MAXIMUM LIKELIHOOD FOR BAYESIAN ESTIMATOR BASED ON α-STABLE FOR IMAGE

X.Huang[1] and A.C.Madoc[2]
[1]School of Electronics and Telecommunications Engineering, University of Canberra, ACT 2601, Australia
[2]Center for Actuarial Studies, Department of Economics, University of Melbourne, VIC 3051, Australia

ABSTRACT

A maximum likelihood for Bayesian estimator based on α-stable is discussed. Closer to a realistic situation, and unlike previous methods used for Bayesian estimator, for the case discussed here it is not necessary to know the variance of the noise. The parameters relative to Bayesian estimators of the model built up are carefully investigated after a discussion of α-stable 3-D simulations for a maximum likelihood. The Bayesian estimator then is established. As an example, an improved Bayesian estimator that is a natural extension of the Wiener solution and other wavelet denoising (soft and hard threshold methods), is presented to illustrate our discussion.

1. INTRODUCTION

Wavelet transform as a powerful tool for recovering signals from noise has been of considerably interest [1-7]. In fact, wavelet theory combines many existing concepts into a global framework and hence becomes a powerful tool for several domains of application.

Those methods that compute the correlation between coefficients at successive scales are based on the assumption that regular signal features show correlated coefficients at different scales, whereas irregularities due to noise do not [2].

As mentioned by Achim *et al.* [8], there are two major drawbacks for thresholding. One is that choice of the threshold is always done in an ad hoc manner; another is that the specific distributions of the signal and noise may not be well matched at different scales.

Donoho gives some minimum thresholds for several threshold schemes, titled "universal thresholds". These explicitly depend on the standard deviation of noise, where the standard deviation is assumed to be known. In practice, the standard deviation can be readily estimated using the methods discussed in [3], [9]. For some applications the optimal threshold can be computed. An approach different from "universal thresholds" is presented by Weyrich and Nason [10], in which cross-validation is used. Two approaches to cross validation are used, namely ordinary cross validation (OCV) and generalised cross validation (GCV): each is used to minimize the least-squares error between the original (which is the unknown value) function and its estimate based on the noisy observation.

Several groups working on multispectral image restoration [11], multichannel image restoration [12], and multiframe image restoration [13], have explored the Wiener filter. It is well known that a classical solution that deals with the noise removal problem is the Wiener filter [14], [15]. However this method, designed mainly for additive noise suppression, has its limitations.

Modelling the statistics of natural images is a challenging task because of the high dimensionality of the signal and the complexity of statistical structures that are prevalent. Numerous papers discuss modelling the statistics of nature images, including Bayesian processing presuppose proper modelling for the prior probability density function of the signal [8, 14, 15,16, 17, 18]. In this paper the maximum likelihood for Bayesian estimator based on α-stable for image is discussed, and a method is offered such that it removes noise images without information about the noise-variance, even though this is one of key-parameters for normal Bayesian estimators.

2. ALPHA-STABLE DISTRIBUTIONS AND MAXIMUM LIKELIHOOD

Tsakalides *et al.* [19], Achim *et al.* [8] and Peterson et al. [20] recently showed that alpha-stable distributions, a family of heavy tailed densities, are sufficiently flexible and rich to appropriately model wavelet coefficients of images in denoising applications. In this section we shall investigate maximum likelihood for an alpha-stable distribution after briefly introduce symmetric alpha-stable distribution ($S\alpha S$).

It is well known that the $S\alpha S$ distribution is defined by its characteristic function:

$$\phi(\omega) = \exp(j\delta\omega - \gamma|\omega|^{\alpha}), \tag{1}$$

The parameters α, γ, and δ describe completely a $S\alpha S$ distribution. The characteristic exponent α controls the heaviness of the tails of the stable density. α can take values in (0,2]; while α = 1 and 2 are the Cauchy and Gaussian cases. There is not closed-form expression known for the general $S\alpha S$ probability density function (PDF). The dispersion parameter γ (γ>0) refers to the spread of the PDF. The location parameter δ is analogous to the mean of the PDF. Figure 1 shows the PDF for alpha-stable with different alpha, 1.0 to 2.0 when gamma takes 0.3. For comparison with the Laplace distribution, which was described in [18] with s =1.5 and p = 0.8 is included in Figure 1. It is clearly shown that the smaller α value implies heavier tails and Laplace distribution is significantly different from that of alpha-stable distributions.

