

Analysis of Image De-Noising Techniques: A Review

Harshit Sahu¹, Prof. Shalini Sahay²

¹Research Scholar, ²Assistant Professor

^{1,2}Department of Electronics Communication and Engineering

^{1,2}SIRT, Bhopal, India

Abstract: A critical issue in image restoration is the problem of de-noising images while keeping the integrity of relevant image information. A large number of image de-noising techniques are proposed to remove noise. Mainly these techniques depend upon the type of noise present in images. So image de-noising still remains an important challenge for researchers because de-noising techniques remove noise from images but also introduce some artifacts and cause blurring. In this paper, we discuss various image de-noising and their features. Some of these techniques provide satisfactory results in noise removal and also preserve edges with fine details present in images. Noise modeling in images is greatly affected by capturing instruments, data transmission media, image quantization, and discrete sources of radiation. Different algorithms are used depending on the noise model. Most of the natural images are assumed to have additive random noise which is modeled as a Gaussian. Speckle noise is observed in ultrasound images whereas Rician noise affects MRI images. The scope of the paper is to focus on noise removal techniques for natural images.

Keywords- Image De-Noising, Artifacts, Data Transmission.

1. INTRODUCTION

Images are corrupted with various types of noises. So it is very difficult to get useful information from noisy images. That is why de-noising techniques are very important subject nowadays. For example, medical images obtained by X-ray or computed tomography CT in adverse conditions, or a mammographic image which may be contaminated with noise that can affect the detection of diseases or the object of interest. The aim of this work is to provide an overview of various de-noising techniques [1]. Some of these techniques provide satisfactory results in removing noise from images and also preserve edges with other fine details present in images. Different methods have been proposed for image restoration depending on the type of noise present in image. Some of these algorithms provide better result for smoothing flat Regions like spatial domain approaches. One of the biggest advantages of these techniques is a Speed but these techniques do not preserve the fine details in the image. On other hand wavelet domain techniques has great advantage of preserving edges and fine details in images. Image noise is a random variation of brightness or color information in images. It can be produced by sensor or circuitry of a scanner or digital camera. Noise in digital images arises during image

acquisition and/ or transmission. Image noise model: in image noise model image degradation and image restoration process are used. In image degradation a degradation function H is applied on input image $f(x,y)$ with some additive noise $n(x,y)$ and produce degraded image $g(x,y)$ [2]. After that image restored with specific techniques and produce an estimated image of original image.

2. EVOLUTION OF IMAGE DE-NOISING RESEARCH

Image De-noising has remained a fundamental problem in the field of image processing. Wavelets give a superior performance in image de-noising due to properties such as sparsely and multi-resolution structure. With Wavelet Transform gaining popularity in the last two decades various algorithms for de-noising in wavelet domain were introduced. The focus was shifted from the Spatial and Fourier domain to the Wavelet transform domain. Ever since Donoho's Wavelet based thresholding approach was published in 1995, there was a surge in the de-noising papers being published. Although Donoho's concept was not revolutionary, his methods did not require tracking or correlation of the wavelet maxima and minima across the different scales as proposed by Mallat [3]. Thus, there was a renewed interest in

wavelet based de-noising techniques since Donoho [4] demonstrated a simple approach to a difficult problem. Researchers published different ways to compute the parameters for the thresholding of wavelet coefficients. Data adaptive thresholds [6] were introduced to achieve the optimum value of the threshold. Later efforts found that substantial improvements in perceptual quality could be obtained by translation-invariant methods based on thresholding of an Un-decimated Wavelet Transform [7]. These thresholding techniques were applied to the non-orthogonal wavelet coefficients to reduce artifacts. Multi-wavelets were also used to achieve similar results. Probabilistic models using the statistical properties of the wavelet coefficient seemed to outperform the thresholding techniques and gained ground. Recently, much effort has been devoted to Bayesian de-noising in the Wavelet Domain. Hidden Markov Models and Gaussian Scale Mixtures have also become popular and more research continues to be published. Tree Structures ordering the wavelet coefficients based on their magnitude, scale, and spatial location have been researched. Data adaptive transforms such as Independent Component Analysis (ICA) have been explored for sparse shrinkage. The trend continues to focus on using different statistical models to model the statistical properties of the wavelet coefficients and their neighbors. The future trend will be towards finding more accurate probabilistic models for the distribution of non-orthogonal wavelet coefficients.

