Adaptive Denoising Techniques for Medical Images in Wavelet Domain

Ashish Khare[1] Umashanker Tiwary[1,2]
ashishkhare@hotmail.com ust@iiita.ac.in

1. Department of Electronics & Communication, University of Allahabad, Allahabad, India
2. Indian Institute of Information Technology, Allahabad, India

ABSTRACT

Medical Images generally have poor contrast and complex nature of noise, as the noise is generated due to various acquisition devices and algorithm used. So denoising of medical images is a particularly delicate and difficult task. Wavelets provide an orthonormal basis for multiresolution analysis and decorrelation of non-stationary time series and spatial process. The decorrelation property of wavelets makes them suitable to denoise the images. There are three basic approaches for denoising: based on computation of simple threshold, regularity detection in wavelet domain and bayesian estimation in wavelet domain. The present work compares these techniques in term of Mean Square Error and discusses the relative advantages and disadvantages for their application to medical images. Simple threshold based methods are simple to implement and they perform good results. Regularity detection based techniques are better for the boundary detection of object as well as denoising. Bayesian estimation based methods perform very well for images where some information about distribution of noise is known. They will perform poor in some cases where the unknown type of noise is present.

I. INTRODUCTION

With a rapid increase in usages of medical images for diagnostic purpose, the major challenge of medical imaging is to provide a way to enable faithful extraction of scientific and clinical information. The images obtained by any imaging system gets various complex types of noises due to acquisition, transmission, display etc. Medical images are generally of poor contrast and they also get complex type of noises due to various acquisition, transmission, storage and display devices and also, due to application of different quantization, reconstruction and enhancement algorithms used [1]. For e.g. X-ray images will generally get Poisson noise, Ultrasound images gets speckle noise and MRI images were reported, corrupted by Rician noise.

The majority of research activities have been devoted on modeling the noise from different sources of noise. Although valuable information from these models can be obtained, but assumption for noise to follow any specific model will not yield good results. In the present paper, we have compared the results of our recently proposed [2]-[3] various better methods [2]-[4] for mix type of noise.

In Image denoising a trade of between noise suppression and preservation of actual image discontinuities must be made. To denoise the image with excessive smoothing of important details, the algorithm should be spatially adaptive. Wavelets can be very useful to denoising the image [5] by preserving important details. As noise commonly manifests itself as fine-grained structure and wavelets transform provides a scale representation by wavelet coefficients at finer scales. So discarding these coefficients would result in a natural filtering. The wavelet representation, due to its sparsity, edge detection and multiresolution property also lead to design adaptive denoising algorithm. Thus setting up some threshold and modifying the wavelet coefficient by certain function and then reconstructing the image will lead denoise image. A lot of work is available to make choice for suitable threshold. There are three basic approaches for denoising: one based on computation of simple threshold, second based on regularity detection in wavelet domain and third apply bayes estimation in wavelet domain.

The present work implements all the three improved techniques, suggested in our recent works [2]-[3] and compare them in terms of Mean Square Error and discusses the relative advantages and disadvantages for their applications to various kind of real life medical images where the nature of noise of unknown.

The rest of paper is organized as follows: Section II discusses the basic theory for computation of Discrete Wavelet Transform and Mallet's multiresolution principle. Section III describes the three classes of denoising schemes. In section IV, experiments and results are given and finally in section V conclusions are given.

II. DISCRETE WAVELET TRANSFORM AND MULTIRESOLUTION ANALYSIS

As the wavelets provide an orthonormal basis for multiresolution analysis and decorrelation of non-stationary time series and spatial process, so it becomes an important tool for signal analysis.

The Continuous Wavelet Transform (CWT) of a function f(t) is defined as [6]

$$CWT_{\psi}(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t)\psi\left(\frac{t-b}{a}\right)dt \tag{1}$$

where, $\psi(t)$ is a transforming function, called mother wavelet.

