Phase Preserving Denoising of Images

Peter Kovesi
Department of Computer Science
The University of Western Australia
Nedlands, W.A. 6907
pk@cs.uwa.edu.au

Abstract

In recent years wavelet shrinkage denoising has become the method of choice for the denoising of images. However, despite much research a number of questions remain.
- Which of the many wavelets that exist should one use?
- How should the threshold be set? and
- How are features in the image affected by the thresholding operation?
This paper explores these issues and argues for the use of non-orthogonal, complex valued, log-Gabor wavelets, rather than the more usual orthogonal or bi-orthogonal wavelets. Thresholding of wavelet responses in the complex domain allows one to ensure that perceptually important phase information in the image is not corrupted. It is also shown how appropriate threshold values can be determined automatically from the statistics of the wavelet responses to the image.

1. Introduction

Denoising of images is typically done with the following process: The image is transformed into some domain where the noise component is more easily identi£ed, a thresholding operation is then applied to remove the noise, and £nally the transformation is inverted to reconstruct a (hopefully) noise-free image.

The wavelet transform has proved to be very successful in making signal and noise components of the signal distinct. As wavelets have compact support the wavelet coef£cients resulting from the signal are *localised*, whereas the coef£cients resulting from noise in the signal are *distributed*. Thus the energy from the signal is directed into a limited number of coef£cients which 'stand out' from the noise. Wavelet shrinkage denoising then consists of identifying the magnitude of wavelet coef£cients one can expect from the noise (the threshold), and then shrinking the magnitudes of all the coef£cients by this amount. What remains of the coef£cients should be valid signal data, and the transform can then be inverted to reconstruct an estimate of the signal [4, 3, 1].

Wavelet denoising has concentrated on the use of orthogonal or bi-orthogonal wavelets because of their reconstructive qualities. However, no particular wavelet has been identi£ed as being the 'best' for denoising. It is generally agreed that wavelets having a linear-phase, or near linear-phase, response are desirable, and this has led to the use of the 'symlet' series of wavelets and bi-orthogonal wavelets.

A problem with wavelet shrinkage denoising is that the discrete wavelet transform is not translation invariant. If the signal is displaced by one data point the wavelet coef£cients do not simply move by the same amount. They are completely different because there is no redundancy in the wavelet representation. Thus, the shape of the reconstructed signal after wavelet shrinkage and transform inversion will depend on the translation of the signal - clearly this is not very satisfactory. To overcome this translation invariant denoising has been devised [1]. This involves averaging the wavelet shrinkage denoising result over all possible translations of the signal. This produces very pleasing results and overcomes pseudo-Gibbs phenomena that is often seen in the basic wavelet shrinkage denoising scheme.

The criteria for quality of the reconstructed noise-free image has generally been the RMS error - though Donoho suggests a side condition that the reconstructed (denoised) signal should be, with high probability, as least as smooth as the original (noise free) signal.

While the use of the RMS error in reconstructing 1D signals may be reasonable, the use of the RMS measure for image comparison has been criticised [2, 10]. Almost without exception images exist solely for the bene£t of the human visual system. Therefore any metric that is used for evaluating the quality of image reconstruction must have relevance to our visual perception system. The RMS error certainly does not necessarily give a good guide to the perceptual quality of an image reconstruction. For example, displacing an image a small amount, or offsetting grey levels by a small amount, will have negligible perceptual effect, but will induce a large RMS error.

As yet no metric that matches human visual perception has been devised. However, one quantity that appears to be very important in the human perception of images is phase. The classic demonstration of the importance of phase was devised by Oppenheim and Lim [9]. They took the Fourier transforms of two images and used the phase information from one image and the magnitude information of the other to construct a new, synthetic Fourier transform which was then back-transformed to produce a new image. The features seen in such an image, while somewhat scrambled, clearly correspond to those in the image from which the phase data was obtained. Little evidence, if any, from the other image can be perceived. A demonstration of this is repeated here in Figure 1.

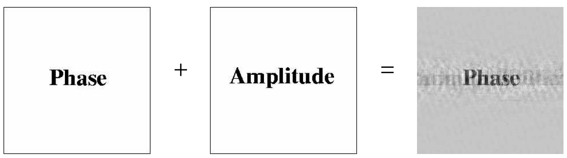


Figure 1. When phase information from one image is combined with magnitude information of another it is phase information that prevails.

