# ULTRASOUND IMAGE DENOISING VIA MAXIMUM A POSTERIORI ESTIMATION OF WAVELET COEFFICIENTS

A. Achim[1], A. Bezerianos[1], and P. Tsakalides[2]
[1]Department of Medical Physics, University of Patras, Greece
[2]Department of Electrical and Computer Engineering, University of Patras, Greece

*Abstract*— **Speckle noise removal by means of digital image processors could improve the diagnostic potential of medical ultrasound. This paper addresses the speckle suppression issue within the framework of wavelet analysis. As a first step of our approach, the logarithm of the original image is decomposed into several scales through a multiresolution analysis employing the 2-D wavelet transform. Then, we design a *maximum a posteriori* (MAP) estimator, which relies on a recently introduced statistical representation for the wavelet coefficients of ultrasound images [1]. We use an alpha-stable model to develop a blind noise-removal processor that performs a non-linear operation on the data. Finally, we compare our technique to current state-of-the-art denoising methods applied on actual ultrasound images and we find it more effective, both in terms of speckle reduction and signal detail preservation.**
*Keywords* **- speckle noise, wavelet transform, alpha-stable distributions, MAP estimation**

## I. Introduction

Nowadays, clinicians are allowed to noninvasively evaluate physiological processes within the human body by means of various medical imaging modalities, such as computed tomography, positron emission tomography, functional magnetic resonance imaging, and ultrasonography. Among them, ultrasound imaging is definitely the one that offers the best price-to-performance ratio receiving thus an added attention. However, one major issue when using this imaging modality is the inherent presence of speckle noise. Its occurrence is often undesirable, since it affects the tasks of human interpretation and diagnosis. On the other hand, its texture carries important information about the tissue being imaged. Speckle filtering is thus a critical pre-processing step in medical ultrasound imagery, provided that the features of interest for diagnosis are not lost.

Current speckle reduction methods are based on temporal averaging [2], [3], median filtering [4], and homomorphic Wiener filtering [5]. However, it is recognized that standard noise filtering methods often result in blurred image features. Indeed, single-scale representations of signals, either in time or in frequency, are often inadequate when attempting to separate signals from noisy data.

Recently, the wavelet transform has been proposed as a useful processing tool for signal recovery [6], [7], [8], [9]. Current state-of-the-art wavelet-based speckle reduction and image enhancement techniques employ a combination of wavelet shrinkage by soft and hard thresholding together with a generalized adaptive gain (GAG) for feature emphasis [8]. Also, in [9] the authors describe a combination of adaptive weighted median filtering [4] and soft thresholding. These methods try to address the inability of the original soft thresholding technique to balance between speckle suppression and signal detail preservation.

In a recent work [1], we have shown that a successful ultrasound imaging algorithm can achieve both noise reduction and feature preservation if it takes into consideration the true statistics of the signal and noise components. Specifically, we have shown that the subband decompositions of ultrasound images have significantly non-Gaussian statistics that are best described by families of heavy-tailed distributions such as the alpha-stable and consequently we designed a Bayesian estimator that exploits these statistics.

The approach presented here is similar to the method reported in [1]. The differences are that: (i) we select a different cost function for the design of the Bayesian processor, which gives raise to slightly different shapes of the nonlinearities applied to the noisy wavelet coefficients, and (ii) we propose a new method for estimating the parameters of the alpha-stable distribution from noisy observations, which is based on Koutrouvelis' [10] regression method.

## II. Preliminaries

Parametric Bayesian processing presupposes proper modeling for the prior probability density function (PDF) of the signal and noise. The statistical properties of speckle noise were studied by Goodman [2]. He has shown that, if the number of scatterers per resolution cell is large, a fully developed speckle pattern can be modeled as the magnitude of a complex Gaussian field with independent and identically distributed (i.i.d.) real and imaginary components. In general, speckle noise has a spatial correlation length on each axis which is roughly the same as the resolution cell size [11]. In order to generate spatially correlated speckle noise for use in simulations, one can lowpass filter a complex Gaussian random field and take the magnitude of the filtered output [12], [13].