- a is real nonzero scale parameter
- b is real parameter that shifts the mother wavelet, so that $CWT_{\psi}(a,b)$ shows the local information about f(t) at t=b.

Thus wavelet transform represent a function f(t) in terms of a set of shifted and dilated basic mother wavelets (called basis functions).

In practical applications, only a discrete version of wavelet transform is employed [6]-[7]. With the Discrete Wavelet Transform (DWT), a dyadic decomposition of a data signal of length N is performed, by setting a scale parameter $a=2^j$ ($j=1,2,\ldots,\log_2 N$) and by using shifted and dilated versions of mother wavelet $\psi(t)$ and of a scaling function $\varphi(t)$, these functions form an orthonormal basis which are used to analyze the input signal by using several resolution steps.

In numerical computation, DWT is performed by using Mallet's pyramidal algorithm [8], known as Multiresolution Analysis. A discrete signal f[n] of length N is decomposed into its scaling (approximation) coefficients s[n] and detail coefficient d[n], by using two Quadrature Mirror filters g[n] and h[n], which can be computed starting from the mother wavelet and the scaling function, respectively

$$\psi(t) = \sqrt{2}\sum_n g[n]\psi(2t-n)$$
$$\phi(t) = \sqrt{2}\sum_n h[n]\phi(2t-n) \tag{2}$$

The detail coefficients represent high-pass filtered versions of the original signal, and the scaling coefficients represent low-pass filtered versions. If the input signal has a frequency band-width $\Delta f = f_s/2$, being f_s the sampling frequency, both the outputs from the g[n] and h[n] present a halved bandwidth $\Delta f/2$; therefore half of the samples can be discarded by subsampling, without any information loss, and the scaling and detail coefficients can be computed as

$$s[n] = \sum_k f[k]h[2n-k]$$
$$d[n] = \sum_k f[k]g[2n-k] \qquad (3)$$

The DWT decomposition halves the time resolution and doubles the frequency resolution. The decomposition process can be repeated and at each level, subsampling and filtering halves both the number of samples and the frequency band. This approach reduces the computational load. Since at each decomposition level, a reduced number of samples are processed, and, at the same time, improves both the time and frequency resolution. In fact high frequencies are better resolved in time domain and lower frequency in frequency domain.

In the DWT, only the scaling coefficients are recursively filtered. The input signal f[n] is split into its scaling and detail coefficients, by passing through h[n] and g[n], followed by subsampling. The scaling coefficients s[n] are further decomposed, whereas the detail coefficient d[n] are not analyzed and considered as the DWT outputs. At the j^{th} level, it has $N/2^j$ elements and spans over $\Delta f/2^j$ frequency bandwidth.

III. DENOISING IN WAVELET DOMAIN

One of major problem of diagnostic imaging systems is presence of noise, which reduces the detectability of problematic portion and noise is also a reason for creating artifacts. Following properties of wavelet transform makes wavelets to be useful for denoising of images [9]:

(i). Locality
(ii). Multiresolutuion
(iii). Edge Detection
(iv). Energy Compaction
(v). Decorrelation

In wavelet domain, the most essential information in image is compressed into relatively few large coefficients, which coincides with the areas of major spatial activity (edges, corners etc.) in the image. On the other hand, noise is spread over all coefficients and at typical noise levels the important coefficients can be well recognized. The wavelet transform can be a powerful method to distinguish the back-scattered signals from the noise, since it provides a nonuniform partitioning of signals in the time-frequency space.

A. Denoising by Wavelet Thresholding:

Wavelet thresholding is a popular approach for denoising due to its simplicity. In its most basic form, the method work as follows: compute the Discrete Wavelet Transform of observed image and each coefficient is threshold by comparing against a threshold and apply a simple nonlinearity (shrunk toward zero) to each wavelet coefficient then compute estimates of original image by applying inverse wavelet transform to transformed coefficients.

