MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No.1078 September 1988

An Optimal Scale for Edge Detection

Davi Geiger and Tomaso Poggio

Abstract:
Many problems in early vision are ill posed[1]. Edge detection is a typical example. This paper applies regularization techniques to the problem of edge detection. We derive an optimal filter for edge detection with a size controlled by the regularization parameter λ and compare it to the Gaussian filter. A formula relating the signal-to-noise ratio to the parameter λ is derived from regularization analysis for the case of small values of λ. We also discuss the method of Generalized Cross Validation for obtaining the optimal filter scale. Finally, we use our framework to explain two perceptual phenomena: coarsely quantized images becoming recognizable by either blurring or adding noise.

This report describes research done within the Artificial Intelligence Laboratory. Support for the A.I. 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 (ONR) contract N00014-85-K-0124. Davi Geiger is supported by Comissao National de Energia Nuclear.

1 Introduction

Edge detection can be considered as a problem of numerical differentiation which can be shown to be an ill-posed problem [3]. One approach to solve such an ill-posed problem is to regularize it. Standard regularization techniques suggest the use of Gaussian-like filters before differentiation[2,4]. In this paper, we address the issue of how to estimate the optimal scale of the filter, that is, the amount of smoothing required by the given image data. More precisely, we will:

1. Obtain a regularization filter for edge detection and show that it can be seen as a generalization of the Poggio, Voorhees and Yuille filter (1985) (PVY filter) [2].

2. Understand the role of the parameter λ (a problem mentioned by T.Poggio and V.Torre 1984 [3]) and find out how to calculate it. We will show that the optimal value of λ, which controls the size of the filter, is related to the signal-to-noise ratio.

3. Compare the optimal filter with the Gaussian filter and PVY filter. We show how stable σ (the size of the gaussian filter) is with respect to λ. This also corresponds to the stability of the width of the optimal filter with respect to the parameter λ.

4. Show that the perceptual phenomena of improved recognizability of coarsely quantized images by either adding noise or blurring can be understood in terms of scale changes in the edge detector's filter.

5. Apply this filter to edge detection. Given the image data, we use Canny's[7] edge detector but using the scale provided by the regularization analysis. More precisely, the noise in the image is first computed, then the signal-to-noise ratio is obtained and finally the λ parameter is used to set the scale of the filter used in computing the edges.

6. Compare the standard regularization techniques with the Generalized Cross Validation method for finding the optimal λ.

7. Propose biological mechanisms for selecting the optimal scale.

We now give an outline of the paper. Chapter 2 is a brief summary of results relating to ill-posed problems, regularization analysis and the variational method. In Chapter 3 we derive the optimal filter. We show that this filter gives the solution of the ill-posed problem of finding the exact signal from noisy data. In Chapter 4 we compare the optimal filter, Gaussian filter and PVY filter. In Chapter 5 we analyze the role of the parameter λ. In Chapter 6 we present an implementation of the edge detection results from Chapter 4 for computational vision. In Chapter 7 we give a possible explanation for the phenomena of improved recognizability of coarsely quantized image when either noise is added or blurring. In Chapter 8 we compare the method of Generalized Cross Validation (GCV) for finding λ with the standard regularization analysis.

2 Framework for Edge Detection

Regularization Techniques for Ill-posed Problems
A problem

$$Az = u \tag{2.1}$$

for which the class S of solutions z, given A and u, is not compact (changes on the right-hand side of the equation can take u outside the set AS) is called ill-posed. The approach suggested by Tikhonov to deal with ill-posed problems is to construct approximate solutions of equation (2.1) that are stable under small changes in the data u.

If the right-hand side of equation (2.1) is known only approximately, we have $u(x,y) = u_T(x,y) + v(x,y)$, where $u_T(x,y)$ is the true solution and $v(x,y)$ is noise. Then,

$$z(\omega,\nu) = \frac{u(\omega,\nu)}{k(\omega,\nu)} = z_T(\omega,\nu) + \frac{v(\omega,\nu)}{k(\omega,\nu)}$$

where

$$Az = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K(x-\xi, y-\tau)z(\xi,\tau)d\xi\, d\tau \tag{2.2}$$

and $k(\omega,\nu)$- is the Fourier transform of $K(x,y)$. It would seem natural to take the solution of equation (2.1) as being

$$z(x,y) = \frac{1}{(2\pi)^2}[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} z_T(\omega,\nu)e^{-i(\omega x+\nu y)}d\omega\, d\nu + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{v(\omega,\nu)}{k(\omega,\nu)}e^{-i(\omega x+\nu y)}d\omega\, d\nu]$$

since $u_T(\omega,\nu) = k(\omega,\nu)z_T(\omega,\nu)$. However, this function may not exist since the last integral may diverge. Furthermore, even if this ratio does have an inverse Fourier transform, the deviation from zero (in the C- or L_2-metric) can be arbitrarily large, and thus, we cannot think of the exact solution of equation (2.1) as an approximate solution of the equation with approximate right-hand side.

Finding edges in an image is in general an ill-posed problem[4], since it involves taking an appropriate derivative of noisy data (notice that we do not specify which derivative operator should be used: it may be a directional derivative [4] or any other desirable differential operator). The differentiation of the function $u(x)$ is ill-posed, since it can be seen as a solution of equation 2.1 for the operator A of the form

$$\int_{-\infty}^{x} z(\xi)d\xi = \int_{-\infty}^{\infty} h(x-\xi)z(\xi)d\xi = u(x)$$

where $h(x)$ is the step function. As described by Rheinsch[9] and by Poggio and Torre [3], this problem can be regularized by smoothing the data before taking derivatives. The idea is to consider the regularized solution z of equation (2.1), with A being the imaging operator, such that z is sufficiently well-behaved for numerical differentiation and as close as possible to the true data.

Tikhonov[1] proves that, for the case of one dimensional image data, to approximate a solution of equation (2.1) one takes the solution of a different problem, the one of minimizing the functional given by the following equation, that is close to the original problem for small values of the error in the data:

$$M^{\lambda}[z,u] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[Az-u]^2 dx\, dy + \lambda\Omega[z] \qquad (2.3)$$

where $u(x,y)$ is the image data, λ is the regularization parameter, A is in our special case a convolution operator and $\Omega[z]$ is the stabilizing operator. We will be considering the special case where

$$\Omega[z] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)^{p/2} z(x,y)\right|^2 dx\,dy.$$

The operator $\Omega[z]$ forces to smooth the data and the motivation to choose this particular stabilizing operator is to guarantee a desired smoothness in the solution. The order of the derivative in $\Omega[z]$, controlled by the parameter p, should be high enough to ensure the appropriate degree of differentiability in z required by the derivative operations that has to be performed next.
It should be pointed out that, whereas the original problem (2.1) does not have the property of stability, the problem of minimizing the functional $M^\lambda[z,u]$ is stable under small changes in the right-hand side u. This stability has been attained by narrowing the class of possible solutions by introducing the stabilizing functional $\Omega[z]$.

