

Wavelets for SAR Image Smoothing

Graham Horgan

Abstract

Wavelets are an increasingly widely used tool in many applications of signal and image processing. This paper reviews the basic ideas of wavelets for representing the information in signals such as time series and images, and shows how wavelet shrinkage may be used to smooth these signals. This is illustrated by application to a synthetic aperture radar image. Some guidelines on using wavelet shrinkage are given.

Introduction

Wavelet analysis has recently been recognized as a tool with important applications in time series, function estimation, and image analysis. Applications in remote sensing have included the combination of images of different resolutions (Garguetduport *et al.*, 1996), image compression (Werness *et al.*, 1994), the provision of edge detection methods (Li and Shao, 1994), and the study of scales of variation (Lindsay *et al.*, 1996). As the development of wavelet methods is recent, the fundamentals are not yet widely understood, and guidance on their practical use is hard to find. Much of the literature is not easily accessible without much mathematical sophistication.

This paper explains the fundamental ideas of wavelets in a non-mathematical way. Their use for smoothing data and images, particularly by wavelet shrinkage, will be reviewed. This will be illustrated by application to a synthetic aperture radar image.

Wavelets

Wavelets arose from signal processing theory, a signal being the variability of some quantity over time. They can be generalized to two-dimensional signals, of which images are a special case. Wavelet decomposition is an alternative way of presenting the details of a signal which differs from specifying the value of the signal at successive times, the so-called *time domain* representation.

Other ways of doing this have been used for a long time. The best known is the Fourier series. This represents a signal in terms of sine waves with frequencies which are multiples (harmonics) of a basic frequency. In many applications this is a useful way of decomposing a signal, and its properties can be understood in terms of these oscillations at different frequencies. The sine waves used are orthogonal to each other (two functions $f(t)$ and $g(t)$ are orthogonal if $\int_{-\infty}^{\infty} f(t)g(t)dt = 0$). For a signal sampled at n points, a full reconstruction can be made from n Fourier components. This is termed the Fourier or *frequency domain* representation of a series.

A drawback to the Fourier representation is that the frequency components apply to the signal as a whole, and the way that the signal variability changes over time may be important. However, it is impossible to say what the frequency components are at a time point, but only in a region about it. To put it another way, a signal cannot be highly localized in

both the time and frequency domains (this limitation gives rise in physics to the *Heisenberg Uncertainty Principle*). One approach is to estimate Fourier components locally, in a set of short intervals, or convolved with a decaying weight function about each point. To do this, however, is to lose the orthogonality and economy of representation of the Fourier method. Wavelets preserve these advantages, while enjoying good spatial and frequency resolution.

The wavelet approach to signal representation is based on representing the signal at different scales or resolutions. At a particular resolution, the signal is approximated by a sum of *scaling functions*. The difference between the resolutions (termed the *detail* at the finer resolution) can be expressed by a sum of *wavelet* functions. For certain scaling and wavelet functions, this hierarchical or *multiresolution* representation can be constructed using scaled versions of the same functions at each resolution.

For example, consider the simplest multiresolution representation, based on the *Haar* scaling function and wavelet. These are illustrated in Figure 1. They are piecewise constant, and it is clear how they may be used to approximate signals. Figure 2 shows a signal crudely approximated by the Haar scaling function at two different resolutions.

It is clear that the signal could be approximated in this way at ever higher resolutions until it is as close to the original signal as we wish. If this is only available at discrete time points, the finest resolution approximation is exact. The *wavelet representation* consists of specifying an approximation to the signal at some coarse resolution, and then specifying the refinements needed at each stage to achieve the next resolution. In Figure 2, we see that the approximation (a) can be refined to (b) by the addition of two Haar wavelets (Figure 2b). One covers the first half of the time domain, and has a negative coefficient. The other covers the second half, and has a slightly smaller positive coefficient. A further four wavelets could then be added to (b) to refine it further. In the wavelet representation, we specify the coefficients of the scaling functions at the coarsest resolution (which could be just the average of the signal) and the coefficients of all the wavelets needed to refine it to the finest resolution. If the signal is of length 2^m , then $1 + 1 + 2 + 4 + \dots + 2^{m-1} = 2^m$ coefficients are needed. (The wavelet representation is most convenient for 2^m time points, and some modifications are needed when this is not so — see Discussion below.)