While phase is not the only quantity important to our perception of images it would seem that an important constraint that should be satis£ed by any image enhancement process, such as denoising, is that it should not corrupt the phase information in an image

2. Phase Preserving Denoising

To be able to preserve the phase data in an image we have to £rst extract the local phase and amplitude information at each point in the image. This can be done by applying (a discrete implementation of) the continuous wavelet transform and using wavelets that are in symmetric/anti-symmetric pairs. Here we follow the approach of Morlet, that is, using wavelets based on complex valued Gabor functions - sine and cosine waves, each modulated by a Gaussian [8]. Using two £lters in quadrature enables one to calculate the amplitude and phase of the signal for a particular scale/frequency at a given spatial location.

However, rather than using Gabor £lters we prefer to use *log Gabor* functions as suggested by Field [5]; these are £lters having a Gaussian transfer function when viewed on the logarithmic frequency scale. Log Gabor £lters allow arbitrarily large bandwidth £lters to be constructed while still maintaining a zero DC component in the even-symmetric £lter. A zero DC value cannot be maintained in Gabor functions for bandwidths over 1 octave. It is of interest to note that the spatial extent of log Gabor £lters appears to be minimized when they are constructed with a bandwidth of approximately two octaves [7, 6]. This would appear to be optimal for denoising as this will minimise the spatial spread of wavelet response to signal features, and hence concentrate as much signal energy as possible into a limited number of coef£cients.

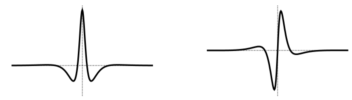


Figure 2. Even and odd log Gabor wavelets, each having a bandwidth of two octaves.

Analysis of a signal is done by convolving the signal with each of the quadrature pairs of wavelets. If we let I denote the signal and M_n^e and M_n^o denote the even-symmetric and odd-symmetric wavelets at a scale n we can think of the responses of each quadrature pair of £lters as forming a response vector,

$$[\, e_n(x), o_n(x) \,] = [\, I(x) * M_n^e, I(x) * M_n^o \,]\,.$$

The values $e_n(x)$ and $o_n(x)$ can be thought of as real and imaginary parts of complex valued frequency component. The amplitude of the transform at a given wavelet scale is given by

$$A_n(x) = \sqrt{e_n(x)^2 + o_n(x)^2}$$

and the phase is given by

$$\Phi_n(x) = atan2(o_n(x), e_n(x)).$$

At each point x in a signal we will have an array of these response vectors, one vector for each scale of £lter. These response vectors form the basis of our localized representation of the signal as shown in Figure 3.

In this domain the denoising process consists of determining a noise threshold at each scale and shrinking the magnitudes of the £lter response vectors appropriately, while leaving the phase unchanged.

It should be noted that shrinkage of complex-valued wavelet response vectors is *not* the same as shrinkage of real-valued discrete wavelet responses. In the proposed phase preserving scheme some component of the even £lter response will always be retained (even if it is very small) as long as the odd £lter response is such that the total amplitude exceeds the noise threshold. This should be compared to the case for real-valued discrete wavelets where a response will only be retained if the magnitude exceeds

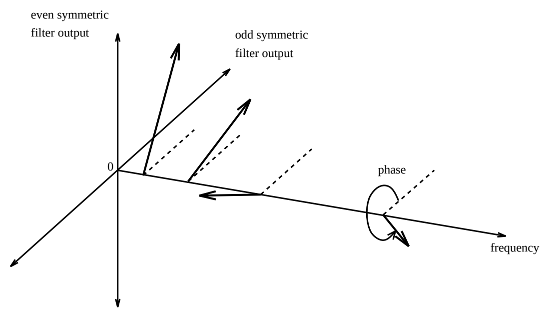


Figure 3. An array of £lter response vectors at a point in a signal can be represented as a series of vectors radiating out from the frequency axis. The amplitude speci£es the length of each vector and the phase speci£es its angle. Note that wavelet £lters are scaled geometrically, hence their centre frequencies vary accordingly.