It is well known that if it is possible to minimize the mean square error (MSE), which is one of our major targets for denoising, when the bias is zero, and then the variance is also minimized. Such estimators are called "minimum variance unbiased" estimators, and they attain an important minimum bound on the variance of the estimator, called minimum variance bound.

0-7803-7304-9/02/$17.00 C2002 IEEE

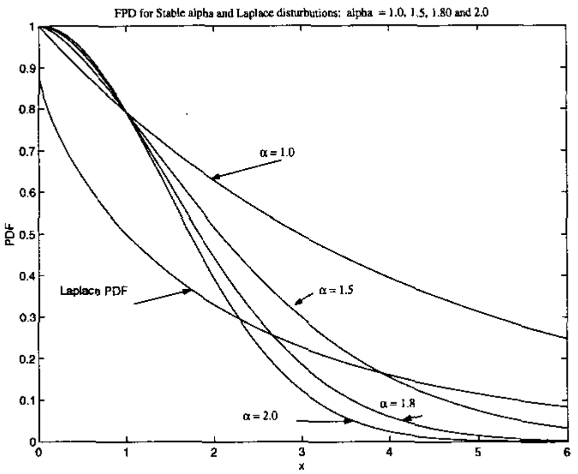

Figure 1: PDF for alpha-stable with (from top to down α = 1.0, 1.5 1.8, and 2.0 and Laplace distribution.

If a variable $\hat{\theta}$ is unbiased it follows that

$$E(\hat{\theta}-\theta)=0 \qquad (2)$$

which can be expressed as:

$$\int_{-\infty}^{\infty}\cdots\int(\hat{\theta}-\theta)\, f_{\mathbf{x};\theta}(\mathbf{x};\theta)d\mathbf{x}=0 \qquad (3)$$

where $\vec{x}(\xi)=[x_1(\xi),x_2(\xi),\dots,x_N(\xi)]^T$ and $f_{\vec{x};\theta}(\vec{x};\theta)$ is the joint density of $\vec{x}(\xi)$, which depends on a fixed but unknown parameter

Following [18] we have

$$\mathrm{var}(\hat{\theta})\geq-\frac{1}{E\{\partial^2\ln f_{\mathbf{x};\theta}(\mathbf{x};\theta)/\partial\theta^2\}} \qquad (4)$$

The function $\ln f_{\vec{x};\theta}(\vec{x};\theta)$ is well known as the *"log likelihood"* function of θ (LLF). Its maximum likelihood estimate can be obtained from the equation:

$$\frac{\partial\ln f_{\vec{x};\theta}(\vec{x};\theta)}{\partial\theta}=0 \qquad (5)$$

The first order of differential log likelihood function with respect to θ is called the maximum likelihood (ML) estimate. If the efficient estimate does not exist, then the ML estimate will not achieve the lower bound and hence it is difficult to ascertain how closely the variance of any estimate will approach the bound.

Figure 2 show that the LLF simulation results for the alpha-stable with α = (0 ,4.5], x =[-6, -6], with γ = 0.1. Its 2nd order LLF distributions are shown in Figure 3, which shows the view from the α axis of Figure 2, from which we can see why generally the value of α has been taken in the range [1,2]. The value about 1.5 is strongly recommended if there is no information about α.

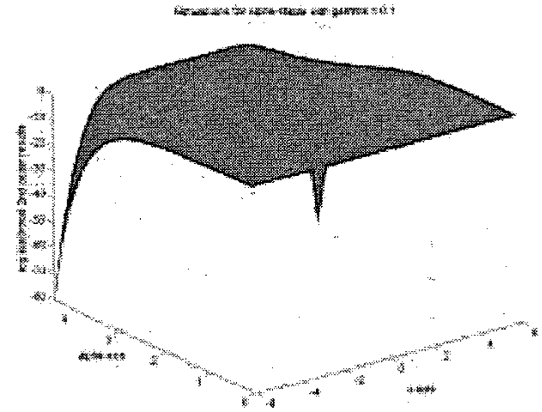

Figure 2: The 2nd order simulations of the LLF for an alpha-stable with g = 0.1.