Among the most important benefits of a wavelet representation are

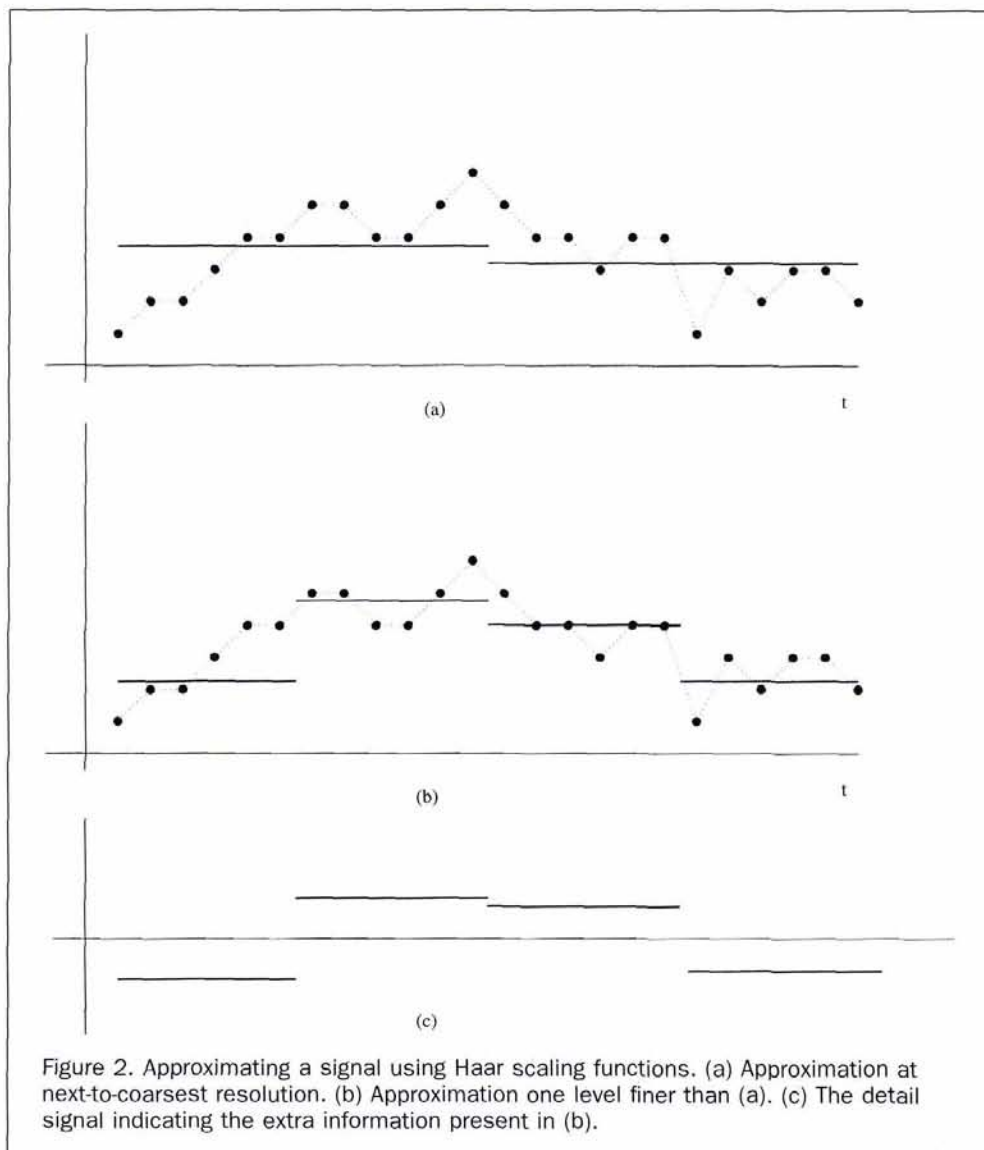
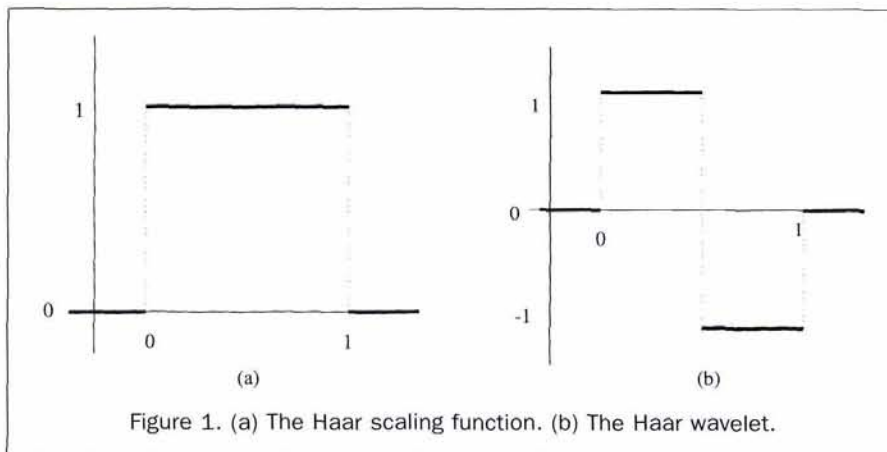
- Single wavelet coefficients provide information about how the signal is changing over time: for example, the coefficient of the wavelet which refines the coarsest (overall average) approximation to obtain the next resolution tells us about the overall increase or decrease of the signal;
- Some wavelet coefficients can be discarded (for economy of

Photogrammetric Engineering & Remote Sensing,
Vol. 64, No. 12, December 1998, pp. 1171–1177.

0099-1112/98/6412-1177\$3.00/0

© 1998 American Society for Photogrammetry
and Remote Sensing

Rowett Research Institute, Biomathematics & Statistics Scotland,
Bucksburn, Aberdeen AB21 9SB, Scotland.



representation) while still retaining the major features of the signal: this is a form of data compression; and

- If one believes that fine details are *noise*, contaminating the true signal of interest, then discarding or reducing them may give a result closer to the true signal.

This is reviewed in the next section.