3 The Optimal Filter

The Fourier transform of $(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2})^{p/2} z$ is $M(\omega,\nu)z(\omega,\nu)$, where $M(\omega,\nu) = (\omega^2+\nu^2)^p$. Using Parseval's theorem we can rewrite equation (2.3) as

$$M^\lambda[z,u] = \frac{1}{(2\pi)^2}\Big\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[k(\omega,\nu)z(\omega,\nu)-u(\omega,\nu)][k(-\omega,-\nu)z(-\omega,-\nu)-u(-\omega,-\nu)]d\omega\,d\nu$$
$$+\lambda\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\omega^2+\nu^2)^p z(\omega,\nu)z(-\omega,-\nu)d\omega d\nu\Big\}.$$

The associated Euler-Lagrange equation is

$$\frac{\partial M^\lambda}{\partial z(-\omega-\nu)} = [k(\omega,\nu)z(\omega,\nu)-u(\omega,\nu)]k(-\omega,-\nu)+\lambda(\omega^2+\nu^2)^p z(\omega,\nu) = 0.$$

For the special case where $p=2$ (stabilize with Laplacian) and the operator A is given by (2.2) with $k(\omega,\nu) = e^{-\frac{b}{2}(\omega^2+\nu^2)}$, we obtain the regularized solution

$$z_\lambda(x,y) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{1+\lambda(\omega^2+\nu^2)^2 e^{b(\omega^2+\nu^2)}}z(\omega,\nu)e^{-i(\omega x+\nu y)}d\omega\,d\nu$$

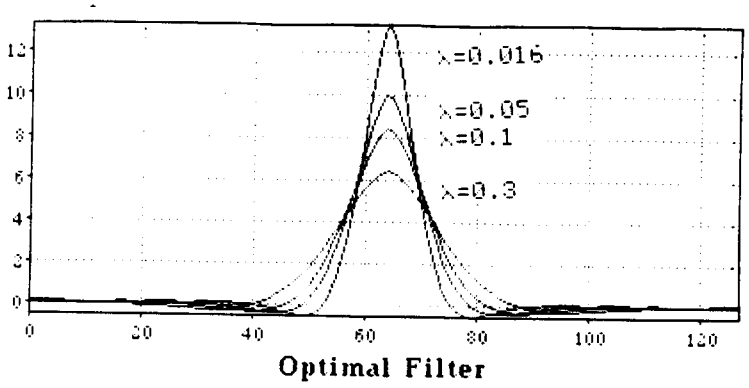


Figure 1: *The optimal filter is plotted in the space domain for different values of the regularization and scale parameter* λ

that corresponds, in the Fourier domain, to filtering the data with

$$f(\omega, \nu, \lambda) = \frac{1}{1 + \lambda(\omega^2 + \nu^2)^2 e^{b(\omega^2+\nu^2)}}. \tag{3.1}$$

The term $k(\omega, \nu)$ approximates the point spread function of the imaging device. For the human eye the $k(\omega, \nu)$ is described by a Bessel function and a Gaussian can be a good approximation. In this case the bandwidth (σ) of the point spread function of the pupil is of the order of 20 seconds of arc which implies $b = 0.5\sigma^2 = 1.5\,10^{-5}$ per degree square. For CCD cameras the bandwidth is less than a pixel, restricting the effect of this term. In our experiments we used $b = 0.5$ per pixel square.
The case b=0.0 gives the PVY filter. In this case the filter is not smooth enough to ensure differentiability with a second order differential operator such as the Laplacian and a higher order stabilizer is needed [2]. For $b > 0$ this problem, however, disappears. The filter is plotted ($b = 0.5$) in the space domain for different values of the parameter λ (see figure 1), which controls the size of the filter.

Appendix 2 shows how to obtain the Wiener filter using the variational method. We can see that the Wiener filter is slightly different from the optimal filter, because of the stabilizer term.

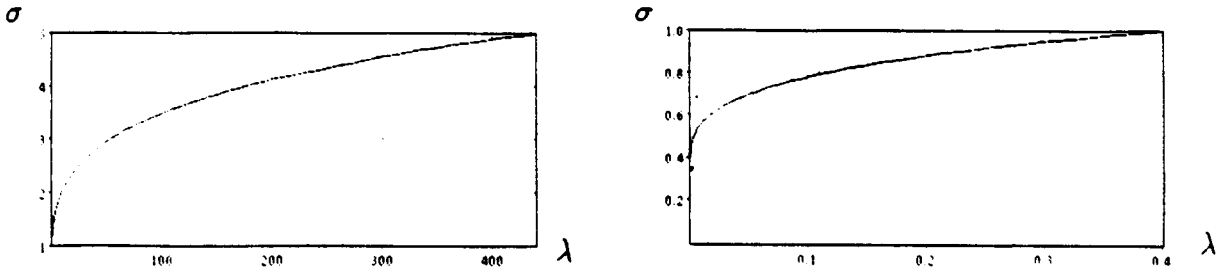


Figure 2: *The relation between λ and σ - the width of the optimal filter and of the Gaussian, respectively.* $\lambda = \frac{\sigma^4}{ln4} e^{\frac{-b(ln4)^{\frac{1}{2}}}{\sigma^2}}$

4 Filters: Optimal, Gaussian, PVY and Wiener

4.1 The relation between λ , σ , α and γ

In order to compare the filters, we have to assume an equivalent criterion for the Optimal filter, the Gaussian filter, PVY filter and Wiener filter. A possible criterion is to choose the same half amplitude in the frequency domain.

1) Gaussian-filter - $\frac{1}{2} = e^{\frac{-\sigma^2\omega_o^2}{2}}$ then $\omega_o = \frac{(ln4)^{\frac{1}{2}}}{\sigma}$

2) Optimal filter - $\frac{1}{2} = \frac{e^{-b\omega_o^2}}{e^{-b\omega_o^2}+\lambda\omega_o^4}$ then $\lambda\omega_o^4 = e^{-b\omega_o^2}$

3) PVY-filter - $\frac{1}{2} = \frac{1}{1+\alpha\omega^4}$ then $\alpha\omega^4 = 1$

4) Wiener-filter - $\frac{1}{2} = \frac{e^{-b\omega_o^2}}{e^{-b\omega_o^2}+\gamma}$ then $\gamma = e^{-b\omega_o^2}$

Then we plot a graph σ x λ for b=0.5 per pixel square (see figure 2)

We notice here the stability of the parameter σ (the bandwidth,in pixel unity,of a filter) with respect to the parameter λ.

Another criterion for equivalence is to chose filters that have the same bandwidth in their second-derivative (see figure 4). This is important when computing zero-crossings, since in this case the desired operator of convolu-

tion is the second derivative of the filter. The value of the parameter λ may vary with these different definitions, but the actual bandwidth of the filter will not vary significantly.

5 The Optimal Scale

In order to find the optimal scale we vary the parameter λ to obtain the closest solution to the true solution. More precisely (see appendix 3), we minimize

$$E\{|z_\lambda(x,y)-z_T(x,y)|^2\} = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\lambda^2(\omega^2+\nu^2)^4 S(\omega,\nu)+e^{-b(\omega^2+\nu^2)}N(\omega,\nu)}{[e^{-b(\omega^2+\nu^2)}+\lambda(\omega^2+\nu^2)^2]^2}d\omega\, d\nu \tag{5.1}$$

where $E\{.\}$ is the expectation value operator, $S(\omega,\nu)$ is the spectral density (power spectrum) of $z_T(x,y)$ and $N(\omega,\nu)$ is the spectral density of $v(x,y)$, assuming that $v(x,y)$ is stationary.
Differentiating equation 5.1 with respect to λ we obtain

$$\int_0^{\infty}\int_0^{\infty}\lambda\frac{e^{-b(\omega^2+\nu^2)}S(\omega,\nu)(\omega^2+\nu^2)^4}{(e^{-b(\omega^2+\nu^2)}+\lambda(\omega^2+\nu^2)^2)^3}d\omega\, d\nu = \int_0^{\infty}\int_0^{\infty}\frac{N(\omega\,\nu)e^{-b(\omega^2+\nu^2)}(\omega^2+\nu^2)^2}{(e^{-b(\omega^2+\nu^2)}+\lambda(\omega^2+\nu^2)^2)^3}d\omega\, d\nu \tag{5.2}$$