3. CLASSIFICATION OF IMAGE DE-NOISING ALGORITHMS

As shown in Figure 1, there are two basic approaches to image de-noising, spatial filtering methods and transform domain filtering methods.

3.1 Spatial Filtering

A traditional way to remove noise from image data is to employ spatial filters. Spatial filters can be further classified into non-linear and linear filters.

I. Non-Linear Filters

With non-linear filters, the noise is removed without any attempts to explicitly identify it. Spatial filters employ a low pass filtering on groups of pixels with the assumption that the noise occupies the higher

region of frequency spectrum. Generally spatial filters remove noise to a reasonable extent but at the cost of blurring images which in turn makes the edges in pictures invisible. In recent years, a variety of nonlinear median type filters such as weighted median [8], rank conditioned rank selection [9], and relaxed median [10] have been developed to overcome this drawback.

II. Linear Filters

A mean filter is the optimal linear filter for Gaussian noise in the sense of mean square error. Linear filters too tend to blur sharp edges, destroy lines and other fine image details, and perform poorly in the presence of signal-dependent noise. The Wiener filtering [11] method requires information about the spectra of the noise and the original signal and it works well only if the underlying signal is smooth. Wiener method implements spatial smoothing and its model complexity control corresponds to choosing the window size. To overcome the weakness of the Wiener filtering, Donoho and Johnstone proposed the wavelet-based denoising scheme in [12, 13].

3.2 Transform Domain Filtering

The transform domain filtering methods can be subdivided according to the choice of the basic functions. The basic functions can be further classified as data-adaptive and non-adaptive. Non-adaptive transforms are discussed first since they are more popular.

3.2.1 Spatial-Frequency Filtering

Spatial-frequency filtering refers to the use of low pass filters using Fast Fourier Transform (FFT). In frequency smoothing methods [11] the removal of the noise is achieved by designing a frequency-domain filter and adapting a cut-off frequency when the noise components are de-correlated from the useful signal in the frequency domain. These methods are time-consuming and depend on the cut-off frequency and the filter function behavior. Furthermore, they may produce artificial frequencies in the processed image.

3.2.2 Wavelet domain

Filtering operations in the wavelet domain can be subdivided into linear and nonlinear methods.

I. Linear Filters

Linear filters such as Wiener filter in the wavelet domain yield optimal results when the signal corruption can be modeled as a Gaussian process and the accuracy criterion is the mean square error (MSE) [14, 15]. However, designing a filter based on this assumption frequently results in a filtered image that

is more visually displeasing than the original noisy signal, even though the filtering operation successfully reduces the MSE. In [16] a wavelet-domain spatially adaptive FIR Wiener filtering for image de-noising is proposed where wiener filtering is performed only within each scale and intra scale filtering is not allowed.