The Haar wavelet is the simplest and easiest to understand. However, there are many possibilities. Fuller discussion of the properties of wavelets may be found in Strang

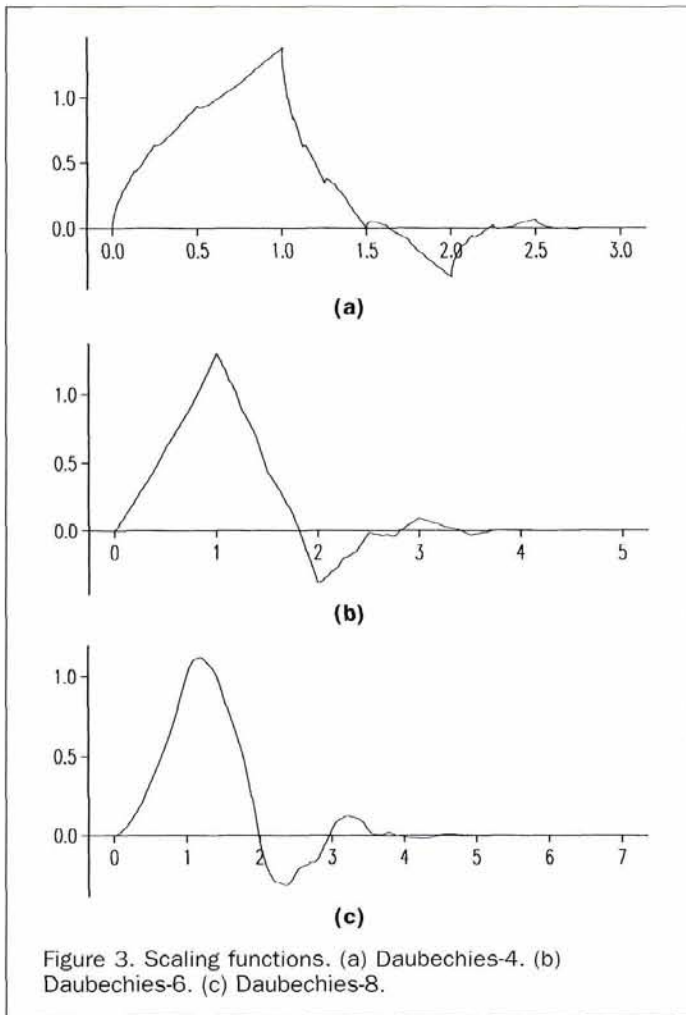


Figure 3. Scaling functions. (a) Daubechies-4. (b) Daubechies-6. (c) Daubechies-8.

(1989), Mallat (1989a; 1989b), Daubechies (1992), and Kay (1994). The two most important properties are that the scaling functions can be represented as the sum of scaled and translated versions of themselves (for example, the Haar scaling function $h(t)$ can be written $h(t) = h(2t) + h(2t - 1)$) and the wavelets should be orthogonal to scaled and translated versions of themselves. Note that by "scaled," we mean by a power of 2, and by "translated," we mean by an integer amount.

The principal disadvantage of the Haar wavelet is its discontinuity. This limits its ability to approximate continuous signals. A commonly used family of continuous scaling func-

tions and wavelets are those of Daubechies (1988). They are the only continuous wavelets with full scaling and translational orthogonality of scaling functions and wavelets which have a *finite support*: i.e., they are non-zero only over a finite range. To achieve these properties requires strange looking functions. Some of these — the Daubechies-4, Daubechies-6, and Daubechies-8 scaling functions — are shown in Figure 3. Although continuous, the Daubechies-4 is differentiable nowhere, and has some fractal self-similarity. It is defined as the solution of the recursion equation

$$d(t) = \frac{1 + \sqrt{3}}{4}d(2t) + \frac{3 + \sqrt{3}}{4}d(2t - 1) + \frac{3 - \sqrt{3}}{4}d(2t - 2) + \frac{1 - \sqrt{3}}{4}d(2t - 3).$$

Although it is not apparent from Figure 3, combinations of this scaling function can approximate linear functions very well. The Daubechies-6 and Daubechies-8 functions, which are smoother, are defined from similar recursions. The Haar wavelet may be thought of as Daubechies-2. These wavelets will be used in the section on Image Filtering. A useful source of computer algorithms for the Daubechies wavelets is Press *et al.* (1993).

Wavelets may also be used to represent higher dimensional signals such as images. One approach to a two-dimensional wavelet representation is to base the scaling functions and wavelets on products of one-dimensional functions. If $\phi(x)$ and $\psi(x)$ are a scaling function and wavelet, respectively, then $\phi(x)\phi(y)$ is a two-dimensional scaling function, and the detail needed to achieve the next finest resolution can be represented by three wavelets, $\phi(x)\psi(y)$, $\psi(x)\phi(y)$, and $\psi(x)\psi(y)$. A fuller description may be found in Mallat (1989a) and Kay (1994). It is also possible to construct two-dimensional wavelets that are not products of one-dimensional functions. For example, Hitchcock (1994) constructs a wavelet with 45° rotational symmetry.