For the discrete case the space is defined in a grid with $N \times M$ pixels and the spatial frequency given by $\omega_i = \frac{2\pi}{N}i$ and $\nu_j = \frac{2\pi}{M}j$ where $i = -\frac{N}{2},...,0,...,\frac{N}{2}$ and $j = -\frac{M}{2},...,0,...,\frac{M}{2}$. We follow the same calculations as in the continuous case and keeping the order of the stabilizer a general parameter p (before we set $p=2$). Using the symmetry with respect to zero of $S(\omega,\nu)$ *and* $N(\omega,\nu)$ one obtain similarly to equation 5.2 the equation

$$\sum_{i=0}^{\frac{N}{2}}\sum_{j=0}^{\frac{M}{2}}\lambda\frac{e^{-b2\pi((\frac{i}{N})^2+(\frac{j}{M})^2)}S(i,j)(2\pi)^{4p}((\frac{i}{N})^2+(\frac{j}{M})^2)^{2p}}{\left(e^{-b2\pi((\frac{i}{N})^2+(\frac{j}{M})^2)}+\lambda(2\pi)^{2p}((\frac{i}{N})^2+(\frac{j}{M})^2)^p\right)^3} =$$

$$\sum_{i=0}^{\frac{N}{2}}\sum_{j=0}^{\frac{M}{2}}\frac{N(i,j)e^{-b2\pi((\frac{i}{N})^2+(\frac{j}{M})^2)}(2\pi)^{2p}((\frac{i}{N})^2+(\frac{j}{M})^2)^p}{\left(e^{-b2\pi((\frac{i}{N})^2+(\frac{j}{M})^2)}+\lambda(2\pi)^{2p}((\frac{i}{N})^2+(\frac{j}{M})^2)^p\right)^3} \tag{5.2b}$$

Within the range of the parameter λ there are two cases two consider. One for very small values, i.e. when $\lambda \to 0$, and the other case for λ not small. The second case is applicable whenever the signal-to-noise ratio is small and to find λ one has to solve equation 5.2b. The first case is applicable whenever the signal-to-noise ratio is very small and the following short cut to find λ can be used. We assume the asymptotic values

$$\lim_{\omega,\nu\to\infty} S(\omega,\nu) = \frac{S_o}{(\omega^2+\nu^2)^c} \qquad c>0 \qquad\qquad \lim_{\omega,\nu\to\infty} N(\omega,\nu) = \frac{N_o}{(\omega^2+\nu^2)^a} \qquad a \geq 0$$

and considering the case with white noise ($a = 0$) we obtain (see appendix 4) λ as the solution of

$$-b(\frac{S_o}{N_o}\lambda)^{\frac{1}{c-p}} = \ln\Big[\lambda^{\frac{c}{c-p}}(\frac{S_o}{N_o})^{\frac{p}{c-p}}\Big(2b(\frac{S_o}{N_o}\lambda)^{\frac{1}{c-p}}+(4p-1)\Big) \Big/ \Big(4b(\frac{S_o}{N_o}\lambda)^{\frac{1}{c-p}}+2p+1\Big)\Big]. \tag{5.3}$$

For the case $b = 0.0$ (the PVY filter) we get $\lambda = (2p+1/4p-1)^{\frac{(c-p)}{c}}(\frac{N_o}{S_o})^{\frac{p}{c}}$. A graph of signal-to-noise ratio versus λ for $b = 0.5$ per pixel square and is plotted in figure 3. Thus the optimal size of the filter can be obtained directly from an estimate of the signal-to-noise ratio for any given image. The equation 5.3, for finding the parameter λ, turns out to be sensitive when we fix the value of the signal-to-noise ratio and vary the values of b or c. However the relevant parameter, namely, the width of the filter is not very sensitive to it. In our experiments the width did not vary much from 0.8 pixels to 1.1 pixels. The parameter c was varied from 2 to 15, the parameter p (degree of smoothness) was varied from 2 to 6 and the parameter b (optical blurring) was varied from 0 to 2. We have to remember that here we are just considering large values of signal-to-noise ratio, typically $\frac{S_o}{N_o} > 100$.

6 Noise Estimation

Here we assume the noise to be caused by the deficiency in the optical system or by artificially adding a random variable to the image. So in this way noise is independent of the recognition task and independent of the scene or a particular scale. To estimate the noise we use a technique developed by H.Voorhees [10]. First, the gradient of the image is computed. A Rayleigh

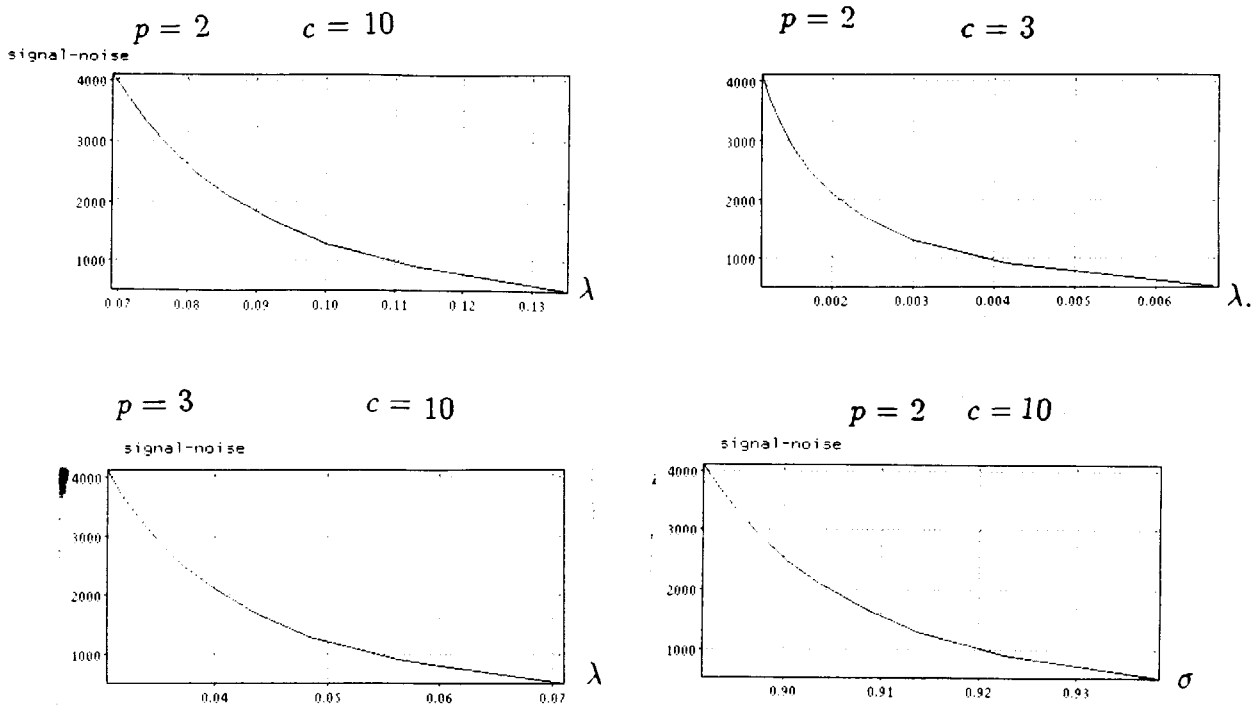


Figure 3: *Signal-to-noise ratio versus λ, for $b = 0.5$. The last curve the x-axis is given by the width of the filter σ*

distribution is then fit to the histogram of the norm of the gradient and the noise parameters are estimated. The signal power is obtained from the variance of the intensity of the image. With this method the program estimates from the image data the corresponding signal-to-noise ratio. This ratio gives λ from relations such as equation 5.3 or 5.2b. The results are shown in figures: 4, 5 and 6. One alternative for noise estimation could be done by taking a sequence of images and from the temporal correlation obtain the noise parameters. Another one is, assuming the noise to be white, to find the average value for large frequencies of the image-spectrum, since typically for large frequencies the noise is constant and the signal values close to zero.