There are various strategies for thresholding. One of them is hard thresholding. The principle is, either to keep or to remove, based on absolute values of wavelet coefficients with respect to a fixed threshold T. This type of thresholding produces a discontinuity at threshold point. An alternative, a simple but performed shrinkage function is soft threshold

$$(w)_{new} = \begin{cases} 0 & ,if\ |w| \le T \\ w - T\,\mathrm{sgn}(w) & ,if\ |w| > T \end{cases} \qquad (4)$$

where T>0 is a threshold. The soft-threshold induces a bias on large coefficients.

A lot of work is available for choice of suitable threshold. Most of methods for estimating the threshold assume Additive White Gaussian Noise (AWGN) or some special type of noise model, and an orthogonal wavelet transform. Among those, well known thresholds are based on minimax analysis [10]. One of optimal threshold is asymptotic to $\sigma\sqrt{2\log(N)}$, is known as Universal threshold and other popular threshold is based on Median absolute deviation [10].

Recently we have proposed a threshold [2], which is independent from type of noise or noise model, which is:

$$T = \left(\frac{1}{2}\right)^{j-1}\left(\frac{\sigma}{\mu}\right)M \quad (5)$$

where, j is no. of level. σ, μ and M are standard deviation, mean and absolute median of wavelet coefficients. The proposed threshold is adaptive, as threshold values depends on mean, median and standard deviation of wavelet coefficients at particular level.

B. Denoising by Regularity Detection in Wavelet Domain:

In images the actual transitions are usually less abrupt or 'softer' than those produced by noise. In mathematical language we can say that local regularity of actual transitions and noise is different. This fact can be effectively exploited for denoising purpose. The rate of increase or decrease of the amplitude of the wavelet transform through resolution scales at a particular spatial position is directly related to the local regularity of the signal at that position i.e. actual edges produces large coefficients across many scales while dies out swiftly as the scale increases. Some of representative work using this idea is reconstruction of denoised image from detected multiscale edges cited in [11], regularity estimation for coefficient selection based on interscale ratio and then denoising the image in [12]. A popular method is to select coefficients based on interscale correlations [13]. They use correlation among wavelet coefficients at adjacent scales in order to detect the edge coefficients. The detected edge coefficients are left unmodified while all others are set to zero; a denoised image is then reconstructed by applying the Inverse Wavelet Transform. In one of our recent work [3], we have proposed a modified denoising scheme for medical image, which is based on fusion of detected multiscale edges to denoised image obtained by thresholding approach, which perform well than ordinary thresholding method or ordinary interscale correlation base denoising method.

C. Denoising by Bayesian Shrinkage based Estimation:

Bayesian approaches to wavelet shrinkage are less ad-hoc than earlier proposals and are generally referred to be effective in various literatures. In general, Bayes rules are shrinkers [14] and their shape in many cases has a desirable property: it can heavily shrink small arguments and slightly shrink large arguments. In Bayesian approach to wavelet shrinkage, the form of shrinkage function is induced by particular choice of prior distributions placed on the wavelet coefficients. There are several distribution models are proposed to model the distribution of image wavelet coefficients. For the natural images the proposed distribution model is a generalized laplacian distribution. Abramovich [4] have proposed the distribution as Normal distribution. We have done experiments with medical images and found that marginal prior models for wavelet coefficients is Gaussian mixture models same as proposed in [14]

$$p_Y(y) = \gamma N(0,(c\tau)^2) + (1-\gamma)N(0,\tau^2)$$

where $0 \le \gamma \le 1$ is a realization of binary random variable Γ, with $P(\Gamma=1) = 1 - P(\Gamma=0) = \pi$, and the parameters π, c and τ depends on the resolution level, but are constant within the given subband. The mixture of these two distributions is a prior with 'heavy' tails, which captures the sparseness common to most wavelet applications. According to [15], this choice of prior give rise to a shrinkage function which varies the amount of shrinkage with the relative magnitude of the coefficients. Smaller coefficients are shrunk essentially to zero, while larger coefficients are shrunk less. Here a separate shrinkage function is used at each resolution level.

