Threshold Estimation for Wavelet Domain Filtering of Signal-dependent Noise

Mitko Kostov, Cvetko Mitrovski, and Momcilo Bogdanov
Faculty of Technical Sciences, Bitola, Macedonia
Faculty of Electrical Engineering and Information Technologies, Skopje, Macedonia
Phone: (389) 47-207 711 (389) 2-3 064 262, Fax: (389) 47-203 370, E-mail: mitko.kostov@uklo.edu.mk

Keywords: Denoising, signal-dependent noise, threshold, wavelet domain filtering.

Abstract - This paper presents a method for signal denoising in the case of varying noise proportional to the local signal intensity. The signals are processed in a wavelet domain with a non-uniform threshold adjusted to the noise level.

1. INTRODUCTION

There are many methods for noise removal from signals, but very few of them focus on removing varying noise that depends on the local intensity of the signal. This kind of signal-dependent noise is commonly found in nuclear medicine (NM) images. Until now, the offered methods have been based on conventional filtering in time and frequency domain and lately, wavelet transforms. Research to date in wavelet-domain filtering has focused on removing additive white Gaussian noise. This type of noise can be removed by using a global threshold or multiscale products of the detail coefficients [1-3], but it is inappropriate for signal-dependent noise. One simple fix would be to work with the square-root of the image, since this operation is variance stabilizing [1]. Another method for Poisson noise removal in the wavelet domain uses a non-uniform threshold matrix for filtering the wavelet coefficients calculated from the raw data [4].

In this paper we propose a novel wavelet based method for removal of signal-dependent noise. It generates a non-uniform threshold adjusted to the noise level. The method uses standard wavelet filtering outlined in Section 2. In Section 3 we discuss how to estimate the varying threshold. In Section 4 we verify the validity of our approach on two deterministic signals contaminated with artificially added noise proportional to the signal intensity.

2. NOTION AND OTHER PRELIMINAIRES

In series expansion of discrete-time function f using wavelets

$$f(t) = \sum_{j=1}^{J} \sum_{k=1}^{2^{-j}M} d_{jk}\psi_{jk}(t) + \sum_{k=1}^{2^{-J}M} a_{Jk}\phi_{Jk}(t)\ , \quad (1)$$

ψ_{jk} and ϕ_{jk} denote wavelet and scaling function, respectively, the indexes j and k are for dilatation and translation, and a_{Jk} and d_{jk} are approximation and detail coefficients.

The most popular form of wavelet-based filtering, commonly known as wavelet shrinkage [1], weights the corresponding wavelet coefficient by a number $0 \le h_{jk} \le 1$:

$$h_{jk}^{(\text{soft})} = \begin{cases} 1 - \dfrac{\tau_j \operatorname{sgn}(d_{jk})}{d_{jk}}, & \text{if } |d_{jk}| \ge \tau_j \\ 0, & \text{if } |d_{jk}| < \tau_j \end{cases} \quad (2)$$

where τ_j is a user-specified threshold level.

3. THE PROPOSED ESTIMATION

Let **y** denotes a noisy signal that consists of a noise-free signal **s** and a noise **n**:

$$\mathbf{y} = \mathbf{s} + \mathbf{n}. \quad (3)$$

Since the wavelet transform (WT) is a linear operation, the wavelet coefficients $\mathbf{D}$ = WT($\mathbf{y}$) satisfy:

$$\mathbf{D} = \mathbf{D_s} + \mathbf{D_n}, \quad (4)$$

where $\mathbf{D_s}$ = WT($\mathbf{s}$) and $\mathbf{D_n}$ = WT($\mathbf{n}$).

Since the noise is proportional to the local signal intensity, a threshold τ_j for filtering of the wavelet coefficients at the level j should not be uniform for all the coefficients $D(i)$, but it should change depending on the noise level. Since the wavelet approximation coefficients $\mathbf{A}$ contain the signal identity and have equal length as the detail coefficients, these coefficients follow the signal shape and can be used to obtain the desired non-uniform threshold $\boldsymbol{\tau}$. In particular, at lower decomposition scales the difference between the approximation coefficients and the signal is smaller. In addition, due to the signal-dependency the noise coefficients $|\mathbf{D_n}|$ have a similar shape to the shape of the approximation coefficients. This implies that a threshold $\boldsymbol{\tau}$ should have some similar form, i.e. to be higher where the signal intensity is higher and vice versa. Owing to this, the approximation coefficients $\mathbf{A}$ can be normalized by multiplying them with a scalar α which will result in normalized coefficients $\alpha\mathbf{A}$ with equal energy as that of the detail coefficients $\mathbf{D}$. Since the coefficients $\mathbf{D}$ and $\alpha\mathbf{A}$ have equal energy and at the same time, the coefficients $\mathbf{D}$ contain narrower and higher peaks compared to the coefficients $\alpha\mathbf{A}$, the coefficients $\alpha\mathbf{A}$ will be smaller than the coefficients $|\mathbf{D}|$ where the signal portion in (4) is bigger, but bigger than coefficients $|\mathbf{D_n}|$ where there is no signal. Consequently, they can be used as a threshold $\boldsymbol{\tau}$ for filtering of the coefficients $\mathbf{D}$:

$$\boldsymbol{\tau} = \alpha\mathbf{A}, \text{ where } \alpha \in R_+ \cup \{0\}. \quad (5)$$

The coefficient α can be obtained from the condition for the coefficients $\mathbf{D}$ and $\alpha\mathbf{A}$ to have equal energy:

$$\sum_i D(i)^2 = \sum_i \left(\alpha A(i)\right)^2 . \quad (6)$$

For the parameter α it is obtained:

$$\alpha = \sqrt{\frac{\sum_i D(i)^2}{\sum_i A(i)^2}} . \quad (7)$$

Since the coefficients **D** and α**A** have equal energy but not exactly same form, it holds that if for some i, $|D(i)| > \alpha A(i)$ (the signal is strong), then for some $j \neq i$, $|D(j)| < \alpha A(j)$ (there is noise).

In general, since the noise **n** is proportional to the local signal intensity, it can be expressed as:

$$n(i) = \beta_n s(i)^n + \cdots + \beta_1 s(i) + \beta_0, \quad i = 0, \cdots, L-1, \quad (8)$$

where L is the length of the vectors **n** and **s**. In addition, the wavelet transform is a linear operation and wavelet coefficients depend on the signal. The absolute value $|D_n(i)|$ can be written as:

$$\left|D_n(i)\right| = \gamma_n A(i)^n + \cdots + \gamma_1 A(i) + \gamma_0, \quad i = 0, \cdots, L-1. \quad (9)$$

For the threshold $\boldsymbol{\tau}$ it follows that:

$$\tau(i) = \alpha_n A(i)^n + \cdots + \alpha_1 A(i) + \alpha_0, \quad i = 0, \cdots, L-1, \quad (10)$$

where $\alpha_0, \alpha_1, \ldots, \alpha_n \in R$.

The coefficients α_0, α_1, … can be obtained by minimizing the error E in the smallest squares sense:

$$E = \frac{1}{2} \sum_i \left(\left|D(i)\right| - \left(\alpha_n A(i)^n + \cdots + \alpha_1 A(i) + \alpha_0 \right) \right)^2 \quad (11)$$

For the purpose of simplicity, the threshold $\boldsymbol{\tau}$ can take the form (5), and in the same time the error function E which is to be minimized is:

$$E = \frac{1}{2} \sum_i \left(\left|D(i)\right| - \alpha A(i) \right)^2 . \quad (12)$$

But, the problem is a bit more complex. This is illustrated through a deterministic 1-D signal (*Fig. 1a*) contaminated with artificially generated noise proportional to the signal intensity (*Fig. 1b*). The first level approximation coefficients are shown in *Fig. 1c*. The approximation coefficients are obtained by using non-decimated wavelet transform [3] and NPR-QMF filters from [5]. From *Fig. 1* it can be noticed that the approximation coefficients follow the signal contour.