Arsenault and April [14] have shown that when an image intensity is logarithmically transformed, the speckle noise is approximately Gaussian additive noise, and it tends to a normal probability much faster than the intensity distribution. Thus, by taking the logarithm of a speckle image we get:

$$\log I(x,y) = \log S(x,y) + \log \eta(x,y), \quad (x,y) \in \mathbf{Z^2} \qquad (1)$$

where by $I(x,y)$ we denoted a noisy observation (i.e. the recorded ultrasound image) of the two-dimensional function $S(x,y)$ (i.e. the noise-free image that has to be recovered) and by $\eta(x,y)$ the corrupting multiplicative speckle noise.

A. Achim is sponsored by the State Scholarship Foundation of Greece (IKY) under grant 542 / 1999.

The work of Dr. Tsakalides was supported by the Greek General Secretariat for Research and Technology under Program ΕΠΕΤ II, Code 97ΕΛ - 152.

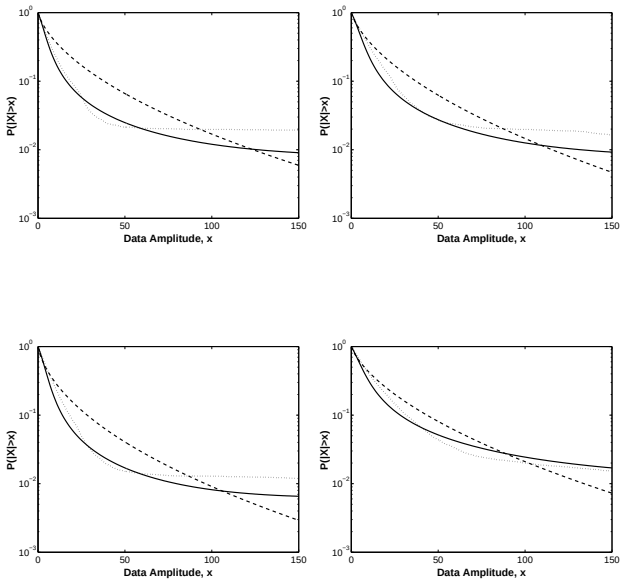


Fig. 1. Examples of APD plots of wavelet subband coefficients corresponding to four different ultrasound images. In each example the empirical APD (dotted line) is fitted with a $S\alpha S$ model (solid line) and the "generalized" Laplacian function (dashed line).

The wavelet transform is a linear operation. Consequently, after applying the *dyadic wavelet transform* (DWT) to (1), we get sets of noisy wavelet coefficients written as the sum of the transformations of the signal and of the noise:

$$d^i_{j,k} = s^i_{j,k} + \xi^i_{j,k}, \tag{2}$$

where $k = 0, ..., 2^{J+j} - 1$ and $-1 < j < -J$ refer to the decomposition level or scale and $i = 1, 2, 3$ refers to the three spatial orientations (see e.g. [15]).

The signal components of the wavelet decomposition in various scales are modeled as symmetric alpha-stable ($S\alpha S$) processes. The $S\alpha S$ distribution is best defined by its characteristic function:

$$\phi(\omega) = \exp(\jmath\delta\omega - \gamma|\omega|^\alpha), \tag{3}$$

where $\alpha$ is the *characteristic exponent*, taking values $0 < \alpha \leq 2$, $\delta$ $(-\infty < \delta < \infty)$ is the *location parameter*, and $\gamma$ $(\gamma > 0)$ is the *dispersion* of the distribution.

The $S\alpha S$ model is suitable for describing signals that have highly non-Gaussian statistics and its parameters can be estimated from noisy observations. For an experimental proof of the appropriateness of the use of $S\alpha S$ models in the context of ultrasound images we refer the reader to [1]. For the purpose of this paper we only show in Fig. 1 several examples of *amplitude probability density* (APD) plots corresponding to subband representations of different ultrasound images. Each data set is modeled using both the $S\alpha S$ family and the "generalized" Laplacian density function [7], [15]. The plots prove that the class of $S\alpha S$ distributions is superior because it provides a better fit to both the mode and the tails of the empirical density of the actual data.

## III. Design of a MAP processor for speckle mitigation

Motivated by the above considerations, we model the signal component of the wavelet coefficients using a two-parameter $S\alpha S$ distribution[1], while we use a zero-mean Gaussian model for the noise component. Also, we consider the signal and noise components to be independent.