The Minimum Mean Square Error (MMSE) estimate, using Posterior mean can be given has an explicit form

$$\hat{y}_{ms} = \left(P(\Gamma = 1 \mid w) \frac{(c\tau)^2}{\sigma^2_n + (c\tau)^2} + P(\Gamma = 0 \mid w) \frac{\tau^2}{\sigma^2_n + \tau^2} \right) w$$

where, $P(\Gamma = 1 \mid w) = \dfrac{\pi\, p_{W|\Gamma}(w \mid 1)}{(1-\pi)\, p_{W|\Gamma}(w \mid 0) + \pi\, p_{W|\Gamma}(w \mid 1)}$

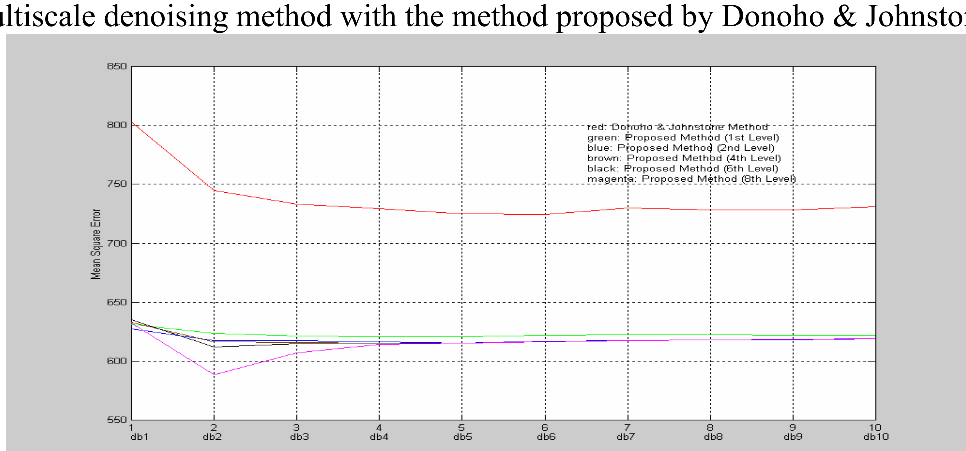
$$p_{W|\Gamma}(w \mid 1) \propto N(0, \sigma^2_n + (c\tau)^2) \quad \text{and} \quad p_{W|\Gamma}(w \mid 0) \propto N(0, \sigma^2_n + \tau^2)$$

since the parameters π, c and τ are different at different resolution level so this method is also level-wise adaptive.

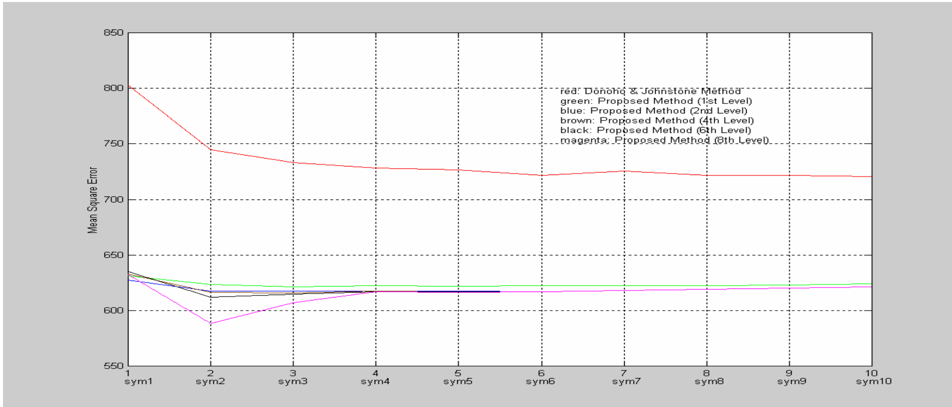
The form of above shrinkage function suggests that it may be interpreted as a smooth version of the hard-threshold function. The shrinkage function is also like as smooth interpolation between two lines of slope $\dfrac{\tau^2}{\sigma_n^2 + \tau^2}$ and $\dfrac{(c\tau)^2}{\sigma_n^2 + (c\tau)^2}$ as w moves from small to large values respectively. The rate at which the change between slope occurs is determined most strongly by size of π; the smaller π, the longer the shrinkage function lingers at the lower slope of $\tau^2 / (\sigma_n^2 + \tau^2)$. Hence hard-thresholding is a limiting case of the shrinkage function defined by above.