7 Two perceptual phenomena: explanations and biological implications

7.1 Coarse quantized images can be better recognized when noise is added

We first discuss the perceptual phenomena of improved recognizability of coarse quantized images when noise is added[5]. Consider an image with 320 by 384 pixels and 8 bits. A coarsely quantized version of it is shown in figure 5a. The optimal filter for figure 5a, estimated as explained above,

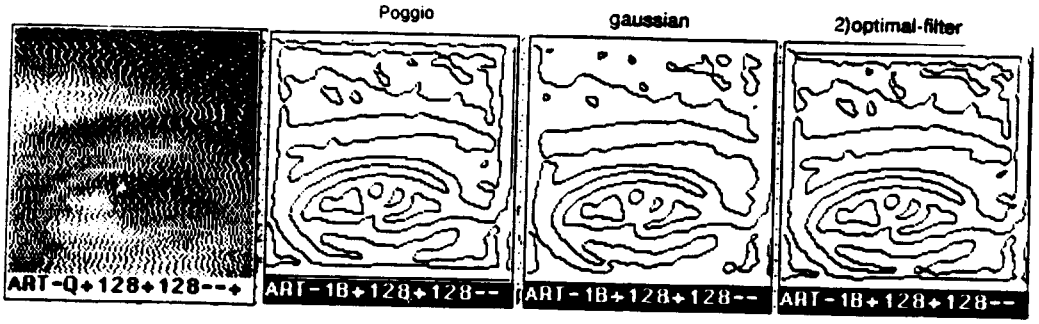


Figure 4: *Zero-crossings , the Laplacian applied to an image that was smoothed using a Gaussian filter and the optimal filter with the same width for the second derivative*

turns out to have a small scale of $\approx$ 1.0 pixel corresponding to very large signal-to-noise ratio of $S_o/N_o = 1000.0$. We then apply the Canny's edge detector using a gaussian filter with the same bandwidth predicted by the regularization. The edges do not easily reveal a face as we see in figure 5c. Gaussian white noise with standard deviation 70 is added to figure 5a (see figure 5b), making recognition easier. Estimation of the optimal scale, using equation 5.2b since the signal-to-noise ratio is small, gives now a width of $\approx$ 6.0 pixels. The corresponding contours (output from Canny's detector) reveal the face in a much better way.

These results may shed some light on what the visual system may be doing. Harmon and Julesz [6] claim that for the quantized image "high frequencies introduced by quantized blocking mask the lower spatial frequencies which convey information about the face, preventing recognition". In our framework two processes determine recognizability of the face. The first process consists of the estimation of the signal-to-noise ratio (S_o/N_o). The second step is to use S_o/N_o to set the optimal λ for computing the corresponding "edges". In the case of the quantized image the ratio S_o/N_o is large. λ is then small, which implies that a large bandwidth channel (in the spatial-frequency domain) is selected. The zero crossings for this channels do not

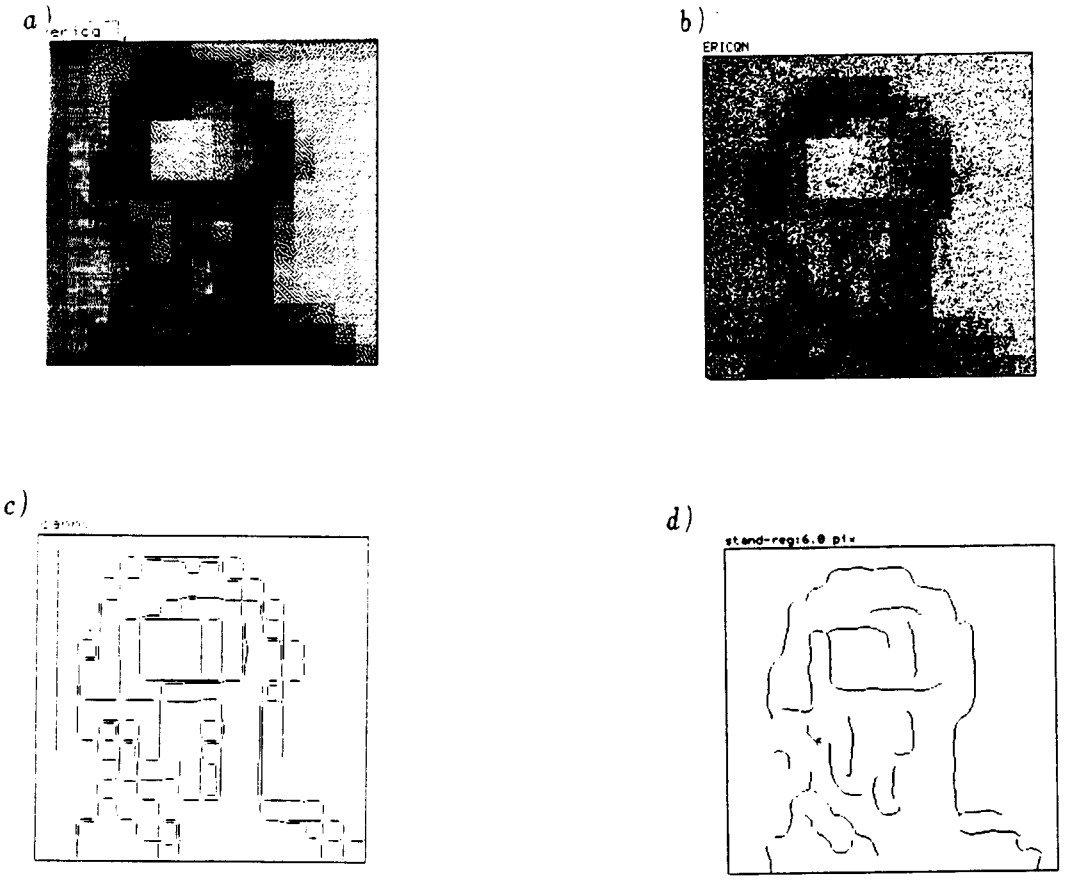


Figure 5: *a,c) Quantized image and edges with $\sigma = 1.0$ pixel. b,d) White noise, standard deviation of 70 units, was added to the quantized image. Now the edges are computed using $\sigma = 6.0$ pixels.*

easily allow face recognition because they mostly capture the box outlines. For noisy quantized images the ratio S_o/N_o is small and correspondingly λ is large. This implies a filter with small bandwidth. In this case the small bandwidth filter suppresses the noise and, as a side effect, also the high frequency outlines of the boxes.

This explanation is not in contrast with the one given by Canny [7] or by the one given by Morrone, Burr and Ross (1983) when they claim "that added noise (more high frequencies) destroys the propensity to organize the image according to its spurious high-frequency structure,...", but is more precise.