In a Bayesian framework, referring to (2), $d_{j,k}$, $s_{j,k}$, and $\xi_{j,k}$ are considered as samples of the random variables $d$, $s$, and $\xi$, respectively. Our goal is to find the Bayes risk estimator $\hat{s}$ that minimizes the conditional risk, which is the loss averaged over the conditional distribution of $s$, given the noisy observation, $d$:

$$\hat{s}(d) = \arg\min_{\hat{s}} \int L[s, \hat{s}(d)]\, P_{s|d}(s\,|d)\, ds \tag{4}$$

Selecting the uniform cost function:

$$L[s, \hat{s}(d)] = \begin{cases} 0, & \text{for } |s - \hat{s}| < \epsilon \\ 1, & \text{otherwise} \end{cases} \tag{5}$$

the MAP estimator can be easily derived as being:

$$\hat{s}(d) = \arg\max_{\hat{s}} P_{s|d}(s\,|d) \tag{6}$$

It is important to underline at this point that under the loss function in (5), expression (4) is well defined for all $S\alpha S$ random variables (with characteristic exponent $\alpha$ taking values in the whole range $0 < \alpha \leq 2$).

Bayes' theorem gives the *a posteriori* PDF of $s$ based on the measured data:

$$P_{s|d}(s\,|d) = \frac{P_{d|s}(d\,|s)\, P_s(s)}{P_d(d)}, \tag{7}$$

where $P_s(s)$ is the *prior* PDF of the signal component of the measurements and $P_{d|s}(d\,|s)$ is the *likelihood* function. Substituting (7) in (6), we get:

$$\begin{aligned} \hat{s}(d) &= \arg\max_{\hat{s}} P_{d|s}(d\,|s)\, P_s(s) = \arg\max_{\hat{s}} P_\xi(d - s) P_s(s) \\ &= \arg\max_{\hat{s}} P_\xi(\xi) P_s(s) \end{aligned} \tag{8}$$

Fig. 2 depicts the numerically computed MAP input-output curves for five different values of the signal characteristic exponent, $\alpha$, namely, $\alpha = 2$ (Gaussian data), $\alpha = 1.95$ (slightly non-Gaussian data), $\alpha = 1.5$, $\alpha = 1$, and $\alpha = 0.5$ (considerably heavy-tailed data). Apart from the case $\alpha = 2$, all curves correspond to a nonlinear "coring" operation, i.e., large-amplitude observations are essentially preserved while small-amplitude values are suppressed. This is expected since small measurement values are assumed to come from signal values close to zero. On inspecting Fig. 2 it can be observed that for a given ratio $\gamma/\sigma$, the amount of shrinkage decreases as $\alpha$ decreases. The intuitive explanation for this behavior is that the smaller the value of $\alpha$, the heavier the tails of the signal PDF and the greater the probability that the measured value is due to the signal.

In order to be able to construct the MAP processor in (8), first one should estimate the parameters of the prior distributions of the signal and noise components of the measurements. Then, the parameters are used to "build" the two prior PDFs $P_\xi(\xi)$ and $P_s(s)$ and the nonlinear (in general) input-output relationship $\hat{s}(d)$. To achieve this, we

[1]Due to the properties of the DWT, the location parameter is in this case $\delta = 0$.

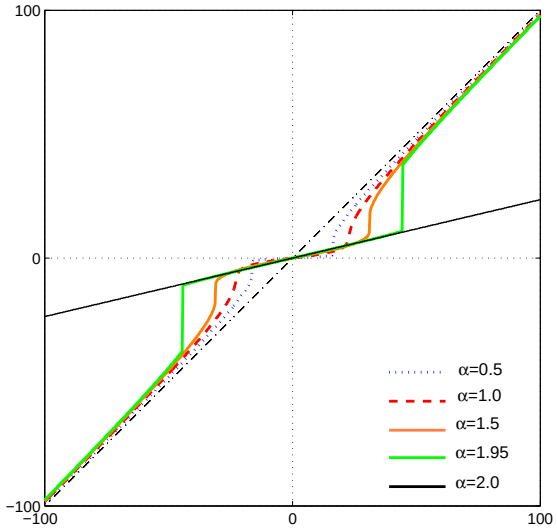


Fig. 2. MAP processor input-output curves for alpha-stable signal ($0.5 \leq \alpha \leq 2$) and Gaussian noise prior distributions. The dash-dotted line indicates the identity function.