Wavelet Shrinkage

Noise in a signal may be simply measurement error, or it could be fluctuation details which are a nuisance when the underlying trends or discontinuities are being investigated. Many methods have been developed for smoothing signals, in the hope that the noise can be suppressed and the significant patterns retained and revealed. These have ranged from simple moving averages or moving medians to methods of considerable mathematical complexity. Wavelets seem to offer a smoothing approach which is relatively simple to use, while adapting well, and automatically, to the form of the signal being smoothed. One way of doing this is by *shrinking* individual wavelet coefficients and reconstructing a signal from these shrunken coefficients.

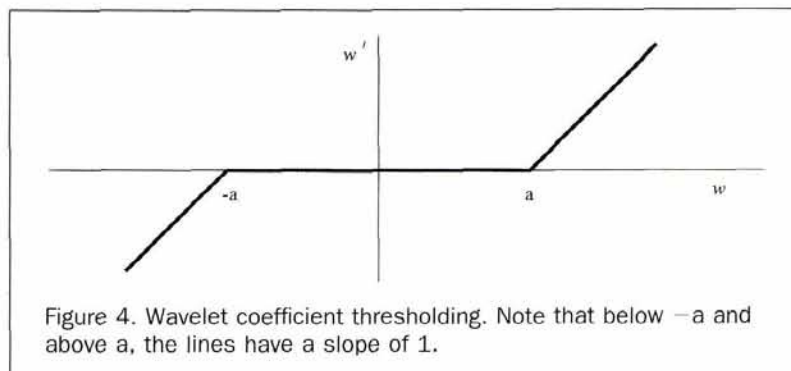


Figure 4. Wavelet coefficient thresholding. Note that below $-a$ and above a , the lines have a slope of 1.

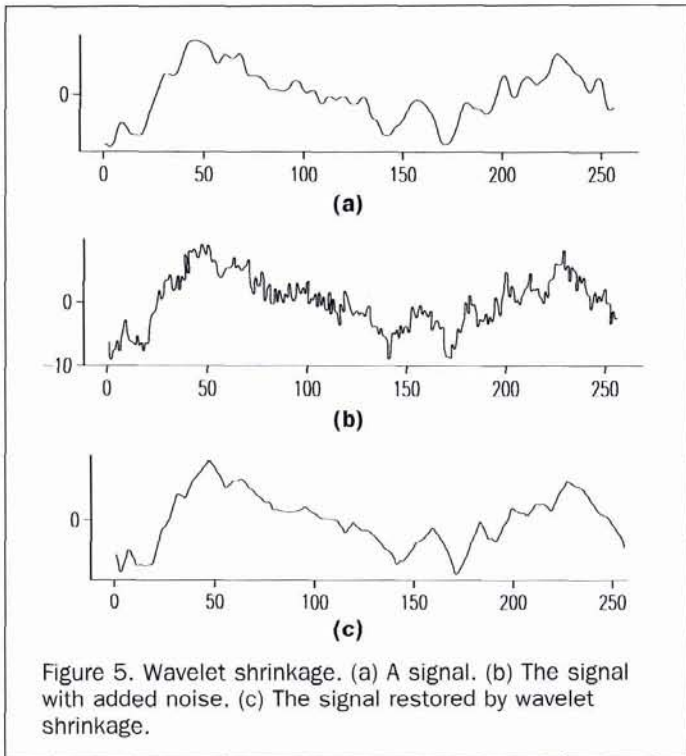


Figure 5. Wavelet shrinkage. (a) A signal. (b) The signal with added noise. (c) The signal restored by wavelet shrinkage.

A full description of the method may be found in Donoho and Johnstone (1994) and Donoho *et al.* (1995). Only the basic idea is presented here. The noisy signal is first decomposed into its wavelet representation. Then all wavelet coefficients finer than some resolution level (which we will denote j_0) are transformed: i.e.,

$$w' = \begin{cases} w + a & \text{if } w \leq -a \\ 0 & \text{if } -a < w < a \\ w - a & \text{if } w \geq a \end{cases}$$

where w is an original wavelet coefficient, w' is the transformed coefficient, and a is some threshold value. In effect, the transform pushes all coefficients towards zero by an amount a , setting them equal to zero if they are already within a of zero. This is illustrated in Figure 4.

The idea behind the transform is that major trends and discontinuities will contribute to large wavelet coefficients, while noise will only generate small coefficients. These small coefficients will be removed by the transform, while those corresponding to the features of interest will remain. An alternative transform, whereby coefficients between $-a$ and a are set to zero and others are left unchanged, has sometimes been used, but it may be shown that the shrinkage transform is a more effective estimator of the true (noise-free) wavelet coefficients under reasonable assumptions, such as their having a unimodal distribution symmetric about zero.

The use of wavelet shrinkage is illustrated in Figure 5. Figure 5a shows a signal sampled at 256 time points. Figure 5b shows the same signal with random $N(0, 1)$ noise added. Figure 5c shows the signal smoothed by wavelet shrinkage, with $j_0 = 4$ (where $j_0 = 0$ is the full signal and $j_0 = 8$ is the signal mean only) and $a = 3.3$. The Daubechies-6 wavelet was used. It is apparent that most of the noise has been removed. The restoration is not perfect — a few of the smaller bumps in Figure 5a have also been lost, and the peaks and troughs are sharper. This latter effect is a consequence of the shape of the Daubechies-6 function.