IV. EXPERIMENTS AND RESULTS

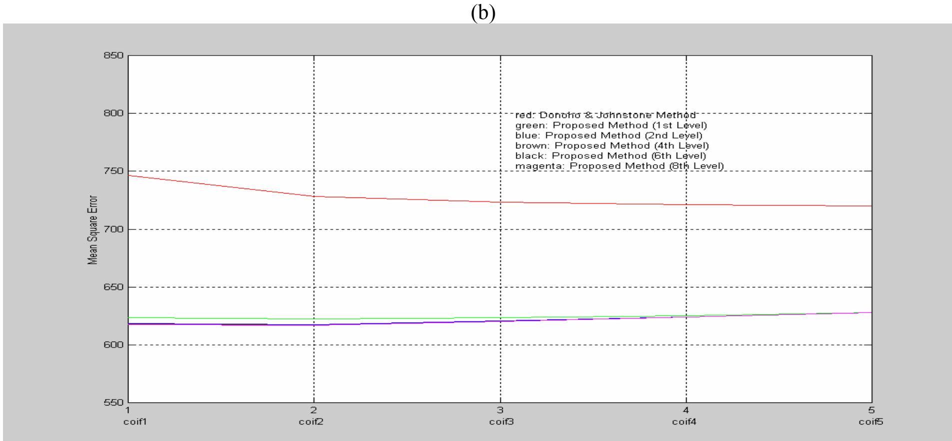
It have been suggested that noises in X-ray, Ultrasound and MRI images are some combination of Gaussian additive, Speckle and Salt-and-Pepper noise [2]. Experiments have been performed with 256×256 size images for all the methods discussed in section II. For investigating the performance of proposed method, reference images have been taken and mixture of different amount of different noises has been added artificially for measuring the performance. Mean Square Error (MSE) has been used for performance measure. Fig.[1] shows a comparison of our recently proposed multiscale denoising method with the method proposed by Donoho & Johnstone.



(a)



(b)



(c)

Fig. 1: Comparison of MSE Performance of proposed method at various levels to that of Donoho & Johnstone ((a): Daubechies wavelets, (b): Symlet wavelets and (c): Coiflet wavelets)

We have also experimented with different wavelets (Daubechies, Symmelet and Coiflet) and found that mother wavelet whose vanishing moment is less gives less smoothing or removes less noise, and the wavelet whose vanishing moment is large, produces distortion. The mother wavelet which has less number of points in basis function, produces coarser approximation and if we select the wavelet whose genus is large, give rise to smoothness. So a mother wavelet must be chosen by making different compromises and from fig.[1] it is obvious.