observe that the PDF of the measured coefficients ($d$) is the convolution between the PDFs of the signal ($s$) and noise components ($\xi$). Consequently, the associated characteristic function of the measurements is given by the product of the characteristic functions of the signal and noise:

$$\Phi_d(\omega) = \exp\left(-\gamma_s|\omega|^{\alpha_s}\right) \cdot \exp\left(-\frac{\sigma^2}{2}|\omega|^2\right) \tag{9}$$

At this point, instead of directly fitting the Fourier transform of the empirical PDF of the measured coefficients with the function $\Phi_d(\omega)$ as we did in [1], we observe that (9) implies:

$$\log[-(\log|\Phi_d(\omega)|^2 + \sigma^2\omega^2)] = \log(2\gamma_s) + \alpha_s \log|\omega| \tag{10}$$

First, we estimate the level of noise $\sigma$ as in [6], then we find the parameters $\alpha_s$ and $\gamma_s$ by regressing $y = \log[-(\log|\Phi_d(\omega)|^2 + \sigma^2\omega^2)]$ on $w = \log|\omega|$ in the model

$$y_k = \mu + \alpha\, w_k + \epsilon_k \tag{11}$$

where $\mu = \log(2\gamma)$, $\epsilon_k$ denotes an error term, and $(\omega_k, k = 1, \ldots, K)$ is an appropriate set of real numbers. We should note here that Koutrouvelis [10] used a similar approach to estimate the parameters of alpha-stable distributions. He proved that its regression method gives very good results in terms of consistency, bias, and efficiency.

## IV. Results

In this section, we show simulation results obtained by processing one ultrasound image, randomly chosen from our database. The original image is shown in Fig. 3(a) and it represents an ultrasound scan of a fetal chest.

In order to obtain speckle images, we degraded the original test image by multiplying it with unit-mean random fields, as explained in Section II. We controlled the correlation length of the speckle by appropriately setting the size of the kernel used to introduce correlation to the underlying Gaussian noise. In practice uncorrelatedness of the noise could be achieved by decimating the image to the theoretical resolution limit of the imaging device [13]. Thus, a short-term correlation obtained with a kernel of size three was sufficient to model reality. We considered three different levels of simulated speckle noise.

We compared the results of our approach with the classical median filter, and wavelet shrinkage denoising using soft thresholding. The soft thresholding scheme was developed using Daubechies' Symmlet 8 mother wavelet as suggested in [8], while for our algorithm we used the Symmlet 4 wavelet. The maximum number of wavelet decompositions we used was five. In order to minimize the effect of pseudo-Gibbs phenomena, we have embedded both wavelet-based methods into the cycle spinning algorithm [16].

In order to quantify the achieved performance improvement, two measures were computed based on the original and the denoised data. For quantitative evaluation, we used the signal-to-mean-square-error (S/MSE) ratio, defined as:

$$S/MSE = 10\log_{10}\left(\sum_{i=1}^{K} S_i^2 \Big/ \sum_{i=1}^{K}(\hat{S}_i - S_i)^2\right) \tag{12}$$

This measure corresponds to the classical SNR in the case of additive noise. Remember that in ultrasound imaging, we are interested in suppressing speckle noise while at the same time preserving the edges of the original image that often constitute features of interest for diagnosis. For this reason, we also considered a qualitative measure for edge preservation [9]:

$$\beta = \frac{\Gamma(\Delta S - \overline{\Delta S}, \widehat{\Delta S} - \overline{\widehat{\Delta S}})}{\sqrt{\Gamma(\Delta S - \overline{\Delta S}, \Delta S - \overline{\Delta S}) \cdot \Gamma(\widehat{\Delta S} - \overline{\widehat{\Delta S}}, \widehat{\Delta S} - \overline{\widehat{\Delta S}})}} \tag{13}$$

where $\Delta S$ and $\widehat{\Delta S}$ are the high-pass filtered versions of $S$ and $\hat{S}$ respectively, obtained with a $3 \times 3$-pixel standard approximation of the Laplacian operator, and

$$\Gamma(S_1, S_2) = \sum_{i=1}^{K} S_{1_i} \cdot S_{2_i}. \tag{14}$$

The correlation measure, $\beta$, should be close to unity for an optimal effect of edge preservation. The results are summarized in Tables I and II respectively. It can be seen that our proposed Bayesian approach exhibits the best performance according to both metrics.