Image Filtering

Wavelet shrinkage can be used to smooth images. The approach is similar to that for one-dimensional signals. At each resolution level, we obtain the coefficients for the products of scaling and wavelet functions in the horizontal and vertical directions (i.e., four combinations). Shrinkage is applied to the coefficients involving a wavelet product.

Figure 6a shows a synthetic aperture radar (SAR) image of an area near Thetford Forest, England, obtained from the Maestro campaign (Ispra, 1989). The speckle in the image is intrinsic to SAR imaging, and its variance is known (Oliver, 1991; Horgan, 1994). Because the speckle contains little useful information, smoothing is desirable. Figures 6b, 6c, and 6d show the result of wavelet shrinkage, using, respectively, the Haar, Daubechies-4, and Daubechies-6 wavelets. A threshold of 28 was used (the pixel intensities ranging from 0 to 255) and $j_0 = 3$. This threshold was about half the minimax-optimal Donoho threshold (see below), which gave too strong a smoothing.

We can see that, in all the wavelet-shrunk images, the variability within the dominant field pattern has been smoothed, while edges and some other features which can be seen near the top left and bottom right corners have been preserved. Smoothing by the Daubechies wavelets appears to produce smoother and more satisfying results than does the Haar wavelet. In particular, the field boundaries are better preserved — they appear more jagged with the Haar wavelet. It seems to matter less whether the Daubechies-4 or Daubechies-6 wavelets are used. Some artifacts of the shape of the Daubechies functions are present. This would be the case with any other functional representation system also.

Practical Use of Wavelet Shrinkage

In using wavelet shrinkage as described here, three choices will be needed. First, which wavelet to use; second, what resolution level j_0 to smooth from; and third, what threshold a to use. Theoretical guidance is available on the last of these. Donoho *et al.* (1995) show that choosing $a = \sigma\sqrt{2 \log n}$ (where σ is the noise standard deviation and n is the number of time points) has some optimal properties, in that it minimizes the maximum reconstruction error that would be expected across a wide class of functions (the *minimax* criterion). However, it is not clear that such a criterion is the most suitable for practical use.

Some guidance on these issues can be obtained from a simulation study (Horgan, 1997). Although carried out with one-dimensional signals, it should be applicable to images also, because the way in which wavelets capture signal information and respond to noise is essentially the same regardless of dimension. Only the Haar and Daubechies wavelets shown in Figure 3 were used. The principal conclusions were

- The main determinant of the behavior of the wavelet shrinkage is whether there are rapid oscillations present in the signal (it is assumed that they are present in the noise). Even though discontinuities generate high frequency components (in the Fourier sense), a signal with discontinuities but no rapid oscillations behaves more like a slowly varying signal.
- For slowly varying signals, the Daubechies-4 wavelet gives the best results. For rapidly varying signals, the Daubechies-8 wavelet gives best results at low noise levels, the Daubechies-6 wavelet at higher levels. The Haar wavelet always gives inferior results. Perhaps the only reason for using it is that it is easy to program.
- The Donoho threshold is known to be optimal according to the minimax criterion. In practical applications, it may be felt that a criterion such as minimizing the mean square error or mean symmetric error (Marron and Tsybakov, 1995) is more appropriate. If so, a threshold less than the Donoho threshold

