AN OPTIMAL SCALE FOR EDGE DETECTION

Davi Geiger and Tomaso Poggio

Massachusetts Institute of Technology
Artificial Intelligence Laboratory

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, showing that the scale of the filter is a function of the signal-to-noise ratio. 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.

1. Introduction

If edge detection is considered as a problem of numerical differentiation, the first step is to regularize it. Standard regularization techniques suggest the use of Gaussian-like filters before differentiation[2,4]. In this paper, we address the important issue of how to estimate the optimal scale of the filter, that is, the amount of smoothing required by the given image data.

2. Framework for Edge Detection

Regularization Techniques for Ill-Posed Problems

A problem

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

for which the class 5 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[1] 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 $v(x,y)$ 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

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

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 $u(x)$ is ill-posed, since it can be viewed 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' 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 to equation (2.1), with *A* being the imaging operator, such that *z* is sufficiently well-behaved for numerical differentiation.

To approximate a solution of equation (2.1) one takes the solution of a different problem, that of minimizing the functional given bv.

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

that is close to the solution of the original problem for small values of the error in the data. Tikhonov[1] proved that for the case of one dimensional image data. Here

$u(x,y)$ 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 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 $M(\omega,\nu) = (\omega^2+\nu^2)^p$. 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 desired derivative operation 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

Using Parseval's theorem, we can rewrite equation (2.3) as

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

The associated Euler-Lagrange equation is

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

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

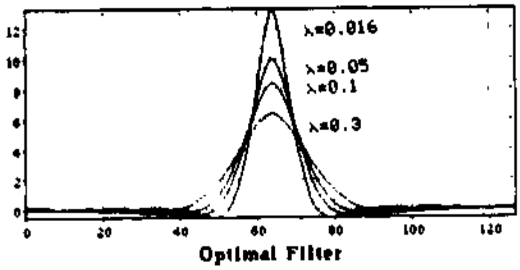
$$z_\lambda(x,y) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{z(\omega,\nu)e^{-i(\omega x+\nu y)}}{1+\lambda(\omega^2+\nu^2)^2 e^{b(\omega^2+\nu^2)}}d\omega\, d\nu.$$

This corresponds, in the Fourier domain, to filtering the input data with

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

The case b=0.0 $(degree)^2$ gives the PVY filter [2]. For $b=0.0$ 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 disappears however. The term $k(\omega,\nu)$ approximates the modulation transfer function of the imaging device. For the human eye it has bandwith ω=60 $(degree)^{-1}$, implying that b=0.2216 $(degree)^2$. The filter is plotted (b=0.2216) in the space domain for different values of λ (figure 1), which controls the size of the filter.

Figure 1. The optimal filter in the space domain for different values of the regularization parameter λ



In order to compare the optimal filter with the Gaussian filter, we have to use some measure for the filter's size. A possible criterion is the frequency at which the amplitude drops to half of its maximun value. This gives Gaussian: $\frac{1}{2} = e^{\frac{-\sigma^2\omega_o^2}{2}}$ and optimal-filter: $\frac{1}{2} = \frac{e^{-b\omega_o^2}}{e^{-b\omega_o^2}+\lambda\omega_o^4}$.

Therefore $\lambda = \frac{\sigma^4}{\ln 4}e^{\frac{-b(\ln 4)^{\frac{1}{2}}}{\sigma^2}}$. Another criterion is to chose filters that have the same width in their (appropriate) derivative. An example is given in figure 3 where the derivative operator is the Laplacian; both filters give very similar zero-crossings.

4. The Optimal Scale

In order to find the optimal scale we vary λ to obtain the closest solution to the true solution. More precisely, we minimize

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

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 wide sense stationary. In practice, it is a good approximation to assume white noise $N(\omega,\nu)=N_o$ and the asymptotic value

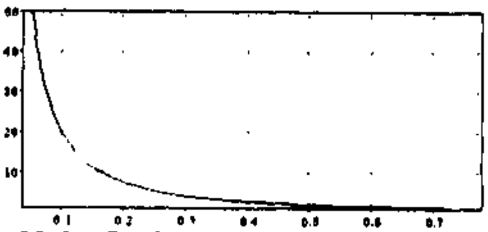
$$\lim_{\omega,\nu\to\infty} S(\omega,\nu) = S_o(\omega^2+\nu^2)^{-c}.$$

For $c=2$, we obtain $\lambda = \frac{N_o}{S_o}$ and for $c=3$, λ is the solution of

$$-b\frac{S_o}{N_o}\lambda = \ln\left(\lambda^3(\frac{S_o}{N_o})^2(2b\frac{S_o}{N_o}\lambda+7)/(4b\frac{S_o}{N_o}\lambda+5)\right). \quad (4.1)$$

For the case $b=0.0$ (the PVY filter) we get $\lambda = (5/7)^{\frac{1}{3}}(\frac{N_o}{S_o})^{\frac{1}{3}}$. A graph of signal-to-noise ratio versus λ for $b=0.2216$ (degree)^2 is plotted in figure 2. Thus the optimal size of the filter can be obtained directly from an estimate of the signal-to-noise ratio for any given image.