TABLE I

Quantitative image enhancement measures obtained using three denoising methods. The tabulated S/MSE metric is given in $dB$.

| | | | |
|---|---|---|---|
| Noisy image | 5.63 | 9.67 | 16.68 |
| Median Filtering | 11.07 | 14.24 | 17.97 |
| Soft Thresholding | 10.95 | 14.70 | 18.38 |
| Bayesian Denoising | 11.66 | 15.48 | 19.93 |

TABLE II

Qualitative image enhancement measures obtained using three denoising methods. Values of $\beta$ close to unity denote optimal edge preservation performance.

| | | | |
|---|---|---|---|
| Noisy image | 0.2577 | 0.3930 | 0.6933 |
| Median Filtering | 0.1989 | 0.4230 | 0.5310 |
| Soft Thresholding | 0.2806 | 0.4985 | 0.7391 |
| Bayesian Denoising | 0.3814 | 0.5203 | 0.7957 |

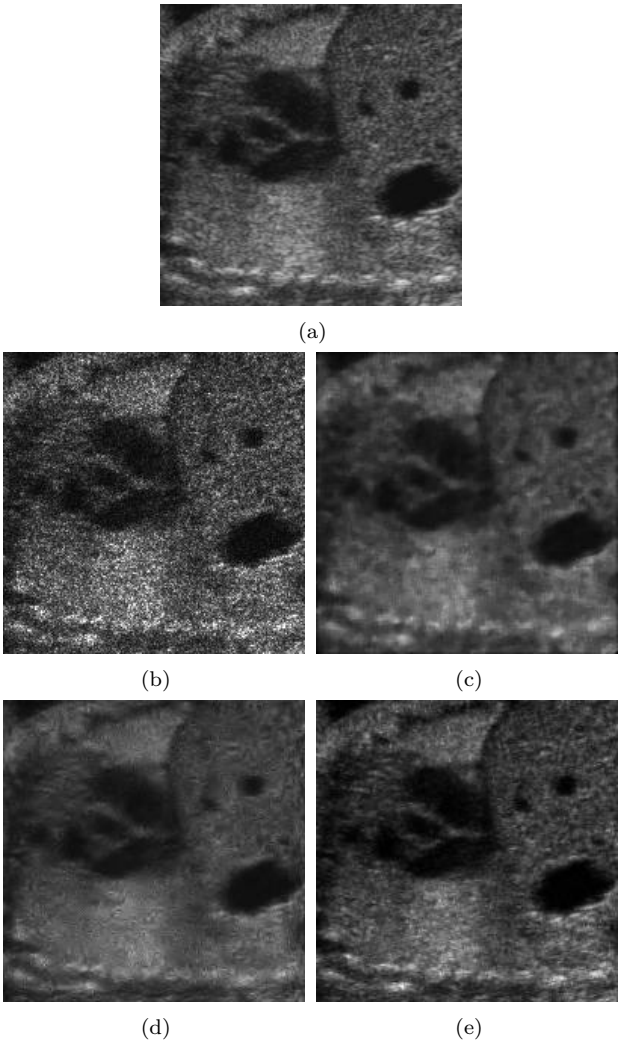

Fig. 3. *Results of various speckle suppressing methods. (a) Original image. (b) Image degraded with simulated speckle noise ($S/MSE = 9.67dB$). (c) Median filtering. (d) Translation-invariant soft thresholding. (e) Bayesian denoising.*

In Fig. 3 we show for visual comparison a representative result from the processing of our test image. Although it achieves a good speckle suppression performance, the median filter looses many of the signal details and the resulting image is blurred (Fig. 3(c)). On the other hand, the image processed by soft thresholding is oversmoothed (Fig. 3(d)). It seems that the Bayesian processor performs like a feature detector, retaining the features that are clearly distinguishable in the speckled data but cutting out anything which is assumed to be constituted by noise (Fig. 3(d)).