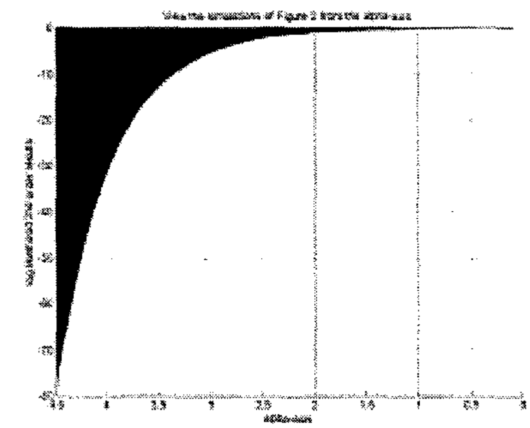

Figure 3: The view of the results of Figure 2 from the front of the alpha-axis.

3. BAYESIAN ESTIMATOR

If we take the probability density of θ as $p(\theta)$; and the posterior density function as $f(\theta\mid x_1,\dots,x_n)$, then the updated probability density function of θ is as follows:

$$f(\theta\mid x_1,\dots,x_n)=\frac{f(\theta,x_1,\dots,x_n)}{f(x_1\dots,x_n)} =\frac{p(\theta)\,f(x_1,\dots,x_n\mid\theta)}{\int f(x_1,\dots,x_n\mid\theta)\,p(\theta)\,d\theta} \qquad (6)$$

If we estimate the parameters of the prior distributions of the signal *s* and noise *q* components of the wavelet coefficients c, we may use the parameters to form the prior PDFs of $P_s(s)$ and $P_q(q)$, hence the input/output relationship can be established by the Bayesian estimator, namely, let input/output of the Bayesian estimator =*BE*, we have:

$$BE = \frac{\int P_q(q) P_s(s) s ds}{\int P_q(q) P_s(s) ds} \quad (7)$$

Figure 4 shows how γ affect the *BE* and Figure 5 shows how σ affects the *BE*.

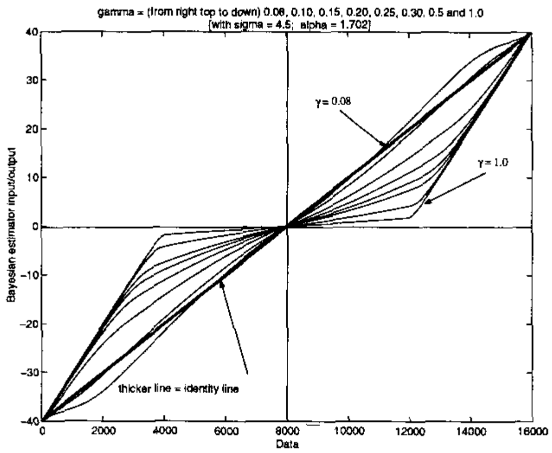

Figure 4: How γ affects *BE*. Here γ ranges from 0.08 to 1.0, with $\sigma = 4.5$, $\alpha = 1.702$.

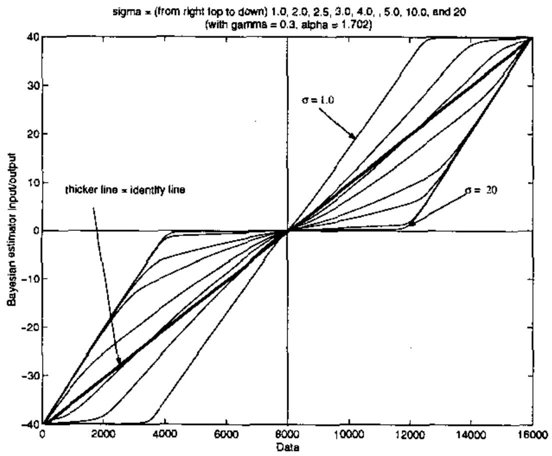

Figure 5: How σ affects *BE*. Here σ ranges from 1.0 to 10, with $\gamma = 0.3$, $\alpha = 1.702$.

As we mentioned above for the Bayesian estimators the three parameters are important and can affect the *BE* strongly. Achim *et al.* [8] tried to take γ/σ as one parameter and showed a simulation by keeping γ/σ constant, but the fact is that even though the ratio is the same the *BE* can be significantly different. However, it is noted that the σ is generally estimated and we may even have no information about it. However, as in section 2 we may take $\alpha = 1.5$ as fist estimate. Hence we can obtain the *BE* from the Figure 6, where γ/σ is taken from 0.1, to 0.8. It is clearly seen that when ratio equals to 0.8 the BE becomes the similar to the case of the hard threshold presented by Donoho. It is to be noticed that the identity line can be a reference for individual *BE* curve.