7.2 Improved recognizability of coarse quantized images by blurring

Blurring coarsely quantized images also improves recognition[6]. The explanation for this second perceptual phenomena is natural in our model. Consider the zero crossings of the blurred image. For the Gaussian filter they are represented by the solution of

$$\nabla^2(G(\sigma_2;x,y) * (G(\sigma_1;x,y) * I(x,y))) = 0$$

Where $I(x,y)$ is the quantized image, $*$ stands for convolution, $G(\sigma;x,y)$ is the Gaussian and the blurring process is $(G(\sigma_1;x,y) * I(x,y))$. Because all of those operations are linear we can first convolve both Gaussians. In the Fourier domain this is equivalent to multiplying the Gaussians, and the multiplication of two Gaussians is a Gaussian with $\sigma = [(\sigma_1)^2 + (\sigma_2)^2]^{1/2}$. So computing zero crossing of a blurred image is the same as computing zero crossings of the same image but with larger σ $(\sigma = [(\sigma_1)^2 + (\sigma_2)^2]^{1/2})$. Therefore blurring is equivalent to using an effectively larger filter for edge detection. This has the effect of suppressing the spurious high frequency edges introduced by coarse quantization. [1]

[1] Notice that blurring the quantized *noisy* image has the effect of increasing the estimated signal-to-noise ratio, thereby reducing σ_2 to a value close to the one obtained for the quantized image.

8 The Method of Generalized Cross Validation and Regularization

When S_o/N_o cannot be directly estimated, it is natural to consider the method of Generalized Cross Validation (GCV)[8]. The GCV method states that the optimal value of λ can be obtained by minimizing the functional (here in one dimension)

$$V(\lambda) = \frac{1}{n}\sum_{i=0}^{n} \frac{[Az_{n,\lambda}(t_i) - u_i]^2}{(1 - a_{kk}(\lambda))^2}\omega_k^2(\lambda) \tag{8.3}$$

where $\omega_k(\lambda) = (1 - a_{kk}(\lambda))/(1 - \frac{1}{n}\sum_{j=0}^{n} a_{jj}(\lambda))$ and $a_{kk}(\lambda) = \frac{\partial}{\partial z_k}(Az_{n,\lambda})(t_k)$. From here on we use the Gaussian filter , since would be computationally more expensive to use the optimal filter. Equation 8.3 reduces to

$$V(\sigma) = \frac{1}{n}\frac{1}{(1 - \frac{1}{\sqrt{2\pi}\sigma})^2}\sum_{i=0}^{n}\left[\frac{1}{\sqrt{2\pi}\sigma}\sum_{k=0}^{n} e^{\frac{-(i-k)^2}{2\sigma^2}} z(k) - z(i)\right]^2.$$

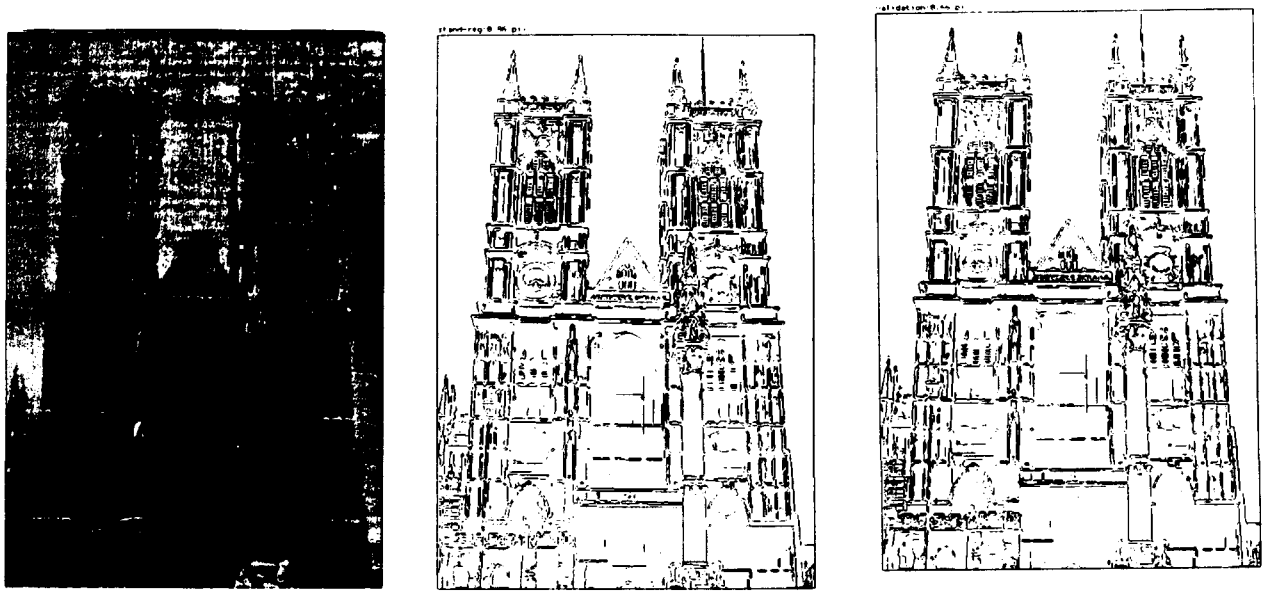
The method is computationally expensive but intrinsically parallel. We have implemented it on the Connection Machine [2]. We tested this method on different images including the ones in figure 5 with various amounts of noise. We notice a consistency of the GCV with the results obtained with the standard regularization. The GCV gave allways (see figure 6) a slightly maller value for the parameter λ (and σ, the width of the filter). We point out that equation 5.3 is just applicable when the signal-to-noise ratio is high (higher than 200) and the other cases we used equation 5.2b. The results of applying Canny's edges detector with both estimates of the parameter suggest a slightly better performance for the standard regularization method. For large amount of noise as in figure 5b we obtained significantly different results. For figure 5c for example the GCV method gave $\sigma = 3.5$ pixels as oppose to 6.0 given by the standard regularization method. This suggests that since no noise estimation is involved in the GCV method whenever the signal-to-noise ratio is small the standard regularization is going to give better results. Typically slices of the image of size 80 pixels require 20 milliseconds for computing $V(\sigma)$. Using Newton's method to find the minimum, the algorithm converges after 10 iterations of $V(\sigma)$. Therefore the GCV method takes in this

[2]The Connection Machine is a massivelly parallel computer with 65,536 processors

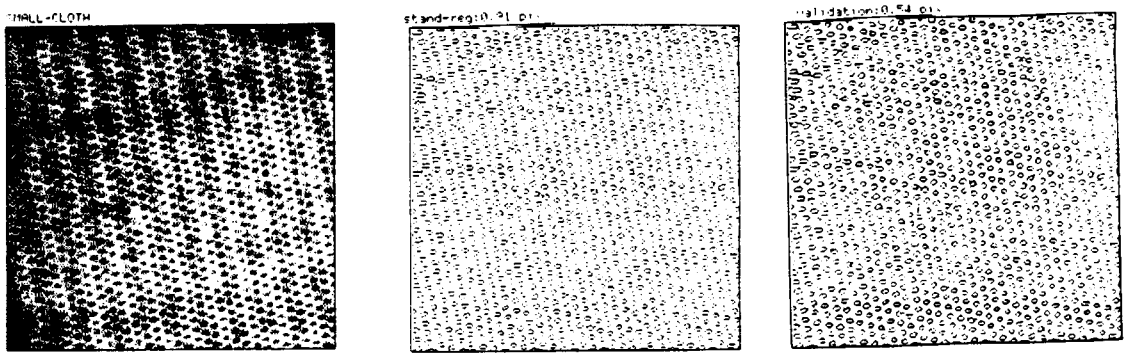
case 0.2 seconds to find the optimal σ. In figure 6 we show several examples where the optimal scale was estimated with the standard regularization and GCV methods.