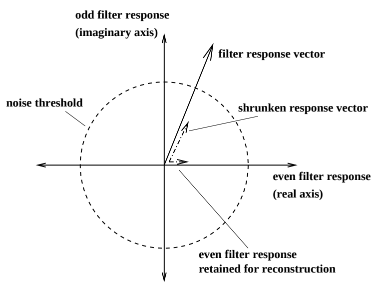


Figure 4. View along the frequency axis illustrating the shrinkage of complex-valued wavelet response vectors.

the threshold because the shrinkage process is constrained to operate only on the real axis. Note also that applying shrinkage only along the real axis will corrupt phase information as the imaginary component will be ignored in deciding how much a wavelet component should be shrunk. It is worth noting that the averaging process in translation-invariant denoising may achieve a similar result to the proposed phase preserving algorithm.

Having shrunk the complex-valued wavelet response vectors an estimate of the signal can then be reconstructed by summing the remaining even-symmetric £lter responses over all scales and orientations. However, there are some issues in the reconstruction of the denoised image from the shrunk response vectors. Complex valued log Gabor £lters do not form an orthogonal basis set. This means that the signal can only be reconstructed over the range of frequencies covered by the £lters, and that the signal can only be reconstructed up to a scale factor. Thus to achieve satisfactory reconstruction the design of the wavelet £lter bank must be such that the transfer functions of all the £lters overlap suf£ciently so that their sum results in an even coverage of the spectrum. In the 2D frequency plane the £lter transfer functions appear as 2D log Gaussians. These can be arranged in a 'rosette' to ensure uniform coverage of the spectrum. Under this arrangement it is dif£cult to have £lters that cover the very low frequencies in the image. However, perceptually this does not appear to be very important. Similarly, the lack of an absolute scale in the reconstructed grey levels is not important perceptually.

3. Determining the Threshold

The most crucial parameter in the denoising process is the threshold. While many techniques have been developed [4, 3] none have proved very satisfactory. Here we develop an automatic thresholding scheme.

First we must look at the expected response of the £lters to a pure noise signal. If the signal is purely Gaussian white noise the positions of the resulting response vectors from a wavelet quadrature pair of £lters at some scale will form a 2D Gaussian distribution in the complex plane. What we are interested in is the distribution of the *magnitude* of the response vectors. This will be a Rayleigh distribution

$$R(x) = \frac{x}{\sigma_g^2} e^{\frac{-x^2}{2\sigma_g^2}} ,$$

where σ_g^2 is the variance of the 2D Gaussian distribution describing the position of the £lter response vectors.

The mean of the Rayleigh distribution is given by

$$\mu_r = \sigma_g \sqrt{\frac{\pi}{2}} ,$$

and the variance is

$$\sigma_r^2 = \frac{4-\pi}{2} \sigma_g^2 .$$

The point to note is that only one parameter is required to describe the distribution; given μ_r one can determine σ_r, and vice-versa. If we can determine the noise response distribution at each £lter scale we could then set the noise shrinkage threshold at each scale to be some number of standard deviations beyond the mean of the distribution

$$T = \mu_r + k\sigma_r ,$$

where k is typically in the range 2 – 3.

How can we determine the noise amplitude distribution? The smallest scale £lter has the largest bandwidth, and as such will give the strongest noise response. Only at feature points will the response differ from the background noise

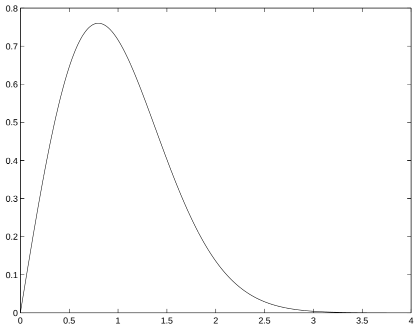


Figure 5. Rayleigh distribution with a mean of one.

response, but the regions where it will be responding to features will be small due to the small spatial extent of the £lter. Thus the smallest scale wavelet quadrature pair will spend most of their time only responding to noise.

Thus, the distribution of the amplitude response from the smallest scale £lter pair across the whole image will be primarily the noise distribution, that is, a Rayleigh distribution with some contamination as a result of the response of the £lters to feature points in the image.