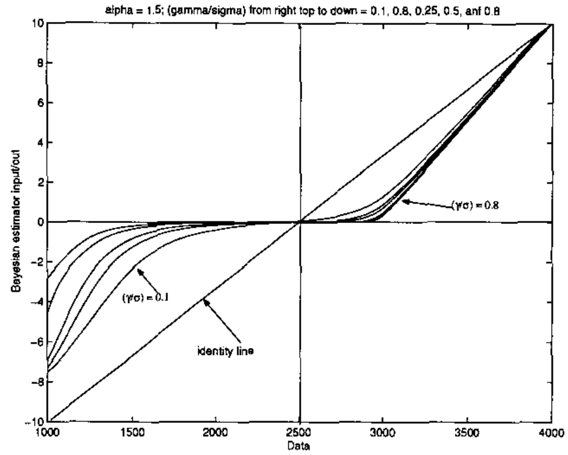

Figure 6: Individual BE with ratio of (γ/σ) equals to 0.1, 0.2, 0.25, 0.5 and 0.8 and $\alpha = 1.5$ taken from section 2. The BE with 0.8 is similar to "hard threshold" case.

4. SOME EXAMPLES

When we make measurements, we have no information about the noise value of the image we obtained. The only information one may have is an experience in judgement of the noise level, which becomes the outline of the denoising strategy.

Figure 7 shows the denoising with Haar wavelet in 2 levels.

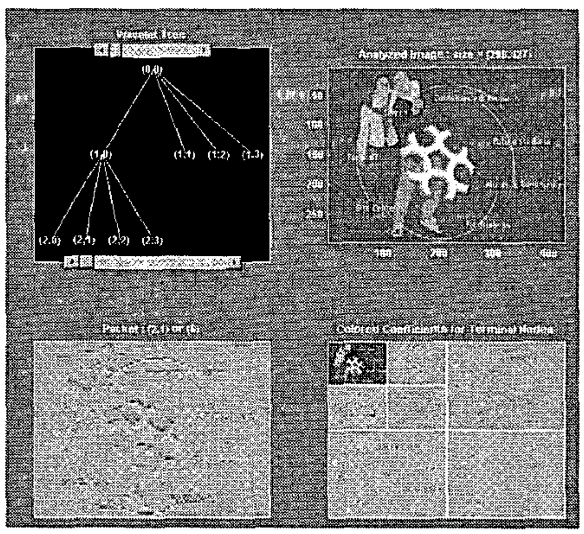

Figure 7: Denoising image of the web page of "University of Canberra" by Haar wavelet in 2 levels.

Figure 8 shows the noisy image. If we did not have any information about the noisy image, we can, as mentioned above, take $\alpha = 1.5$ and $(\gamma/\sigma) = 0.5$ and the denoised image is in Figure 9. The results by other methods, together with the Bayesian estimator for α-stable method, are given in the table 1.

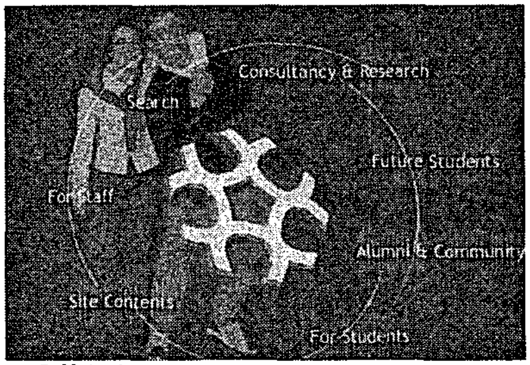

Figure 8: Noisy image

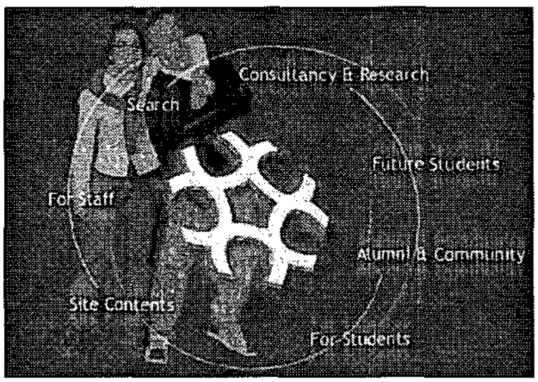

Figure 9: The denoised image by Bayesian estimator based on α-stable distribution.

Comparisons of other denoising results are in table 1.