It is interesting to notice that for the figure 6.a the optimal scale was for $\lambda = 0.1$ ($\sigma = 0.91$ pixels), however there are two distinct scales for this image that could be considered optimal. The coarser scale was not captured with any regularization method. A possible way of obtaining both scales would be by reconsidering the definition of noise or signal. For instance if the noise was defined to be the medium and high frequency of the spectrum then the noise estimation would be higher than the one we used and the coarser scale would be optimal.

Figure 6. Several examples of applying the standard regularization method and the GCV method to obtain the optimal scale

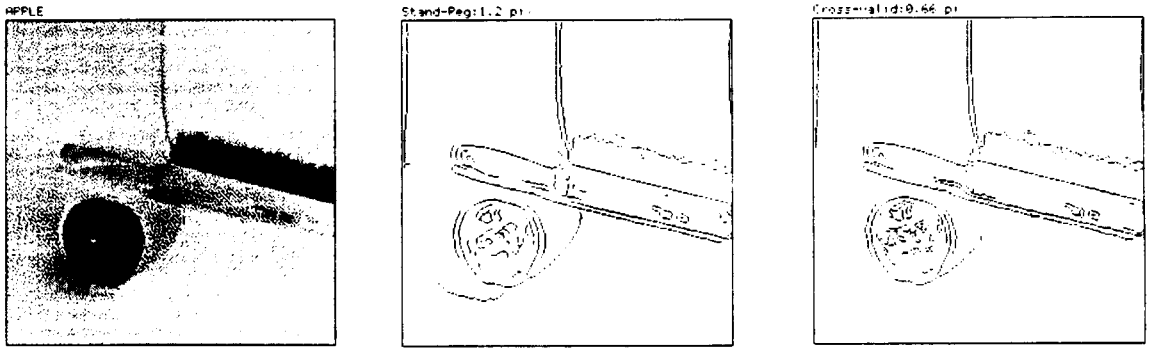


Westminst and Canny's edges with $\sigma = 0.86$ and $\sigma = 0.66$. Cross-validation: $\sigma = 0.66$ pixels, standard-regularization: $\sigma = 0.86$ pixels



Cloth and Canny's edges with $\sigma = 0.91$ and $\sigma = 0.54$. Cross-validation: $\sigma = 0.54$ pixels, standard-regularization: $\sigma = 0.91$ pixels.

Walter and Canny's edges with $\sigma = 1.63$ and $\sigma = 1.57$. Cross-validation: $\sigma = 1.57$ pixels, standard-regularization: $\sigma = 1.62$ pixels. We solved equation 5.b since So/No ratio smaller than 200.



Apple and Canny's edges with $\sigma = 1.2$ and $\sigma = 0.66$. Cross-validation: $\sigma = 0.66$ pixels, standard-regularization: $\sigma = 1.2$ pixels. We solved equation 5.b since So/No ratio smaller than 200.

Ericqn and Canny's edges with $\sigma = 6.00$ *and* $\sigma = 3.5$. *Cross-validation:* $\sigma = 3.48$ *pixels, standard-regularization:* $\sigma = 6.08$ *pixels. We solved equation 5.b since So/No ratio smaller than 200.*

9 Conclusion

We have derived rigorously the optimal way of filtering images prior to numerical differentiation. We also obtained the precise relation between the scale of the filter λ and the signal-to-noise ratio of the image. Some biological implications were also considered. In particular we suggested that humans can estimate the signal-to-noise ratio in the image from which the scale λ is computed. Only channels with the appropriate spatial-frequency band are then used, the others being inhibited. In this framework it is possible to understand the perceptual phenomena of improved recognizability of coarsely quantized images when noise is added. When the signal-to-noise ratio is large, the estimated λ is small and the associated edges do not provide good information for recognition. When S_o/N_o is smaller, the estimated λ is larger: the edges then provide better information for recognizing the face in the image. When the signal-to-noise ratio cannot be estimated, it is possible to use the method of Cross Validation for estimating the optimal λ.

Acknowledgments: We are grateful to A. Yuille for many discussions (with D.G.). Thanks to E. Hildreth for the suggestions, to E. Gamble, M. Gennert, and H. Voorhees who were always helpful and to J. Little for the last comments.

Appendix 1: Deviation of the regularized solution from the exact one and the stability of the solution

The regularized solution can be written in the form (see Chapter 3)

$$z_\lambda(x,y) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{f(\omega,\nu,\lambda)}{K(\omega)K(\nu)}u_T(\omega,\nu)e^{-i(\omega x+\nu y)}d\omega\,d\nu + \\ +\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{f(\omega,\nu,\lambda)}{K(\omega)K(\nu)}v(\omega,\nu)e^{-i(\omega x+\nu y)}d\omega\,d\nu$$

then

$$z_\lambda(x,y) - z_T(x,y) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\frac{\{f(\omega,\nu,\lambda)-1\}}{K(\omega)K(\nu)}u_T(\omega,\nu)e^{-i(\omega x+\nu y)}d\omega\,d\nu + \\ +\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\frac{f(\omega,\nu,\lambda)}{K(\omega)K(\nu)}v(\omega,\nu)e^{-i(\omega x+\nu y)}d\omega\,d\nu \qquad (1.a1)$$

Now one defines

$$\Delta_S(t,\lambda) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\frac{\{f(\omega,\nu,\lambda)-1\}}{K(\omega)K(\nu)}u_T(\omega,\nu)e^{-i(\omega x+\nu y)}d\omega\,d\nu \qquad (1.a2)$$

$$\Delta_N(t,\lambda) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{f(\omega,\nu,\lambda)}{K(\omega)K(\nu)}v(\omega,\nu)e^{-i(\omega x+\nu y)}d\omega\,d\nu \qquad (1.a3)$$

so that $z_\lambda(x,y) - z_T(x,y) = \Delta_S(x,y,\lambda) + \Delta_N(x,y,\lambda)$ One notices that equation (1.a2) characterizes the influence of the regularization and equation (1.a3) characterizes the influence of the noise.
Under these assumptions, the variance of the random function Δ_N is

$$E\{\Delta_N^2(t,\lambda)\} = \varphi^2(\lambda) = \frac{1}{2\pi^2}\int_0^{\infty}\frac{f^2(\lambda,\omega,\nu)}{L(\omega,\nu)}N(\omega,\nu)d\omega\,d\nu$$

and $\Delta^2(\lambda) = sup_t\Delta_S^2(t,\lambda)$
How stable is the regularized solution; more precisely, how much does the regularized solution change with a change in the order of the stabilizer p ?
It is possible to show that:
1) The asymptotic ($\omega,\nu \to \infty$) behavior of Δ_S and Δ_N are the same for stabilizers of the form

$$M(\omega,\nu) = (\omega^2+\nu^2)^p + a_{p-1}(\omega^2+\nu^2)^{p-1} + ... + a_0 \qquad or \qquad M(\omega,\nu) = (\omega^2+\nu^2)^p$$