We can obtain a robust estimate of the mean of the amplitude response of the smallest scale £lter via the median response. The median of a Rayleigh distribution is the value x such that

$$\int_0^x \frac{x}{\sigma_g^2} e^{\frac{-x^2}{2\sigma_g}} = \frac{1}{2}$$

$$\Rightarrow \text{median} = \sigma_g \sqrt{-2\ln(1/2)}\,.$$

Noting that the mean of the Rayleigh distribution is $\sigma_g \sqrt{\frac{\pi}{2}}$ we obtain the expected value of the amplitude response of the smallest scale £lter (the estimate of the mean)

$$\begin{aligned} \mathbb{E}(A_N) &= \frac{\sigma_g \sqrt{\pi/2}}{\sigma_g \sqrt{-2\ln(1/2)}}\,.\,\text{median} \\ &= \frac{1}{2}\sqrt{\frac{-\pi}{ln(1/2)}}\,.\,\text{median}\,, \end{aligned}$$

where N is the index of the smallest scale £lter. Given that $\sigma_g = \frac{\mathbb{E}(A_N)}{\sqrt{\pi/2}}$ we can then estimate μ_r and σ_r for the noise response for the smallest scale £lter pair, and hence the shrinkage threshold.

We can estimate the appropriate shrinkage thresholds to use at the other £lter scales if we make the following observation: If it is assumed that the noise spectrum is uniform then the wavelets will gather energy from the noise as a function of their bandwidth which, in turn, is a function of their centre frequency. For 1D signals the amplitude response will be proportional to the square root of the £lter centre frequency. In 2D images the amplitude response will be directly proportional to the £lter centre frequency.

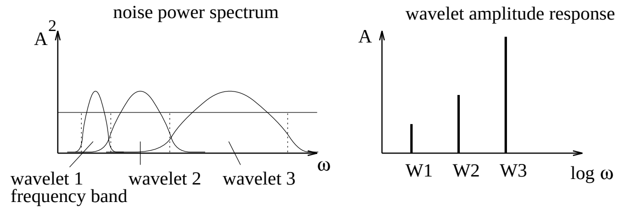


Figure 6. If the noise spectrum is uniform the response of a wavelet to the noise will be a function of its bandwidth.

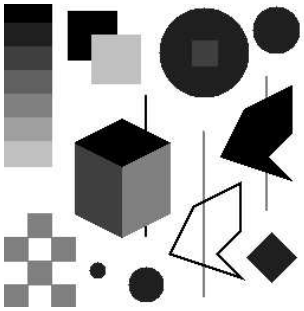
Thus having obtained an estimate of the noise amplitude distribution for the smallest scale £lter pair we can simply scale this appropriately to form estimates of the noise amplitude distributions at all the other scales. This approach proves to be very successful in allowing shrinkage thresholds to be set automatically from the statistics of the smallest scale £lter response over the whole image.

4. Results

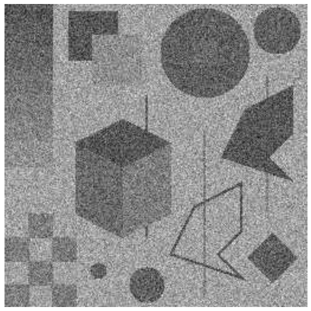
Figure 7 shows a synthetic test image with grey values ranging between 0 and 255. Gaussian white noise with a standard deviation of 80 grey levels was added to the image. The result of applying the phase preserving denoising algorithm to the image (using a k value of 2 to set the threshold) is shown along with the result obtained by applying a standard discrete wavelet denoising scheme (the MATLAB `wdencmp` function using the ‘symlet8’ wavelet and a manually derived threshold of 60).

Figure 8 shows the 1D sections at row 150 (out of 256) on each of the four images shown in Figure 7. Note the vertical scale for the plot of the phase preserved denoised image does not match that for the original image. The reconstruction from the complex-valued log Gabor wavelets cannot cover the very low, and zero frequency, components of the signal. Also the signal can only be recovered up to a scale factor. Despite this the *shape* of the reconstructed signal is very good. A major part of the success of the seemingly astonishing reconstruction is due to the fact that the denoising process is taking place in 2D. The reconstruction of row 150 in the image makes use of information from above and below that row. Such a result would not be possible working solely in 1D.