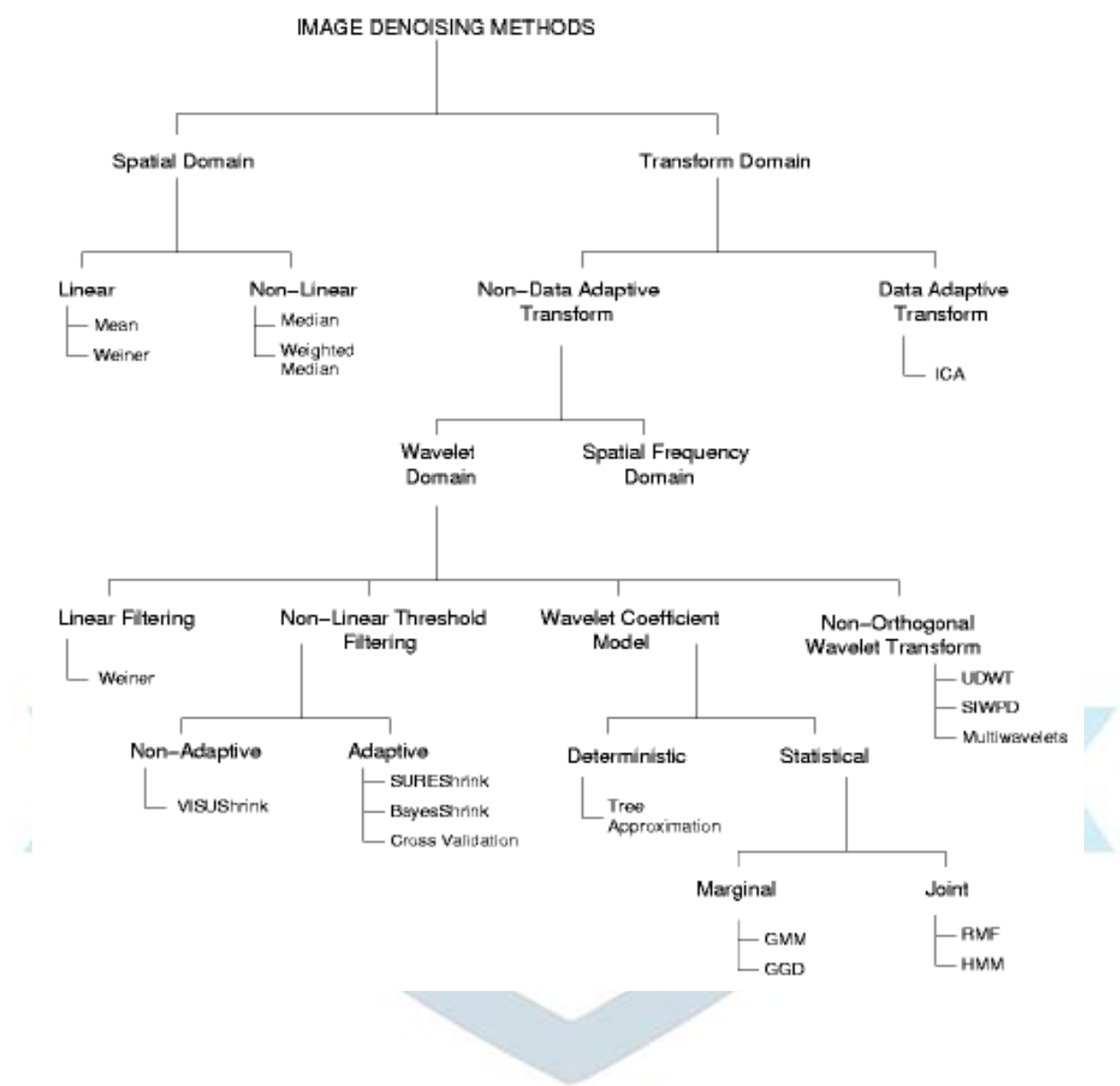


Figure 1 – Classification of Image De-Noising Methods

II. Non-Linear Threshold Filtering

The most investigated domain in de-noising using Wavelet Transform is the non-linear coefficient thresholding based methods. The procedure exploits sparsely property of the wavelet transform and the fact that the Wavelet Transform maps white noise in the signal domain to white noise in the transform domain. Thus, while signal energy becomes more concentrated into fewer coefficients in the transform domain, noise energy does not. It is this important principle that enables the separation of signal from noise. The procedure in which small coefficients are removed while others are left untouched is called Hard Thresholding [5]. But the method generates spurious blips, better known as artifacts, in the images as a result of unsuccessful attempts of removing moderately large noise coefficients. To overcome the demerits of hard thresholding, wavelet transform using soft thresholding was also introduced in [5]. In this scheme, coefficients above the threshold are shrunk by the absolute value of the threshold itself. Similar to soft thresholding, other techniques of applying thresholds are semi-soft thresholding and Garrote thresholding [6]. Most of the wavelet shrinkage literature is based on methods for choosing the optimal threshold which can be adaptive or non-adaptive to the image.

A. Non-Adaptive Thresholds

VISUShrink [12] is non-adaptive universal threshold, which depends only on number of data points. It has asymptotic equivalence suggesting best performance in terms of MSE when the number of pixels reaches infinity. VISUShrink is known to yield overly smoothed images because its threshold choice can be unwarrantedly large due to its dependence on the number of pixels in the image.

B. Adaptive Thresholds

SUREShrink [12] uses a hybrid of the universal threshold and the SURE [Stein's Unbiased Risk Estimator] threshold and performs better than VISUShrink. BayesShrink [17, 18] minimizes the Baye's Risk estimator function assuming Generalized Gaussian prior and thus yielding data adaptive threshold. BayesShrink outperforms SUREShrink most of the times. Cross Validation [19] replaces wavelet coefficient with the weighted average of

neighborhood coefficients to minimize generalized cross validation (GCV) function providing optimum threshold for every coefficient. The assumption that one can distinguish noise from the signal solely based on coefficient magnitudes is violated when noise levels are higher than signal magnitudes. Under this high noise circumstance, the spatial configuration of neighboring wavelet coefficients can play an important role in noise-signal classifications. Signals tend to form meaningful features (e.g. straight lines, curves), while noisy coefficients often scatter randomly.