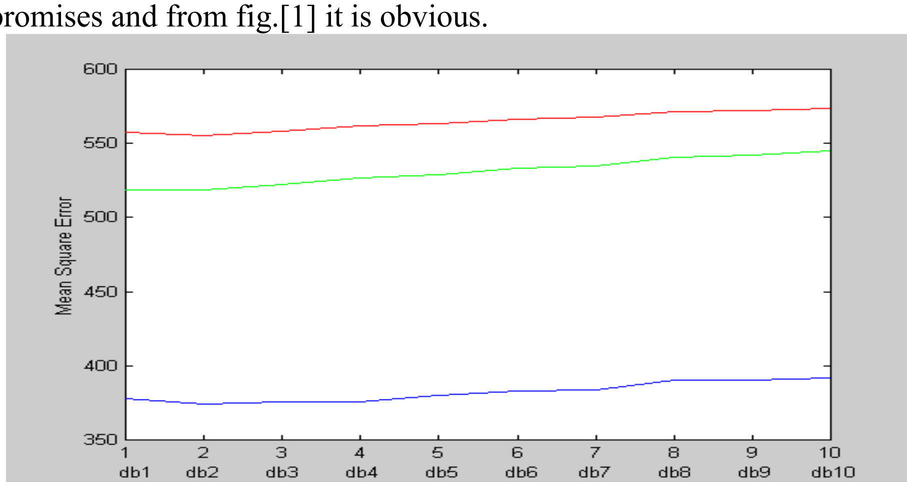


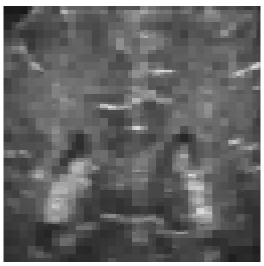
Fig. 2: Comparison of methods (Red, Green and Blue represent denoising by thresholding, multiscale edge based denoising and Bayesian thresholding respectively)

Fig.[2] shows the comparison of methods suggested in [2], [3] and [4]. For implementation of Bayesian shrinkage based method, we have chosen $\pi = 0.05$, $\sigma^2 = 1$, $\tau^2 = 0.01$ and $(c\tau)^2 = 50$.
Fig.[3] shows the one representative case of denoising for each of the methods discussed.

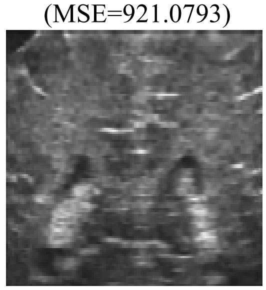
(a). Original Image

(b). Noisy Image
(MSE=921.0793)

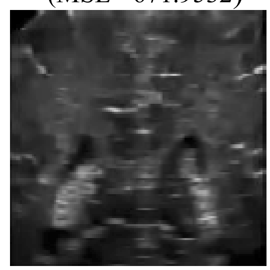
(c)
(MSE= 671.9532)



(d)
(MSE=789.6030)



(e)
(MSE=654.7823)



(f)
(MSE=459.7266)

Fig. 3 (a-Original Image, b-Noisy Image, c-denoised image by method in [2], d- denoised image by method in [10], e-denoised image by method in [3], f- denoised image by method in [4])

We have also observed that our recently proposed method [2] work well than method proposed in [4], for some cases of unknown noise, where Gaussian noise is not dominating. This is shown in Table[1], where approximately 80% of noise is due to speckle noise and a very little amount of Gaussian noise.

Table 1

Image	MSE of Noisy Image	Best MSE by method of [2]	Best MSE by method of [4]
test_image1	957.5682	621.8710	678.9800
test_image2	812.8340	538.7603	573.1548
test_image3	514.5122	389.8217	423.9492
test_image4	314.0231	267.8977	275.0104

V. CONCLUSION

Noises in medical images have variability from one condition to other condition (e.g machine specification, detector specification, surroundings etc.). So it is always very difficult to suggest a general method for restoration of medical images. One of most appealing aspect of wavelet shrinkage methods in image processing is their simplicity and processing each wavelet coefficients independently. Our main conclusions are as follows:

- Methods based on Denoising by wavelet thresholding are very simple to implement. These methods are commonly used for solving denoising problems. The proposed method based on multilevel soft-thresholding is adaptive, simple and outperform than other popular methods [10].