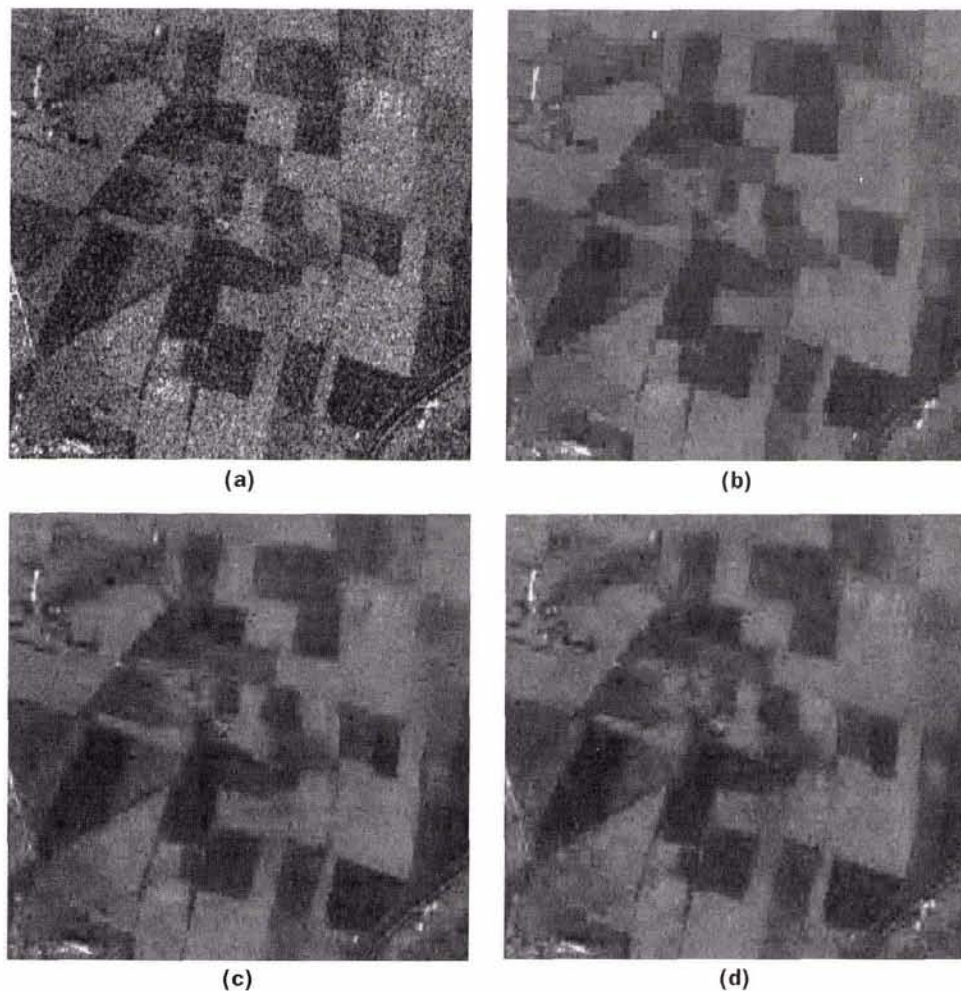


Figure 6. Image smoothing by wavelet shrinkage. (a) A SAR image. Smoothing by shrinkage is shown based on (b) the Haar, (c) the Daubechies-4, and (d) the Daubechies-6 wavelets.

a_D should be used, particularly for low noise levels. This will apply when a_D is based on a signal of length 256, and also on a wide range of other lengths because $\sqrt{\log n}$ varies very slowly.

- Shrinkage should be performed only at the finest resolution levels: 2 or 3, or perhaps 4, for slowly varying signals.

Another issue unrelated to these is that of dealing with images whose dimensions are not powers of two. The recursive nature of wavelet algorithms works best with power-of-two sizes, and some *ad hoc* adaptation is required for all other sizes. The easiest modification is to fill the image out to the next highest power of two. This can be done by filling the extra space with zeros, the image mean, the first or last rows or columns, reflection, or wrapping. In Figure 6, the images were 250 rows by 256 columns, and an extra six rows were created from the reflection of rows 245 to 250. These were removed after shrinkage.

For the SAR image used in our example, the noise standard deviation can be derived from theory to be 16.03 ($= 25\sqrt{\frac{\pi^2}{24}}$, for four-look-averaged log-transformed imagery scaled by 25 for display purposes). Although the theory is a simplification, examination of the pixel values within the

apparently homogeneous fields in Figure 6a showed good agreement. Autocorrelation in pixel values was also negligible. When the noise variance is not known, or theory needs confirmation, some means of estimating the noise variance is required. Donoho *et al.* (1995) suggest using the median absolute deviation (MAD) of the wavelet coefficients at the finest resolution divided by 0.645 (the MAD of the standard normal distribution). The rationale is that the wavelets at the finest resolution are mostly noise, and the few which are not will not much affect the robust estimation involved in the median. If the Haar wavelet is used, the finest wavelet coefficients are simply the differences between adjacent pixels divided by $\sqrt{2}$. If this estimation is applied to the SAR image in Figure 6, the estimated standard deviation is 16.44, close to the theoretical value.

Discussion

There have been many other methods proposed for smoothing SAR images. A review of many may be found in Oliver (1991). Simple linear filters (Glasbey and Horgan, 1995, Ch.3) are the most widely available, and have properties which are well understood. Their principal disadvantage is that edges are blurred and small but important features may be