III. Non-orthogonal Wavelet Transforms

Un-decimated Wavelet Transform (UDWT) has also been used for decomposing the signal to provide visually better solution. Since UDWT is shift invariant it avoids visual artifacts such as pseudo-Gibbs phenomenon. Though the improvement in results is much higher, use of UDWT adds a large overhead of computations thus making it less feasible. In [20] normal hard/soft thresholding was extended to Shift Invariant Discrete Wavelet Transform. In [21] Shift Invariant Wavelet Packet Decomposition (SIWPD) is exploited to obtain number of basic functions. Then using Minimum Description Length principle the Best Basis Function was found out which yielded smallest code length required for description of the given data. Then, thresholding was applied to de-noise the data. In addition to UDWT, use of Multi-wavelets is explored which further enhances the performance but further increases the computation complexity. The Multi-wavelets are obtained by applying more than one mother function (scaling function) to given dataset. Multi-wavelets possess properties such as short support, symmetry, and the most importantly higher order of vanishing moments. This combination of shift invariance & Multi-wavelets is implemented in [22] which give superior results for the Lena image in context of MSE.

IV. Wavelet Coefficient Model

This approach focuses on exploiting the multi-resolution properties of Wavelet Transform. This technique identifies close correlation of signal at different resolutions by observing the signal across multiple resolutions. This method produces excellent output but is computationally much more complex

and expensive. The modeling of the wavelet coefficients can either be deterministic or statistical.

A. Deterministic

The Deterministic method of modeling involves creating a tree structure of wavelet coefficients with every level in the tree representing each scale of transformation and nodes representing the wavelet coefficients. This approach is adopted in [23]. The optimal tree approximation displays a hierarchical interpretation of wavelet decomposition. Wavelet coefficients of singularities have large wavelet coefficients that persist along the branches of trees. Thus if a wavelet coefficient has a strong presence at a particular node then the case of it being a signal, its presence should be more pronounced at its parent nodes. If it is a noisy coefficient, for instance, spurious blip, then such a consistent presence will be missing. Lu et al. [24], tracked wavelet local maxima in scale-space, by using a tree structure. Another denoising method based on wavelet coefficient trees is proposed by Donoho [25].

B. Statistical Modeling of Wavelet Coefficients

This approach focuses on some more interesting and appealing properties of the Wavelet Transform such as multi-scale correlation between the wavelet coefficients, local correlation between neighborhood coefficients etc. This approach has an inherent goal of perfecting the exact modeling of image data with use of Wavelet Transform. A good review of statistical properties of wavelet coefficients can be found in [22] and [23]. The following two techniques exploit the statistical properties of the wavelet coefficients based on a probabilistic model.

I. Marginal Probabilistic Model

A number of researchers have developed homogeneous local probability models for images in the wavelet domain. Specifically, the marginal distributions of wavelet coefficients are highly kurtotic, and usually have a marked peak at zero and heavy tails. The Gaussian mixture model (GMM) [28] and the generalized Gaussian distribution (GGD) [17] are commonly used to model the wavelet coefficients distribution. Although GGD is more accurate, GMM is simpler to use. In [30], authors proposed a methodology in which the wavelet coefficients are assumed to be conditionally

independent zero-mean Gaussian random variables, with variances modeled as identically distributed, highly correlated random variables. An approximate Maximum A Posteriori (MAP) Probability rule is used to estimate the marginal prior distribution of wavelet coefficient variances. All these methods mentioned above require a noise estimate, which may be difficult to obtain in practical applications. Simoncelli and Adelson [18] used a two-parameter generalized Laplacian distribution for the wavelet coefficients of the image, which is estimated from the noisy observations. Chang et al. [19] proposed the use of adaptive wavelet thresholding for image denoising, by modeling the wavelet coefficients as a generalized Gaussian random variable, whose parameters are estimated locally (i.e., within a given neighborhood).