2) From formula 5.1

$$T(\lambda,p)=\frac{1}{2\pi^2}\left\{\int_0^\infty \frac{\lambda^2(\omega^2+\nu^2)^p S(\omega,\nu)}{[L(\omega,\nu)+\lambda(\omega^2+\nu^2)^p]^2}d\omega\,d\nu+\int_0^\infty \frac{L(\omega)N(\omega)}{[L(\omega,\nu)+\lambda(\omega^2+\nu^2)^p]^2}d\omega\,d\nu\right\}=$$
$$=\Delta^2(\lambda)+\varphi^2(\lambda)$$

obviously, λ is a function of p. Consequently, $T(\lambda,p)=\psi(p)$ and the function $\psi(p)$ has a unique minimum at $p=p_o=c-a$, and increases monotonically in the interval $p>p_o$ with a finite limit for $p\to\infty$. Tikhonov [1] has shown these results for the one dimensional case.
Defining the measure of stability for arbitrary $p>0$ as being

$$\frac{\psi(p)-\psi(p_o)}{\psi(p_o)},$$

one has from Tikhonov [1]

$$0\le\frac{\psi(p)-\psi(p_o)}{\psi(p_o)}<\frac{\gamma_o}{1-\gamma_o}$$

where $\gamma_o=\frac{2c-1}{2n+2c-2a}$ and n is the order of smoothness of the operator $K(\omega)$, i.e., how fast $k(\omega)\to 0$ when $\omega\to\infty$.
For the case where $k(\omega)=e^{-b\omega^2}$ then $n\to\infty$. Therefore $\frac{\psi(p)-\psi(p_o)}{\psi(p_o)}\to 0$, which is to say that the function $\psi(p)$ is weakly dependent on the order p of the regularization. Consequently, one can quite justifiably replace the optimal order of regularization p_0 with an arbitrary order $p>p_o$. For white noise $p_o=c$ and for $p=2$ one will be safe whenever working with images that at high frequencies do not fall to zero faster than ω^{-2}.
It is clear is that if dealing with images with different asymptotic behavior, one can reset p to give a good solution, and the stability for higher p will be guaranteed.

In practice we are interested in the stability of the scale of the filter and we notice that the actual width of the filter did not vary more than 0.4 pixels when the parameters c, p and b (the point spread function parameter) were varied.

Appendix 2: The connection between the regularization method and optimal Wiener filtering.

The regularized solutions of equation 2.1 is

$$z_\lambda(x,y) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\frac{k(-\omega)k(-\nu)u(\omega,\nu)}{L(\omega,\nu)+\lambda M(\omega,\nu)}e^{-i(\omega x+\nu y)}d\omega$$

and can be written in the form of a convolution

$$z_\lambda(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} L_\lambda(t-\tau)u(\tau)d\tau$$

where

$$L_\lambda(t-\tau) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{k(-\omega)k(-\nu)}{L(\omega,\nu)+\lambda M(\omega,\nu)}e^{-i(\omega x+\nu y)}d\omega.$$

Now taking $T(\lambda M)$ (formula 5.1), from the condition that this functional be minimized on the set of functions $M(\omega,\nu)$, one finds by elementary calculations that the minimum is attained with the function

$$M(\omega,\nu) = M_o(\omega,\nu) = \frac{1}{\lambda}\frac{N(\omega,\nu)}{S(\omega,\nu)}.$$

Therefore the approximate (regularized) solution $z_{op}(x,y)$ of equation 2.1 obtained is

$$z_{op}(x,y) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{k(-\omega)k(-\nu)u(\omega,\nu)}{L(\omega,\nu)+\frac{N(\omega,\nu)}{S(\omega,\nu)}}e^{-i(\omega x+\nu y)}d\omega,$$

which is independent of the parameter λ. It coincides with the result of applying the optimal Wiener filtering to find $z(x,y)$ from $u(x,y) = u_T(x,y)+v(x,y)$.

Appendix 3: The expression for $E\{|z_\lambda(t) - z_T(t)|^2\}$

We show the result for a more general class of solutions where the smoothing term is given by $M(\omega,\nu) = (\omega^2+\nu^2)^p$. Notice that $L(\omega,\nu) = k(-\omega)k(-\nu)k(\omega)k(\nu)$. We have

$$z_\lambda(x,y) - z_T(x,y) =$$
$$= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{\frac{k(-\omega)k(-\nu)[u_T(\omega,\nu)+v(\omega,\nu)]}{L(\omega,\nu)+\lambda M(\omega,\nu)} - z_T(\omega,\nu)\}e^{-i(\omega x+\nu y)}d\omega d\nu =$$
$$= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{\frac{L(\omega,\nu)z_T(\omega,\nu)+k(-\omega)k(-\nu)v(\omega)}{L(\omega,\nu)+\lambda M(\omega,\nu)} - z_T(\omega,\nu)\}e^{-i(\omega x+\nu y)}d\omega d\nu =$$
$$= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{\frac{k(-\omega)k(-\nu)v(\omega,\nu)}{L(\omega,\nu)+\lambda M(\omega,\nu)} - \frac{\lambda M(\omega,\nu)z_T(\omega,\nu)}{L(\omega,\nu)+\lambda M(\omega,\nu)}\}e-i(\omega x+\nu y)d\omega d\nu$$

since $u_T(\omega,\nu) = k(\omega)k(\nu)z_T(\omega,\nu)$. Therefore

$$E\{|z_\lambda(x,y) - z_T(x,y)|^2\} =$$
$$= E\Big\{\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[\frac{k(-\omega)k(-\nu)v(\omega,\nu) - \lambda M(\omega,\nu)z_T(\omega,\nu)}{L(\omega,\nu)+\lambda M(\omega,\nu)}]e^{-i(\omega x+\nu y)}d\omega d\nu.$$
$$\cdot\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[\frac{k(-\omega')k(-\nu')v(\omega',\nu') - \lambda M(\omega',\nu')z_T(\omega',\nu')}{L(\omega',\nu')+\lambda M(\omega',\nu')}]e^{-i(\omega' x+\nu' y)}d\omega' d\nu'\Big\} =$$
$$= \frac{1}{(4\pi^2)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[\frac{\lambda^2 M(\omega,\nu)M(\omega,\nu')E\{z_T(\omega,\nu)z_T(\omega,\nu')\}}{[L(\omega,\nu)+\lambda M(\omega,\nu)][L(\omega',\nu')+\lambda M(\omega',\nu')]} +$$
$$+ \frac{k(-\omega)k(-\nu)k(-\omega')k(-\nu')E\{v(\omega)v(\omega',\nu')\}}{[L(\omega,\nu)+\lambda M(\omega,\nu)][L(\omega',\nu')+\lambda M(\omega',\nu')]}]e^{-i(\omega+\omega')x}e^{-i(\nu+\nu')y}d\omega d\nu d\omega' d\nu'$$

since $E\{v(\omega,\nu)\} = E\{v(\omega',\nu')\} = 0$. For stationary random processes,

$$E\{z_T(\omega,\nu).z_T(\omega',\nu')\} = S(\omega,\nu)\delta(\omega+\omega')\delta(\nu+\nu')$$
$$E\{v(\omega,\nu)v(\omega,\nu')\} = N(\omega,\nu)\delta(\omega+\omega')\delta(\nu+\nu'),$$

where $\delta(\omega+\omega')$ is Dirac's delta function. Performing the integration in the last expression with respect to ω' and using the properties of the delta function and the fact that $M(w)$ and $L(\omega)$ are even functions, one obtains the value of the deviation

$$E\{|z_\lambda(x,y)-z_T(x,y)|^2\} = T(\lambda M) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\lambda^2 M^2(\omega,\nu)S(\omega,\nu)+L(\omega,\nu)N(\omega,\nu)}{[L(\omega,\nu)+\lambda M(\omega,\nu)]^2}d\omega d\nu \quad .$$

Appendix 4: Solving equation 5.2, based on Tikhonov [1]