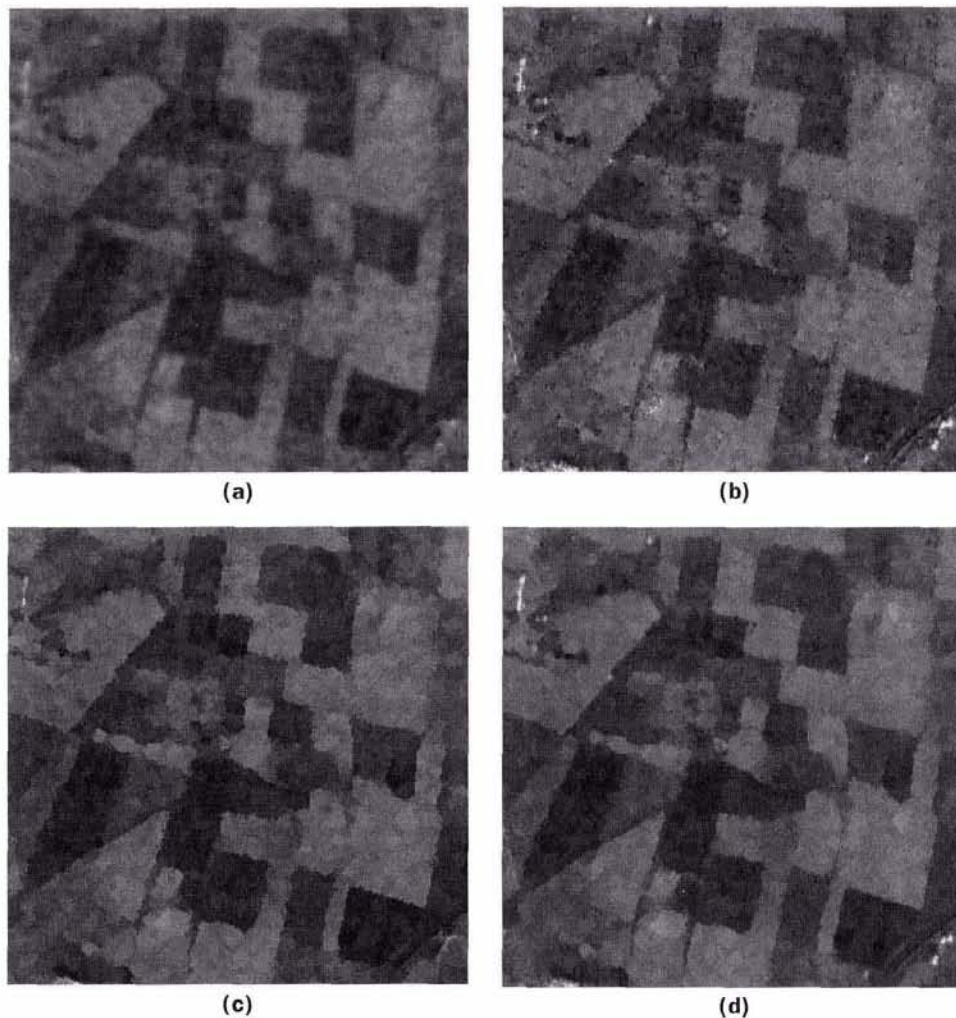


Figure 7. Alternative smoothing filters. (a) Moving median filter in a square 5 by 5 window. (b) Lee's filters in a square 5 by 5 window. (c) Tomita and Tsuji's filter using octagonal sub-windows of radius 2. (d) Modified version of (c) proposed by Glasbey and Jones.

lost. Filters such as the median (Pratt, 1978) avoid some of these problems, but for a more substantial improvement, filters need to be adaptive — they operate differently depending on local image properties. There tend to be two elements to such filters. One is that they inspect the amount of local variation in pixel intensities. Where this is highest, it assumes that there is an edge nearby and less or no smoothing is done than when local variation is lower. Examples of such filters are given by Lee (1981; 1983). The second idea is that, by examining several windows near a pixel, the one to which the pixel is most similar ("belongs" in the ground cover sense) may be selected, and smoothing done in that window. For an example, see Tomita and Tsuji (1977). Glasbey and Jones (1997) provide some fast algorithms for computing these filters. These ideas continue to develop. For recent work, see Martin and Turner (1993) and Smith (1996).

Figure 7 shows some of these filters applied to the SAR image. In the absence of an objective measure of ground truth, or how smooth the truth is, any comparison is subjective. The median filter, shown in Figure 7a preserves boundaries well and is widely available in image processing software. However, it is unable to preserve small details. The Lee filter, shown in Figure 7b does preserve small details, while smoothing homogeneous regions. It is somewhat simi-

lar to the Daubechies wavelet-shrunken images in Figures 6c and 6d. It appears more speckly, particularly near field borders. The Tomita and Tsuji filter, shown in Figure 7c, very effectively preserves field boundaries, but is poor at preserving small details. The modified version suggested by Glasbey and Jones (1997), shown in Figure 7d, is better in this regard, but still not as good as the Lee filter or Daubechies wavelet-shrunken images.

Beyond filtering, there is a large literature on segmentation — the assigning of pixels to discrete groupings. Pointers to the literature are given by Oliver (1991), Dobson *et al.* (1995), and Smith (1996).

These approaches can give good results. However, their adaptive nature, which is their strength in particular applications, means that they often need to be tailored to each image source, land-cover pattern, and application for which they are used. We have presented wavelets as an alternative approach to image smoothing. Although their use as described here is a sort of filter, and can be seen as adaptive in some ways, this adaptation is automatic, at least with regard to the resolution of the image. This derives from the nature of wavelets — they are handled in a similar way at all resolutions. The choice of threshold, a , may need more care.

It is our expectation that, as wavelets become more

widely understood and used, they will have an increasing impact on all aspects of remote sensing. In this paper, one application has been discussed. The results have been impressive, in that an effective smoothing which preserves features of importance has been achieved without the use of detailed knowledge of the nature of the underlying signal. The results appear fairly robust to changes in the parameters of the shrinkage algorithm, so careful fine-tuning is not needed. For image smoothing, these properties can only be considered as ideal.

Acknowledgment

This work was supported by the Scottish Office Agriculture, Environment and Fisheries Department.

References

Antoniadis, A., G. Gregoire, and I.W. McKeague, 1994. Wavelet methods for curve estimation, *Journal of the American Statistical Association*, 89:1340-1353.

Daubechies, I., 1988. Orthonormal bases of compactly supported wavelets, *Communications on Pure and Applied Mathematics*, 41:909-996.

———, 1992. *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics.

Dobson, M.C., F.T. Ulaby, and L.E. Pierce, 1995. Land-cover classification and estimation of terrain attributes using synthetic aperture radar, *Remote Sensing of Environment*, 51:99-214.

Donoho, D.L., and I.M. Johnstone, 1994. Ideal spatial adaptation by wavelet shrinkage, *Biometrika*, 81:425-455.