II. Joint Probabilistic Model

Hidden Markov Models (HMM) models are efficient in capturing inter-scale dependencies, whereas Random Markov Field [20] models are more efficient to capture intra-scale correlations. The complexity of local structures is not well described by Random Markov Gaussian densities whereas Hidden Markov Models can be used to capture higher order statistics. The correlation between coefficients at same scale but residing in a close neighborhood are modeled by Hidden Markov Chain Model whereas the correlation between coefficients across the chain is modeled by Hidden Markov Trees. Once the correlation is captured by HMM, Expectation Maximization is used to estimate the required parameters and from those, de-noised signal is estimated from noisy observation using well known MAP estimator. In [21], a model is described in which each neighborhood of wavelet coefficients is described as a Gaussian scale mixture (GSM) which is a product of a Gaussian random vector, and an independent hidden random scalar multiplier. Strela et al. [22] described the joint densities of clusters of wavelet coefficients as a Gaussian scale mixture, and developed a maximum likelihood solution for estimating relevant wavelet coefficients from the noisy observations. Another approach that uses a Markov random field model for wavelet coefficients was proposed by Jansen and Bulthel [23]. A disadvantage of HMT is the computational burden of the training stage. In order

to overcome this computational problem, a simplified HMT, named as uHMT [24] was proposed.

3.2.3 Data-Adaptive Transforms

Recently a new method called Independent Component Analysis (ICA) has gained wide spread attention. The ICA method was successfully implemented in [25] in de-noising Non-Gaussian data. One exceptional merit of using ICA is it's assumption of signal to be Non-Gaussian which helps to de-noise images with Non-Gaussian as well as Gaussian distribution. Drawbacks of ICA based methods as compared to wavelet based methods are the computational cost because it uses a sliding window and it requires sample of noise free data or at least two image frames of the same scene. In some applications, it might be difficult to obtain the noise free training data.

4. DISCUSSION

Performance of de-noising algorithms is measured using quantitative performance measures such as peak signal-to-noise ratio (PSNR), signal-to-noise ratio (SNR) as well as in terms of visual quality of the images. Many of the current techniques assume the noise model to be Gaussian. In reality, this assumption may not always hold true due to the varied nature and sources of noise. An ideal de-noising procedure requires a priori knowledge of the noise, whereas a practical procedure may not have the required information about the variance of the noise or the noise model. Thus, most of the algorithms assume known variance of the noise and the noise model to compare the performance with different algorithms. Gaussian Noise with different variance values is added in the natural images to test the performance of the algorithm. Not all researchers use high value of variance to test the performance of the algorithm when the noise is comparable to the signal strength. Use of FFT in filtering has been restricted due to its limitations in providing sparse representation of data. Wavelet Transform is the best suited for performance because of its properties like sparsely, multi-resolution and multi-scale nature. In addition to performance, issues of computational complexity must also be considered. Thresholding techniques used with the Discrete Wavelet Transform are the simplest to implement. Non-orthogonal

wavelets such as UDWT and Multi-wavelets improve the performance at the expense of a large overhead in their computation. HMM based methods seem to be promising but are complex. When using Wavelet Transform, Nason emphasized that issue such as choice of primary resolution (the scale level at which to begin thresholding) and choice of analyzing wavelet also have a large influence on the success of the shrinkage procedure. When comparing algorithms, it is very important that researchers do not omit these comparison details. Several papers did not specify the wavelet used neither the level of decomposition of the wavelet transform was mentioned. It is expected that the future research will focus on building robust statistical models of non-orthogonal wavelet coefficients based on their intra scale and inter scale correlations. Such models can be effectively used for image de-noising and compression.

REFERENCES

- [1] Zhaoming Kong and Xiaowei Yang, "Color Image and Multispectral Image Denoising Using Block Diagonal Representation", IEEE Transactions on Image Processing, 2019.
- [2] H. Guo, J. E. Odegard, M. Lang, R. A. Gopinath, I. W. Selesnick, and C. S. Burrus, "Wavelet based speckle reduction with application to SAR based ATD/R," First Int'l Conf. on Image Processing, vol. 1, pp. 75-79, Nov. 1994.
- [3] Robert D. Nowak, "Wavelet Based Rician Noise Removal", IEEE Transactions on Image Processing, vol. 8, no. 10, pp.1408, October 1999.
- [4] S. G. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," IEEE Trans. Inform. Theory, vol. 38, pp. 617-643, Mar. 1992.
- [5] D. L. Donoho, "De-noising by soft-thresholding", IEEE Trans. Information Theory, vol.41, no.3, pp.613- 627, May1995. <http://wwwstat.stanford.edu/~donoho/Reports/1992/denoiserelease3.ps.Z>
- [6] Imola K. Fodor, Chandrika Kamath, "Denoising through wavelet shrinkage: An empirical study", Center for applied science computing Lawrence Livermore National Laboratory, July 27, 2001.
- [7] R. Coifman and D. Donoho, "Translation invariant de-noising," in Lecture Notes in Statistics:

Wavelets and Statistics, vol. New York: Springer-Verlag, pp. 125--150, 1995.

