IMAGE NOISES REMOVAL ON ALPHA-STABLE VIA BAYESIAN ESTIMATOR

X.Huang$^1$ and A.C.Madoc$^2$

$^1$School of Information Sciences and 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 Bayesian estimator that recovers the signal component of the wavelet coefficients from original images by using an a-stable signal prior distribution is discussed. As we discussed in our earlier paper that the Bayesian estimator can approximate impulsive noise more accurately than other models and that the general case of the Bayesian processor does not have a closed-form expression. The attention drawn by this paper is the behaviors of $\alpha \in (0,1]$ following we discussed $\alpha \in [1,2]$ in our earlier paper [18]. Closer to a realistic situation, and unlike conventional 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 an investigation of a-stable simulations for a maximum likelihood estimator. 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 for illustration purposes.

INTRODUCTION

It is well known that linear filtering techniques have been used in many image-processing applications, which demonstrates their attractive natures such as mathematical simplicity and efficiency in the presence of additive Gaussian noise. However, they also blur sharp edges, distort lines and fine image details, less effectively remove tailed noise, and poorly treat the presence of signal-dependent noise.

Wavelet transform as a powerful tool for recovering signals from noise has been of considerably interest [1-6]. 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 [3].

As mentioned by Achim et al. [7], there are two major drawbacks for thresholding. One is that of the 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" [8]. These explicitly depend on the standard deviation of the 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], [8]. For certain applications the optimal threshold can be computed. An approach different from "universal thresholds" is discussed by Weyrich and Nason [9], in which cross-validation is used. Two approaches of 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 [10], multichannel image restoration [11], and multiframe image restoration [12], have explored the Wiener filter. It is well known that a classical solution that deals with the noise removal problem is the Wiener filter [13], [14]. 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 the modelling of the statistics of nature images, including Bayesian processing presuppose proper modelling for the prior probability density function of the signal [7, 13, 14, 15, 16, 17].

In this paper a maximum likelihood Bayesian estimator that recovers the signal component of the wavelet coefficients in original images by using an alpha-stable signal prior distribution is further discussed. As we known that the Gaussian ($\alpha = 2$) and the Cauchy ($\alpha = 1$) distributions are the only symmetric $\alpha$-stable distributions that have close-form probability density functions. This is, in fact, useful when using the principle of maximum likelihood estimation. Our previous paper basically investigated the natures of $\alpha \in [1, 2]$ for the Bayesian estimator [18], hence, it is a natural extension to discussed the situation when $\alpha \in (0,1]$ of the Bayesian estimator for images here.

2. ALPHASTABLE DISTRIBUTIONS AND MAXIMUM LIKELIHOOD ESTIMATOR

Tsakalides et al. [19], Achim et al. [7] 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 estimator for an $\alpha$-stable distribution after briefly introducing the symmetric $\alpha$-stable distribution ($\alpha$S).

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

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

(1)

The parameters $\alpha$, $\chi$ and $\delta$ describe completely a $\alpha$S distribution. The characteristic exponent $\alpha$ controls the heaviness.
of the tails of the stable density. \( \alpha \) can take values in (0,2); while \( \alpha = 1 \) and 2 define the Cauchy and Gaussian cases. There is not a closed-form expression known for the general \( \alpha \) stable probability density function (PDF). Thus, it is useful when using the principle of maximum likelihood estimation. The dispersion parameter \( \gamma (\gamma > 0) \) refers to the spread of the PDF. The location parameter \( \delta \) is analogous to the mean of the PDF. As discussed before that to a realistic situation, i.e. there is no further information about the PDF of the noise you are interested in, it would be helpful to try \( \alpha = 1.5 \) [15, 18]. However, since generally there is no closed form for the probability density function, we need to discuss the natures of \( \alpha \in (0,1) \). Figure 1 shows PDF for \( \alpha \)-stable with different values of \( \alpha \). To compare the case in [18], two curves are added, namely when \( \alpha = 0.3 \) and 0.5, otherwise the rest are the same, i.e. \( \gamma = 0.3 \). Also for comparison with the Laplace distribution, which was described in [17], it is shown in Figure 1. It is obvious that smaller \( \alpha \) values imply heavier tails and the 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.

\[
\text{PDF for stable alpha distribution: } \alpha = 0.3, 0.5, 1.0, 1.5, 1.8, \text{ and } 2.0 \text{ and Laplace distribution.}
\]

Assume a variable \( \hat{\theta} \) is unbiased it follows that

\[
E(\hat{\theta} - \theta) = 0
\]

which can be expressed as:

\[
\frac{1}{\gamma^2} \sum_{i=1}^{N} f_{\hat{\theta}}(\hat{\theta} | \hat{\mathbf{x}}) \hat{\mathbf{x}} = 0
\]