Donoho, D.L., I.M. Johnstone, G. Kerkycharian, and D. Picard, 1995. Wavelet shrinkage - asymptopia, *Journal of the Royal Statistical Society B*, 57:301-337.

Garguetduport, B., J. Girel, J.M. Chassery, and G. Pautou, 1996. The use of multiresolution analysis and wavelets transform for merging spot panchromatic and multispectral image data, *Photogrammetric Engineering & Remote Sensing*, 62:1057-1066.

Glasbey, C.A., and G.W. Horgan, 1995. *Image Analysis for the Biological Sciences*, Wiley, Chichester.

Glasbey, C.A., and R. Jones, 1997. Fast computation of moving average and related filters in octagonal windows, *Pattern Recognition Letters*, 18:555-565.

Hitchcock, D., 1994. A construction for 2-dimensional continuous wavelets, *BioSS Report* (BioSS, JCMB, King's Buildings, Mayfield Rd., Edinburgh EH9 3JZ, UK).

Horgan, G.W., 1994. Optimal Linear filters for SAR, *International Journal of Remote Sensing*, 15:1053-1064.

———, 1997. A simulation study of wavelet shrinkage, (submitted

for publication) and copies available at URL <http://www.bioss.sari.ac.uk/BioSS/Research/Image/remote.html>.

Ispra, 1989. *The Maestro-1 Campaign Report*, Report IRSA/MWT/4.90EU, Joint Research Centre, Ispra, Italy.

Kay, J., 1994. Wavelets, *Advances in Applied Statistics, 1994: Supplement to Vol 21 of Applied Statistics*, pp. 209-224.

Lee, J.S., 1981. Refined filtering of image noise using local statistics, *Computer Graphics and Image Processing*, 15:380-389.

———, 1983. A simple speckle smoothing algorithm for synthetic aperture radar images, *IEEE Transactions on Systems, Man and Cybernetics*, 13:85-89.

Li, D.R., and J.L. Shao, 1994. The wavelet and its application in image edge-detection, *ISPRS Journal of Photogrammetry and Remote Sensing*, 4:4-11.

Lindsay, R.W., D.B. Percival, and D.A. Rothrock, 1996. The discrete wavelet transform and the scale analysis of the surface-properties of sea-ice, *IEEE Transactions on Geoscience and Remote Sensing*, 34:771-787.

Mallat, S.G., 1989a. A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:674-693.

———, 1989b. Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbb{R})$, *Transactions of the American Mathematical Society*, 315:69-87.

Marron, J.S., and A.B. Tsybakov, 1995. Visual error criteria for qualitative smoothing, *Journal of the American Statistical Association*, 90:499-507.

Martin, F.J., and R.W. Turner, 1993. SAR speckle reduction by weighted filtering, *International Journal of Remote Sensing*, 14:1759-1774.

Nagao, M., and T. Matsuyama, 1979. Edge preserving smoothing, *Computer Graphics and Image Processing*, 9:394-407.

Oliver, C.J., 1991. Information from SAR images, *Journal of Physics D: Applied Physics*, 24:1493-1514.

Pratt, W.K., 1978. *Digital Image Processing*, Wiley, New York.

Press, W.H., B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, 1993. *Numerical Recipes in C*, Cambridge University Press.

Smith, D.M., 1996. Speckle reduction and segmentation of synthetic-aperture radar images, *International Journal of Remote Sensing*, 17:2043-2057.

Strang, G., 1989. Wavelets and dilation equations: a brief introduction, *SIAM Review*, 31:614-627.

Tomita, F., and S. Tsuji, 1977. Extraction of multiple regions by smoothing in selected neighborhoods, *IEEE Transactions on Systems, Man and Cybernetics*, 7:107-109.

Werness, S.A., S.C. Wei, and R. Carpinella, 1994. Experiments with wavelets for compression of SAR data, *IEEE Transactions Geoscience and Remote Sensing*, 32:197-201.

(Received 18 July 1997; accepted 24 February 1998; revised 19 March 1998)

Certification Seals & Stamps

- Now that you are certified as a remote sensor, photogrammetrist or GIS/LIS mapping scientist and you have that certificate on the wall, make sure everyone knows!
- An embossing seal or rubber stamp adds a certified finishing touch to your professional product.
- You can't carry around your certificate, but your seal or stamp fits in your pocket or briefcase.
- To place your order, fill out the necessary mailing and certification information. Cost is just \$35 for a stamp and \$45 for a seal; these prices include shipping and handling. *Please allow 3-4 weeks for delivery.*

SEND COMPLETED FORM WITH YOUR PAYMENT TO:

ASPRS Certification Seals & Stamps, 5410 Grosvenor Lane, Suite 210, Bethesda, MD 20814-2160

NAME: _____	PHONE: _____
CERTIFICATION #: _____	DATE: _____
ADDRESS: _____	
CITY: _____	STATE: _____
POSTAL CODE: _____	COUNTRY: _____
PLEASE SEND ME: <input type="checkbox"/> Embossing Seal \$45 <input type="checkbox"/> Rubber Stamp \$35	
METHOD OF PAYMENT: <input type="checkbox"/> Check <input type="checkbox"/> Visa <input type="checkbox"/> MasterCard	
CREDIT CARD ACCOUNT NUMBER _____	EXPIRES _____
SIGNATURE _____	DATE _____