- Generally in diagnostic imaging, one has to also identify the boundaries of object in the region of interest. Methods based on denoising by wavelet thresholding create some blur at the boundaries [3], so the detection of object boundaries is difficult, so the idea for denoising the image, except at multiscale edges works extremely well.
- Bayesian wavelet shrinkage methods perform well than other approaches, but this approach is highly dependent on the assumptions made for the distribution of wavelet coefficients and the choice of estimators. Generally the assumptions about estimators and distribution parameter vary from case to case, and hence, the model dependence nature of these methods is main drawback. This is shown in table[1] for some representative cases, thus we can say that Bayesian estimation based method poorly perform for unknown type of noise.

We believe that the current denoising models can be improved in a number of ways for medical images, and our study for making comparison between some good methods will be helpful to image processing researchers. A greater challenge consists in to devise improved algorithms for unknown type of noise or general denoising models. Finally, from a longer-term perspective, major improvements are likely to come from statistical modes that capture some important properties of local image features, by including additional dependencies such as phase congruency between coefficients of complex multiscale transforms.

REFERENCES

[1] B.Jahne, H. HauBeker and P. GeiBler, "Handbook of Computer Vision and Applications, Vol 1 – Sensors and Imaging", *Academic Press*, 1999.

[2] A. Khare and U.S. Tiwary, "Soft-thresholding for denoising of Medical Images- A multiresolution approach" Accepted for publication in *International Journal on Wavelets, Multiresolution and Information Processing, World Scientific Pub.*

[3] A. Khare and U.S. Tiwary, "Multiscale Edge Preserving Soft-thresholding in Wavelet Domain for Denoising of Medical Images" to appear in *SIAM Conference on Imaging Science,* 2004, Philadelphia.

[4] F. Abramovich, T. Sapatinas and B.W. Silverman, "Wavelet Thresholding via a Bayesian approach", *Journal of Royal Statistical Society B*, vol. 60, pp. 725-749, 1998.

[5] S. Mallat, "Wavelet for Vision", *Procc. on IEEE: Special Issue on Wavelets*, vol. 84, pp. 604-614, 1996.

[6] R.K. Young, "Wavelet Theory and its Applications", Kluwer Academic Pub., 1993.

[7] I. Daubechies, "Ten Lectures on Wavelets", SIAM Pub. 1992.

[8] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation", *IEEE Trans. on PAMI*, vol. 11, pp. 674-693, 1989.

[9] J.K. Romberg, H.Choi and R.G. Baraniuk, "Bayesian Tree-structured Image Modeling using Wavelet-domain Hidden Markov Models", *IEEE Trans. on Image Processing*, vol. 10, no. 7, 2001.

[10] D.L. Donoho and I.M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage", *Biometrica,* vol. 81, pp. 425-455, 1994.

[11] S. Mallat and W.L. Hwang, "Singularity detection and processing with Wavelets", *IEEE Trans. on Information Theory*, vol. 38, no. 2, pp. 617-643, 1992.

[12] T.C. Hsung, D.P.K. Lun and W.C. Siu, "Denoising by singularity detection", *IEEE Trans. on Signal Processing,* vol. 47, no. 11, pp. 3139-3144, 1999.

[13] Y. Xu, J.B. Weaver, D.M. Healy and J. Lu, "Wavelet transform domain filters: a spatially selective noise filtration technique", *IEEE Trans. on Image Processing*, vol. 3, pp. 133-155, 1998.

[14] J. Portilla, V.Strela, M.J. Wainwright and E.P. Simoncelli, "Image Denoising using Scale Mixtures of Gaussians in the Wavelet Domain", *IEEE Trans. on Image Processing*, vol. 12, no. 11, 2003.

[15] H.A. Chipman, E.D. Kolaczyk and R.E. McCulloch, "Adaptive Bayesian wavelet shrinkage", *Journal of the American Statistical Association*, vol. 92, pp. 1413-1421, 1997.