The argument for using the asymptotic expressions (p is the order of differentiation in the smoothing term) follows immediately from

$$\varphi^2(\lambda) = \int_0^\infty \int_0^\infty \lambda \frac{e^{-b(\omega^2+\nu^2)} S(\omega,\nu)(\omega^2+\nu^2)^p}{(e^{-b(\omega^2+\nu^2)} + \lambda(\omega^2+\nu^2)^p)^3} d\omega\, d\nu$$

since $\varphi^2(\lambda) \to \infty$ as $\lambda \to 0$. On the other hand, for any fixed $\lambda > 0$, the integral $\varphi^2(\lambda)$ converges. Therefore, for sufficiently small values of the parameter λ , the basic contribution to φ is provided by large frequencies ω. Then one can replace equation 5.2 by

$$\int_0^\infty \int_0^\infty \lambda \frac{e^{-b(\omega^2+\nu^2)} S_o(\omega^2+\nu^2)^{p-c}}{(e^{-b(\omega^2+\nu^2)} + \lambda(\omega^2+\nu^2)^p)^3} d\omega\, d\nu = \int_0^\infty \int_0^\infty \frac{e^{-b(\omega^2+\nu^2)} N_o(\omega^2+\nu^2)^{p-a}}{(e^{-b(\omega^2+\nu^2)} + \lambda(\omega^2+\nu^2)^p)^3} d\omega\, d\nu \qquad (5.2b)$$

with the change of variables $\omega = \xi\, cos\theta$ and $\nu = \xi\, sin\theta$ which imply $\xi^2 = \omega^2 + \nu^2$ and $d\omega\, d\nu = \xi\, d\xi\, d\theta$. We can rewrite equation (5.2b) as

$$\int_0^\infty \lambda \frac{L_{as}(\xi) S_o \xi^{2p-2c+1}}{(L_{as}(\xi) + \lambda\xi^p)^3} d\xi = \int_0^\infty \frac{N_o L_{as}(\xi) \xi^{2p-2a+1}}{(L_{as}(\xi) + \lambda\xi^p)^3} d\xi \qquad (5.3b)$$

where $L_{as}(\xi) = e^{-b\xi^2}$. The integrals

$$I_1 = \int_0^\infty \frac{L_{as}(\xi)\xi^{2p-2a+1}}{(L_{as}(\xi) + \lambda\xi^p)^3} d\xi \qquad I_2 = \int_0^\infty \frac{L_{as}(\xi)\xi^{2p-2c+1}}{(L_{as}(\xi) + \lambda\xi^p)^3} d\xi$$

are evaluated by the method of steepest descent and are equal to

$$I_1 = \frac{\xi_1^{2p-2a+1} L_{as}(\xi_1)}{\{L_{as}(\xi_1) + \lambda\xi_1^p\}^3} [\sqrt{\frac{\pi}{|f_1''(\xi_1,\lambda)|}} + O(\lambda^{\frac{3}{2}})]$$

$$I_2 = \frac{\xi_1^{p-2a+1} L_{as}(\xi_1)}{\{L_{as}(\xi_1) + \lambda\xi_1^p\}^3} [\sqrt{\frac{\pi}{|f_1''(\xi_1,\lambda)|}} \xi_1^{p-2c+2a} + O(\lambda^{\frac{3}{2}})],$$

where

$$f_1(\xi,\lambda) = \lambda ln \frac{\xi^{2p-2a+1} L_{as}(\xi)}{\{L_{as}(\xi) + \lambda\xi^p\}^3}$$

and the double prime denotes the second derivative with respect to ξ. Here, ξ_1 is a root of the equation

$$(-2L_{\alpha\alpha} + \lambda\xi^p)\frac{dL_{\alpha\alpha}}{d\xi} - (4p + 2\alpha)\lambda\xi^{p-1}L_{\alpha\alpha} + \frac{2p - 2\alpha}{\xi}L_{\alpha\alpha}^2 = 0 \quad .$$

Subetituting the values found for I_1 and I_2 into equation (5.3b) and keeping only the principal terms, one obtains the following equation for determining the optimal λ

$$\lambda = \frac{N_\alpha}{S_\alpha}[\xi_1(\lambda)]^{-2p+2\alpha-2\alpha} \quad .$$

References

1. Tikhonov, A.N. "Solution of Incorrectly Formulated Problems and the Regularization Method," *Soviet Math. Dokl.* 4, 1035-1038, 1963.

2. Poggio, T. and Voorhees, H. and Yuille, A. "A Regularized Solution to Edge Detection," Artificial Intelligence Lab. Memo, No. 833, MIT, Cambridge, MA, 1985.

3. Poggio, T. and Torre, V. "Ill-Posed Problems and Regularization Analysis in Early Vision," Artificial Intelligence Lab. Memo, No. 773, MIT, Cambridge, MA, 1984.

4. Torre, V. and Poggio, T. "On Edge Detection," Artificial Intelligence Lab. Memo, No. 768, MIT, Cambridge, MA, 1984.

5. Morrone, M.C. and Burr, D.C. and Ross, J. "Added Noise Restores Recognizability of Coarse Quantized Images," *Nature*, Vol. 305, 15 Sept. 1983.

6. Harmon, L.D. and Julesz, B. *Science* 180, 1194-1197, 1973.

7. Canny, J.F. "Finding Edges and Lines in Images," Artificial Intelligence Lab. Technical Report No. 720, MIT, Cambridge, MA, 1983.

8. Wahba, G. "Ill Posed Problems," Technical Report No. 595, 1980.

9. Reinsh, C.H. "Smoothing by Spline Functions," *Numer. Math.,* 10, 177-183, 1967.

10. Voorhees, H. "Computing Texture Boundaries," S.M. Thesis, MIT, Cambridge, MA, 1987 (in preparation).

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

12. Terzopoulos, D. "Computing Visible-Surface Representations," Artificial Intelligence Lab. Memo, No. 833, MIT, Cambridge, MA, 1984.

13. Marroquin, J. "Probabilistic Solution of Inverse Problems," Artificial Intelligence Lab. Technical Report No. 860, MIT, Cambridge, MA, 1985.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AIM 1078	2. GOVT ACCESSION NO. AD-A202 147	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) An Optimal Scale for Edge detection		5. TYPE OF REPORT & PERIOD COVERED memorandum
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Davi Geiger and Tomaso Poggio		8. CONTRACT OR GRANT NUMBER(s) N00014-85-K-0214
9. PERFORMING ORGANIZATION NAME AND ADDRESS Artificial Intelligence Laboratory 545 Technology Square Cambridge, MA 02139		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency 1400 Wilson Blvd. Arlington, VA 22209		12. REPORT DATE September 1988 13. NUMBER OF PAGES 23
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Office of Naval Research Information Systems Arlington, VA 22217		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution is unlimited

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

Unlimited

18. SUPPLEMENTARY NOTES

None

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

edge detection
cross-validation
regularization

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

See back of page

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

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block 20 (cont'd)

Many problems in early vision are ill posed[1]. Edge detection is a typical example. This paper applies regularization techniques to the problem of edge detection. We derive an optimal filter for edge detection with a size controlled by the regularization parameter λ and compare it to the Gaussian filter. A formula relating the signal-to-noise ratio to the parameter λ is derived from regularization analysis for the case of small values of λ. We also discuss the method of Generalized Cross Validation for obtaining the optimal filter scale. Finally, we use our framework to explain two perceptual phenomena: coarsely quantized images becoming recognizable by either blurring or adding noise.

Scanning Agent Identification Target

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

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