[8] R. Yang, L. Yin, M. Gabbouj, J. Astola, and Y. Neuvo, "Optimal weighted median filters under structural constraints," IEEE Trans. Signal Processing, vol. 43, pp. 591–604, Mar. 1995.

[9] R. C. Hardie and K. E. Barner, "Rank conditioned rank selection filters for signal restoration," IEEE Trans. Image Processing, vol. 3, pp.192–206, Mar. 1994.

[10] A. Ben Hamza, P. Luque, J. Martinez, and R. Roman, "Removing noise and preserving details with relaxed median filters," J. Math. Imag. Vision, vol. 11, no. 2, pp. 161–177, Oct. 1999.

[11] A.K.Jain,Fundamentals of digital image processing. Prentice-Hall,1989.

[12] David L. Donoho and Iain M. Johnstone,"Ideal spatial adaption via wavelet shrinkage", Biometrika, vol.81, pp 425-455, September 1994.

[13] David L. Donoho and Iain M. Johnstone., "Adapting to unknown smoothness via wavelet shrinkage", Journal of the American Statistical Association, vol.90, no432, pp.1200-1224, December 1995. National Laboratory, July 27, 2001.

[14] V. Strela. "Denoising via block Wiener filtering in wavelet domain". In 3rd European Congress of Mathematics, Barcelona, July 2000. Birkhäuser Verlag.

[15] H. Choi and R. G. Baraniuk, "Analysis of wavelet domain Wiener filters," in IEEE Int. Symp. Time-Frequency and Time-Scale Analysis, (Pittsburgh), Oct. 1998.
<http://citeseer.ist.psu.edu/article/choi98analysis.html>

[16] H. Zhang, Aria Nosratinia, and R. O. Wells, Jr., "Image denoising via wavelet-domain spatially adaptive FIR Wiener filtering", in IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing, Istanbul, Turkey, June 2000.

[17] E. P. Simoncelli and E. H. Adelson. Noise removal via Bayesian wavelet coring. In Third Int'l Conf on Image Proc, volume I, pages 379-382, Lausanne, September 1996. IEEE Signal Proc Society.

[18]. H. A. Chipman, E. D. Kolaczyk, and R. E. McCulloch: 'Adaptive Bayesian wavelet shrinkage', J. Amer. Stat. Assoc., Vol. 92, No 440, Dec. 1997, pp. 1413-1421.

[19] MarteenJansen, Ph.D.Thesis in "Wavelet thresholding and noise reduction" 2000.

[20] M. Lang, H. Guo, J.E. Odegard, and C.S. Burrus, "Nonlinear processing of a shift invariant DWT for noise reduction," SPIE, Mathematical Imaging: Wavelet Applications for Dual Use, April 1995.

[21] I. Cohen, S. Raz and D. Malah, Translation invariant denoising using the minimum description length criterion, Signal Processing, 75, 3, 201-223, (1999).

[22] T. D. Bui and G. Y. Chen, "Translation-invariant denoising using multi-wavelets", IEEE Transactions on Signal Processing, Vol.46, No.12, pp.3414-3420, 1998.

[23] R. G. Baraniuk, "Optimal tree approximation with wavelets," in Proc. SPIE Tech. Conf. Wavelet Applications Signal Processing VII, vol. 3813, Denver, CO, 1999, pp. 196-207.

[24] J. Lu, J. B. Weaver, D.M. Healy, and Y. Xu, "Noise reduction with multiscale edge representation and perceptual criteria," in Proc. IEEE-SP Int. Symp. Time- Frequency and Time-Scale Analysis, Victoria, BC, Oct. 1992, pp. 555–558.

[25] D. L. Donoho, "CART and best-ortho-basis: A connection," Ann. Statist., pp. 1870–1911, 1997.