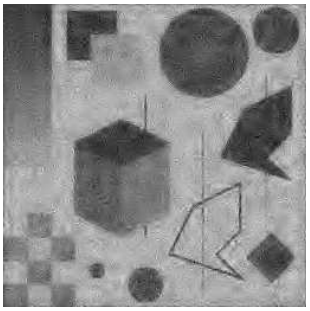
Figure 9 shows the phase preserving denoising process applied to a poor quality surveillance image of a hold-up. It should be noted that video images consist of two interlaced images. If there is any motion (there was a small amount in this image) the interlacing will result in ‘tooth comb’ edges



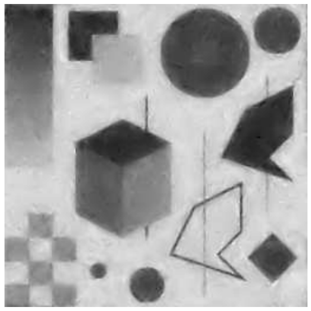
Test image



Noisy test image



Denoised using symlet8 wavelet



Denoised with phase preserved

Figure 7. Denoising of a test image

around objects. To overcome this the individual images that make up the video frame can be obtained by extracting just the even, or just the odd, numbered scan lines from the image prior to denoising.

5. Conclusion

We have presented a new denoising algorithm, based on the decomposition of a signal using complex-valued wavelets. This algorithm preserves the perceptually important phase information in the signal. In conjunction with this a method has been devised to automatically determine the appropriate wavelet shrinkage thresholds from the statistics of the amplitude response of the smallest scale £lter pair over the image. The automatic determination of thresholds overcomes a problem that has plagued wavelet denoising schemes in the past.

The RMS measure is not always the most appropriate metric to use in the development of image processing algorithms. Indeed it could be argued that more time should be spent optimising the choice of the optimisation criteria in general. For images it would appear that the preservation of phase data is important, though of course, other factors must also be important. The denoising algorithm presented here does not seek to do any optimisation, it has merely been constructed so as to satisfy the *constraint* that phase should not be corrupted. Given that it satis£es this constraint, it should be possible to develop it further so that it does incorporate some optimisation, say, the minimisation of the distortion of the signal's amplitude spectrum. What should also be investigated is the possible relationship between this phase preserving algorithm and translation invariant denoising.

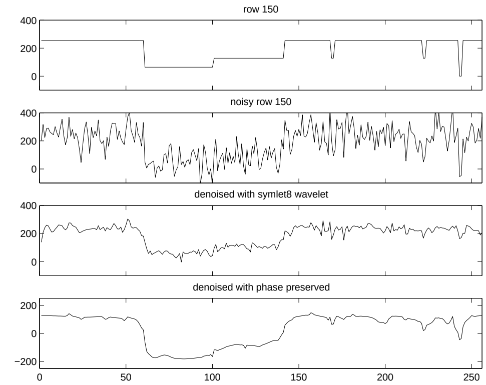


Figure 8. Section along row 150 in the test image

References

[1] R. R. Coifman and D. Donoho. Time-invariant wavelet denoising. In A. Antoniadis and G. Oppenheim, editors, *Wavelets and Statistics*, volume 103 of *Lecture Notes in Statistics*, pages 125–150, New York, 1995. Springer-Verlag.

[2] S. Daly. The visible diferences predictor: An algorithm for the assessment of image £delity. In A. Watson, editor, *Digital Images and Human Vision*, pages 179–206. MIT Press, 1993.

[3] D. L. Donoho. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627, 1995.

[4] D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.

[5] D. J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of The Optical Society of America A*, 4(12):2379–2394, December 1987.

[6] P. D. Kovesi. Image features from phase congruency. *Videre: Journal of Computer Vision Research*, to appear. http://mitpress.mit.edu/e-journals/Videre/.

[7] P. D. Kovesi. *Invariant Measures of Image Features From Phase Information*. PhD thesis, The University of Western Australia, May 1996.

Original surveillance image

Grey scale enhanced surveillance image

Image obtained from just the odd numbered scan lines

Denoised with phase preserved

Figure 9. Denoising of a surveillance image

[8] J. Morlet, G. Arens, E. Fourgeau, and D. Giard. Wave propagation and sampling theory - Part II: Sampling theory and complex waves. *Geophysics*, 47(2):222–236, February 1982.

[9] A. V. Oppenheim and J. S. Lim. The importance of phase in signals. In *Proceedings of The IEEE 69*, pages 529–541, 1981.

[10] D. Wilson, A. Baddeley, and R. Owens. A new metric for grey-scale image comparison. *International Journal of Computer Vision*, 24(1):5–17, 1997.