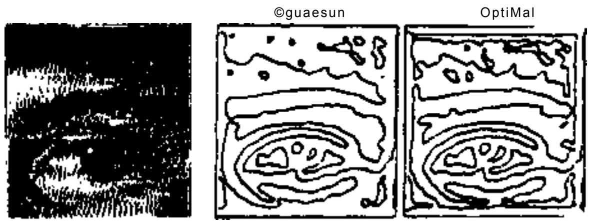
Figure 2. Signal-to-noise ratio versus λ, for b = 0.2216



5. Noise Estimation

To estimate the noise we use a technique .developed by Voorhees[10]. First, the gradient of the image is computed. A Rayleigh distribution is then fitted to the histogram of the norm of the gradient and the noise parameters are estimated. The signal power is obtained from the standard deviation of the histogram of the image. With this method the program estimates the signal-to-noise ratio from the image data. This ratio gives the parameter λ from relations such as equation 4.1. The results are shown in figures 3 and 4.

Figure S. Zero-crossings using Gaussian filter and optimal-filter with the same width for the second-derivative

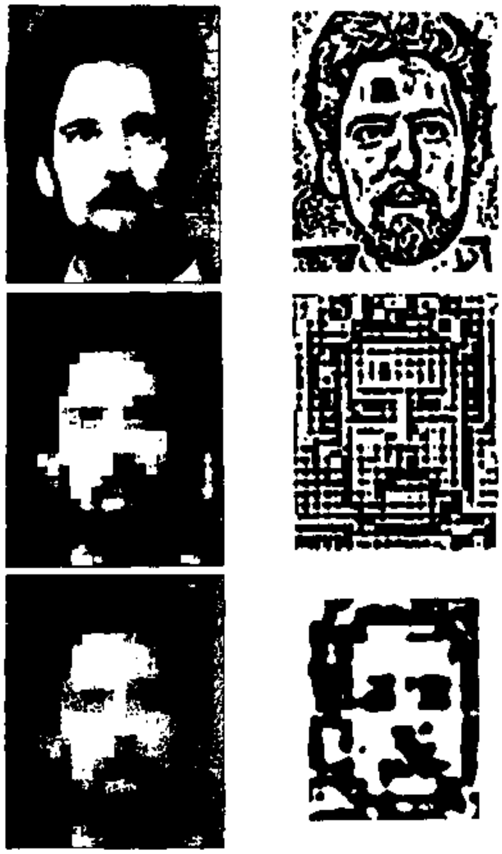


6. Two Perceptual Phenomena

6.1. 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 the image of figure 4a with 320 by 384 pixels. A coarsely quantized version of it is shown in figure 4c. The optimal filter for figure 4c, estimated as explained above, turns out to have a small scale of ≈ 2.0 pixels corresponding to very low noise $S_o/N_o = 100.0$. The sign of the zero-crossings (figure 4d) do not easily reveal a face. Gaussian white noise with standard deviation 70 is the added to figure 4c (see figure 4e), making recognition easier. Estimation of the optimal scale gives now a width of ≈ 6.5 pixels. The corresponding zero-crossing contours 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".

Figure 4. a,b) Image of a face and sign of the zero-crossings (szc) using Laplacian operator. c,d) Quantized image and szc with $\sigma = 2.0$ pixels. e,f) White noise, standard deviation of 70 units, was added to the quantized image. Now szc is $\sigma = 6.5$ pixels.



In our framework two process 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 then computing an appropriate derivative and 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 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 imply 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 Morrone, Burr and Ross 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.

6.2. Improved recognizability of coarse quantized images by blurring

Blurring coarsely quantized images also improve recognition[6]. The explanation for this second perceptual phenomena is natural in our model. 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.*

7. The Method of Generalized Cross Validation and Regularization

When S_0/N_0 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 A, can be obtained by minimizing the functional (here in one dimension)

$$V(\lambda) = \frac{1}{n}\sum_{i=1}^{n}\frac{[Az_{n,\lambda}(t_i) - u_i]^2}{(1 - a_{kk}(\lambda))^2}\omega_k^2(\lambda) \qquad (7.3)$$

where $\omega_k(\lambda) = (1 - a_{kk}(\lambda))/(1 - \frac{1}{n}\sum_{j=0}^{n} a_{jj}(\lambda))$ and $a_{kk}(\lambda) = \frac{\partial}{\partial z_k}(Az_{n,\lambda})(t_k)$

Assuming the filter to be Gaussian-like, using the optimal-filter would be computationally more expensive, equation 7.3 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 Connection Machine. We tested this method on different images including the ones in figure 4 with various amounts of noise. The important result was the consistence of the GCV with the results obtained by a method described earlier. Slices of the image 80 pixels require 20 miliseconds for computing $V(\sigma)$. Using Newton's method to find the minimun, the algorithm converges after 10 intensions of $V(\sigma)$. Therefore the GCV method takes in this case 0.2 seconds to find the optimal *a.*

8. 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 channels with the appropriate spatial-frequency band

* Notice that blurring the quantized *noisy* image has the effect of increasing the estimate of signal-to-noise ratio, thereby reducing 0_2 to a value close to the one obtained for the quantized image.

4. are then used, the others being inhibited. In this framework it is possible to understand the perceptual phenomena of improved recognizability of coarsely quantized image when noise is added. When the signal-to-noise ratio is large, the estimated λ is small and the associated zero-crossings do not provide good information for recognition. When S_o/N_o is smaller, the estimated A is larger: the zero-crossings provide then a 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. Gamble, M. Gennert, and H. Voorhees who were always helpful.

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 Boudaries," 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.