Method	1	2	3	4
S/MSE	12.21	12.30	12.01	12.89

Table 1: Comparison of denoising results with BE in signal to mean square error (S/MSE) in dB. Here 1= soft thresholding; 2 = Hard thresholding; 3 = Homomorphic Wiener; 4 = Bayesian estimator based on α-stable

4. CONCLUSION

A new technique for denoising an image has been developed. The technique uses the statistician's Bayesian estimator theory to simplify the selection of parameters, and in some situations it provides more precise images than other methods.

5. REFERENCES

1. D.L. Donoho and I.M. Johnstone, "Wavelet shrinkage: Asymptopia?", Journal of the Royal Statistical Society, Series B, vol. 57, 301-369, 1995.
2. Y. XU, J.B. Weaver, D.M. Healy, and J. Lu, "Wavelet transform domain filtration technique", IEEE Transactions on Image Processing, vol. 3, 747-758, 1994.
3. D. L. Donoho, "Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect and noisy data", In Proceeding of Symposia in Applied Mathematic, American Mathematical Society vol. 00, 173-205, 1993.
4. D. L. Donoho, "Denoising by soft-thresholding," IEEE Trans. Inform. Theory, vol. 41, pp613-627, May 1995.
5. E.P. Simoncelli and E. H. Adelson, "Noise removal via Bayesian wavelet coring," Third Int. Conf. On Image Processing, vol. 1, pp. 379-382, September 1996.
6. X. Zong, A. F. Laine, and E. A. Geiser, "Speckle reduction and contrast enhancement of echocardiograms via multiscale nonlinear processing," IEEE Trans. Med. Imag., vol. 17, pp. 532-540, Aug. 1998.
7. M. Popescu, P. Cristea, and A. Bezerianos, "Multiresolution distributed filtering: A novel technique that reduces the amount of data required in high resolution electrocardiography," Future Gen. Comp. Sys., no. 15, pp. 195-209, 1999.
8. A. Achim, A. Bezerianos, and P. Tsakalides, "An Aplha-stable based Bayesian algorithm for speckle noise removal in the wavelet domain," Proc. NSIP-01, June 03-06, 2001, Baltimore, Maryland USA.
9. D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage", Biometrika, Also Tech. Report of Statistics, July, 1992, revised April 1993.
10. G. P. Nason, "Wavelet regression by cross-validation", Technical Report 447, Stanford University, Statistics Department, March 1994.
11. B. R. Hunt and O. Kubler, "Karhunen-Loeve multispectral image restoration, Part I: theory", IEEE Trans. On Acoustic Speech and Signal. Proc., ASSP vol. 32, 592-600, 1984.
12. N. P. Galatssanos and R. T. Chin, "Digital restoration of multichannel images", IEEE Trans. Acoust. SP., and Sign. Proco., vol. 37, 415-426, 1989.
13. M. K. Ozhan, A, T. Erden, M. I. Sezan, and A. M. Tekalp, "Efficient multiframe Wiener restoration of blurred and noisy image sequences", IEEE Trans. Image Proc., vol. 1, 453-476, 1992.
14. Eero P. Simoncelli, "Baysian denoising of visual images in the wavelet domain", Bayesian Inference in Wavelet Based Models eds., Pmuller and B Vidakovic, Chapter 18, pp291-308, lecture notes in statistics, vol. 141, Springer-Verlag, New York, 1999.
15. X. Huang and G. A. Woolsey, "Image denoising using Wiener filtering and wavelet thresholding", IEEE International Conference, ICME 2000, Proc. Vol. 3 pp1759-1762, 2000.
16. X. Huang, "Denoising Image via Minimum Variance Bound Bayesian Estimator," IEEE International Conf. ICME2001, Proc., pp533-536, 2001.
17. E. P. Simoncell and E. H. Adelson, "Noise removal via Bayesian wavelet coring", In Third Int. Conf. On Image Proc., vol. I, pp379-382, Lausanne, September 1996. IEEE Sig. Proc. Society.
18. S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," IEEE Trans. Pattern Anal. Machine Intell. Vol. 11, pp-674-692, July 1989.
19. P. Tsakalides, P. Reveliotis, and C. L. Nikias, "Scalar quantization of heavy-tailed signals," IEE Proc.-Vision, Imag. Sign. Proc., vol. 147, pp. 475-484, Oct. 2000.
20. S. R. Peterson, Y. H. Lee, S. A. Kassam, "Some statistical properties of alpha-trimmed mean and standard type M filters," IEEE on Acoust., Speech, Sig. Proc., vol. 36, no.5, 00707-713, 1988.