## V. Discussion

We introduced a novel technique for speckle noise removal using nonlinear processing of wavelet coefficients. The proposed processor is based on solid statistical theory, and it does not depend on the use of any *ad hoc* parameter.

Our algorithm was tested and found to be more effective than thresholding methods, which do not allow for an exact matching of the signal and noise distributions at different scales and orientations. The method proposed in Section III for choosing the "coring" nonlinearity could be thus considered as a principled way of shrinking noisy data, relying on the true statistics of the signal and noise wavelet coefficient. For example, the curve corresponding to $\alpha = 0.5$ in Fig. 2 mimics quite accurately the flavor of a hard thresholding operator.

Finally, we note that our algorithm could be easily adapted for the purpose of denoising other types of biomedical images where the noise can be (eventually after an appropriate transformation) modeled as additive Gaussian and signal-independent.

## Aknowledgment

The ultrasound image of the fetal chest has been provided by Acuson Corporation (Mountain Wiew, CA).

## References

[1] A. Achim, A. Bezerianos, and P. Tsakalides, "Novel Bayesian multiscale method for speckle removal in medical ultrasound images," *IEEE Trans. Med. Imag.*, vol. 20, pp. 772–783, Aug. 2001.

[2] J. W. Goodman, "Some fundamental properties of speckle," *J. Opt. Soc. Amer.*, vol. 66, pp. 1145–1150, November 1976.

[3] J. G. Abbott and F. L. Thurstone, "Acoustic speckle: Theory and experimental analysis," *Ultrason. Imag.*, vol. 1, pp. 303–324, 1979.

[4] T. Loupas, W. N. Mcdicken, and P. L. Allan, "An adaptive weighted median filter for speckle suppression in medical ultrasonic images," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 129–135, Jan. 1989.

[5] A. K. Jain, *Fundamental of Digital Image Processing.* NJ: Prentice-Hall, 1989.

[6] D. L. Donoho, "Denoising by soft-thresholding," *IEEE Trans. Inform. Theory*, vol. 41, pp. 613–627, May 1995.

[7] E. P. Simoncelli, "Bayesian denoising of visual images in the wavelet domain," in *Bayesian Inference in Wavelet Based Models* (P. Muller and B. Vidakovic, eds.), ch. 18, pp. 291–308, New York: Springer-Verlag, June 1999.

[8] 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.

[9] X. Hao, S. Gao, and X. Gao, "A novel multiscale nonlinear thresholding method for ultrasonic speckle suppressing," *IEEE Trans. Med. Imag.*, vol. 18, pp. 787–794, Sep. 1999.

[10] I. A. Koutrouvelis, "Regression-type estimation of the parameters of stable laws," *J. Amer. Statist. Assoc.*, vol. 75, pp. 918–928, Dec. 1980.

[11] R. F. Wagner, M. F. Insana, and S. W. Smith, "Fundamental correlation lengths of coherent speckle in medical ultrasonic images," *IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 35, pp. 34–44, Jan. 1988.

[12] F. Sattar, L. Floreby, G. Salomonsson, and B. Lovstrom, "Image enhancement based on a nonlinear multiscale method," *IEEE Trans. Image Processing*, vol. 6, pp. 888–895, June 1997.

[13] R. N. Czerwinski, D. L. Jones, and W. D. O'Brien, Jr, "Line and boundary detection in speckle images," *IEEE Trans. Image Processing*, vol. 7, pp. 1700–1714, Dec. 1998.

[14] H. H. Arsenault and G. April, "Properties of speckle integrated with a finite aperture and logarithmically transformed," *J. Opt. Soc. Amer.*, vol. 66, pp. 1160–1163, November 1976.

[15] 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.

[16] R. R. Coifman and D. L. Donoho, "Translation-invariant denoising," in *Wavelets and Statistics* (A. Antoniadis, ed.), Springer-Verlag Lecture Notes, 1995.