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

Following [17] we have

\[
\text{var(\hat{\theta})} \geq \frac{1}{E[\hat{\theta}^2] f_{\hat{\theta}}(\hat{\theta}; \hat{\mathbf{x}})^2]}
\]

The function \( \ln f_{\hat{\theta}, \theta}(\hat{\theta}; \theta) \) is well known as the “log likelihood” function of \( \theta \) (LLF). Its maximum likelihood estimate can be obtained from the equation:

\[
\frac{\partial \ln f_{\hat{\theta}, \theta}(\hat{\theta}; \theta)}{\partial \theta} = 0
\]

The first order differential of the log likelihood function with respect to \( \theta \) is called the maximum likelihood (ML) estimate. If an 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.

As paper [18] suggests that the value at around 1.5 is strongly recommended if there is no credible information about \( \alpha \). This is evidenced by the 2nd order simulations of the LLF for an \( \alpha \)-stable with \( \gamma = 0.1 \) shown in Figure 2.

\[
\text{Figure 2: The results of the 2nd order simulations of the LLF for an \( \alpha \)-stable with } \gamma = 0.1 \text{ viewed from the front of the } \alpha-\text{axis.}
\]

3. BAYESIAN ESTIMATOR

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 the input/output of the Bayesian estimator \( BE \), we have:

\[
BE = \frac{\int P_q(q)P_s(s)ds}{\int P_q(q)P_s(s)ds}
\]

\( P_s(s) \) is the prior PDF of the signal component of the wavelet coefficients of the ultrasound image and \( P_q(q) \) is the PDF of the wavelet coefficients corresponding to the noise.

In order to be able to construct the Bayesian processor in (6), we can estimate the parameters of the prior distributions of the signal (s) and noise (q) components of the wavelet coefficients (d). Then, we use the parameters to obtain the two prior PDFs \( P_s(s) \) and \( P_q(q) \) and the nonlinear input-output relationship \( BE \).
As Figure 2 may predict that when \( \alpha \in (0,1) \), in comparison with the case in [18], the input/output of the Bayesian estimator allows bigger \( \alpha \) and \( \gamma \) values. Figure 3 shows how \( \gamma \) affects the BE, where \( \alpha = 0.5 \), \( \sigma = 20 \) and Figure 4 shows the same case but with \( \alpha = 1.702 \) and \( \sigma = 4.5 \), as discussed in [18].

![Figure 3: How \( \gamma \) affects the BE. Here \( \gamma \) ranges from 0.08 to 1.0, with \( \alpha = 0.5 \) and \( \sigma = 20.0 \).](image1)

![Figure 4: How \( \gamma \) affects BE. Here \( \gamma \) ranges from 0.08 to 1.0, with \( \alpha = 1.702 \) and \( \sigma = 4.5 \).](image2)

It is obvious that for the curves with \( \gamma = 2.0 \) and \( \gamma = 3.0 \) in Figure 3 approximates respectively to the cases called "semisoft" and "Garrote" shrinkage functions for denoising images.

Figure 5 shows how \( \sigma \) affects BE when \( \alpha \in (0,1) \). In comparison with the cases shown in [18], we can confirm the above result that the input/output of the Bayesian estimator will allow bigger \( \gamma \) values in similar situations.

![Figure 5: How \( \sigma \) affects BE. Here \( \sigma \) ranges from 1.0 to 10, with \( \alpha = 0.5 \) and \( \gamma = 4.5 \).](image3)

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 \( \gamma \sigma \) as one parameter and showed a simulation by keeping \( \gamma \sigma \) 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 \( \sigma \) is generally estimated and we may even have no credible information about it. However, as in [18] we may take \( \alpha = 1.5 \) as first estimate. For the case when \( \alpha \in (0,1) \), we found no significant changes as expected.

4. SOME EXAMPLES

When we make measurements, we have no credible information about the noise value of the image we obtained. The only information one may have is experience and judgement of the noise level, which outlines of the denoising strategy. For the purpose of comparisons, we use a similar set of images as shown in [18].

Figure 6 shows the noisy image. If we did not have any information about the noisy image, we can, as mentioned above, initially take \( \alpha = 1.5 \), since we have more room due to \( \alpha \in (0,1] \) so we can have better results as showed the denoised image in Figure 7. The results by other methods, together with the Bayesian estimator for \( \alpha \)-stable method, are given in the table 1.

<table>
<thead>
<tr>
<th>Method</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>S/MSE</td>
<td>12.41</td>
<td>12.50</td>
<td>12.31</td>
<td>13.49</td>
</tr>
</tbody>
</table>

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 \( \alpha \)-stable
5. CONCLUSION

A new technique for denoising an image has been further 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.

6. REFERENCES