functions*, Academic Press, New York, 1–36, 1969.

Grimson, W. E. L. "A computational theory of visual surface interpolation," *Phil. Trans. R. Soc. Lond., B*, **298**, 395–427, 1982.

Grimson, W. E. L. *From Images to Surfaces*, MIT Press, 1981.

Grimson, W. E. L. and Hildreth, E. "Comments on 'Digital step edges from zero crossings of second directional derivatives'," *IEEE Trans. PAMI*, **7**, 121–127, 1985.

Haralick, R. M. "Edge and region analysis for digital image data," *Comp. Graphics and Image Proc.*, **12**, 60–73, 1980.

Haralick, R. M. "The digital edge," *Proc. 1981 Conf. on Pattern Recognition and Image Processing*, Dallas, Texas, 285–294, 1981.

Haralick, R. M. "Zero-crossings of second directional derivative edge operator," *SPIE Proc. on Robot Vision*, Arlington, Va., 1982.

Hildreth, E. C. "Implementation of a theory of edge detection," *AI-TR-579*, MIT AI Lab, 1980.

Lunscher, W. H. H. "The asymptotic optimal frequency domain filter for edge detection," *IEEE Trans. PAMI*, **6**, 678–680, 1983.

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

Marr, D. and Poggio, T. "A computational theory of human stereo vision," *Proc. R. Soc. Lond. B*, **204**, 301–328, 1979 (also *AI Memo 451*, MIT AI Lab, 1977).

Poggio, T. and Torre, V. "Ill-posed problems and regularization analysis in early vision," *AI Memo 773*, MIT AI Lab, 1984.

Reinsch, C. H. "Smoothing by spline functions", *Numer. Math.*, **10**, 177–183, 1967.

Rosenfeld, A. and Kak, A. C. *Digital Picture Processing*, 2nd ed., Academic Press, New York, 1982.

Schoenberg, I. J. "Contributions to the problem of approximation of equidistant data by analytic functions," *Quart. Appl. Math.*, **4**, 45–99, 112–141, 1946.

Schoenberg, I. J. "Spline functions and the problem of graduation," *Proc. Nat. Acad. Sci. USA*, **52**, 947–950, 1964.

Shanmugam, K. F., Dickey, F. M. and Green, J. A. "An optimal frequency domain filter for edge detection in digital pictures," *IEEE Trans. PAMI*, **1**, 37–49, 1979.

Terzopoulos, D. "Multiresolution computation of visible-surface representations," Ph.D. Thesis, Dept. of EECS, MIT, 1984a.

Terzopoulos, D. "Controlled Smoothness Stabilizers for Ill-Posed Visual Problems Involving Discontinuities" Proc. Image Understanding Workshop, SAIC, 1984b.

Tikhonov, A. N. "Solution of incorrectly formulated problems and the regularization method," *Soviet Math. Dokl.*, **4**, 1035–1038, 1963.

Tikhonov, A. N. and Arsenin, V. Y. *Solutions of ill-posed problems*, Winston and Sons, Washington, D.C., 1977.

Torre, V. and Poggio, T. "On edge detection," *AI Memo 768*, MIT AI Lab, 1984.

Wahba, G. "Ill-posed problems: numerical and statistical methods for mildly, moderately, and severely ill-posed problems with noisy data," *Tech. Report 595*, Univ. of Wisconsin, Madison, 1980.

Witkin, A. "Scale-Space Filtering," *Proc. IJCAI*, 1019–1021, Karlsruhe, 1983.

Yuille, A. L. and Poggio, T. "Scaling Theorems for Zero-crossings," *AI Memo 722*, MIT AI Lab, 1983a.

Yuille, A. L. and Poggio, T. "Fingerprints Theorems for Zero-crossings," *AI Memo 730*, MIT AI Lab, 1983b.