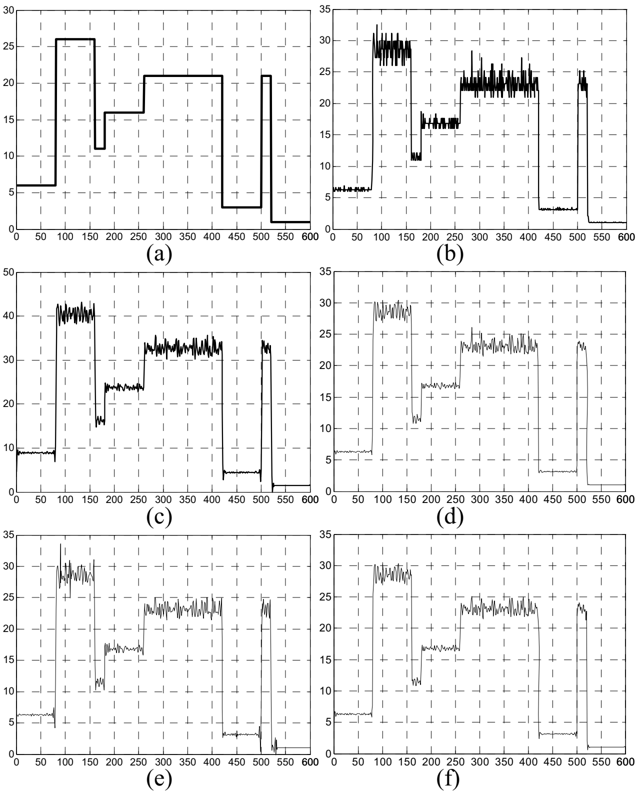


Fig. 1. (a) Deterministic signal; (b) noisy signal; (c) first level approximation coefficients; (d) reconstructed signal by using the proposed approach; (e) reconstructed signal by using the universal global threshold with db7; (f) reconstructed signal by using multiscale product.

In *Fig. 2* the first level approximation and detail coefficients, **A** and **D,** are given on a part of the interval. The points in *Fig. 2b* denote the coefficients values. By comparing *Fig. 2a* and *Fig. 2b* it can be seen that the

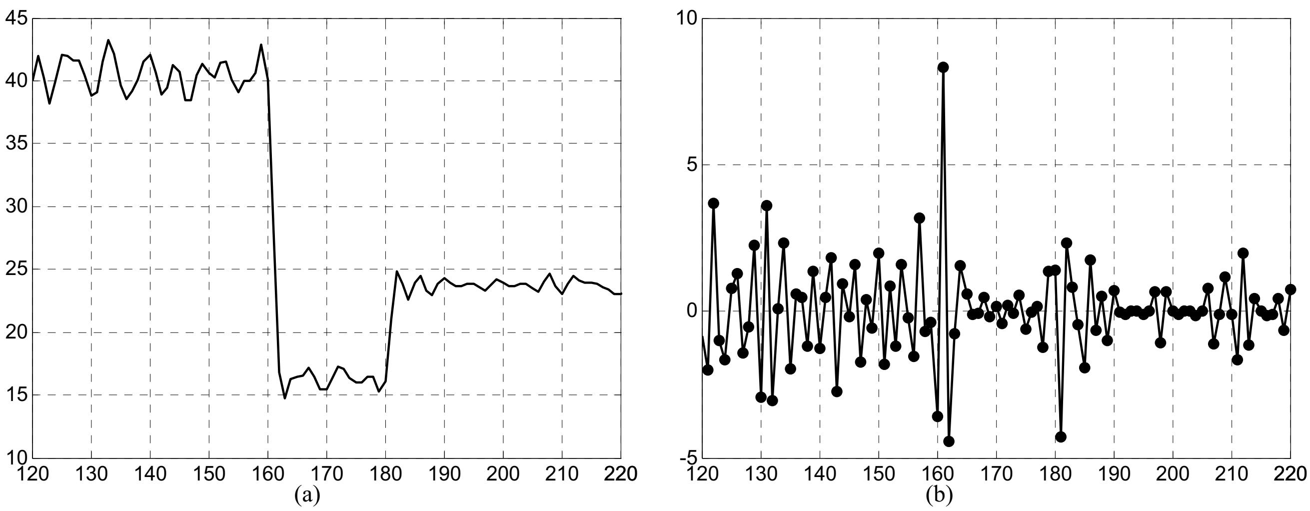
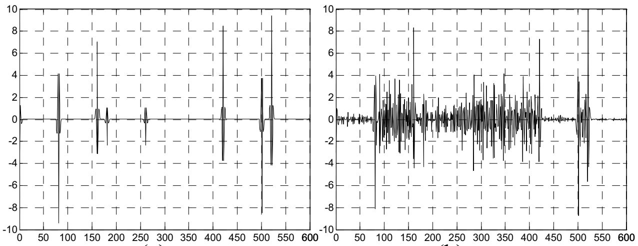


Fig. 2. Part of the first level wavelet coefficients from the signal shown in *Fig. 1b* on the interval. (a) Approximation coefficients **A**; (b) detail coefficients **D**.



(a) (b)

Fig. 3. (a) First level detail coefficients of the noise-free signal; (b) first level detail coefficients of the noisy signal.

coefficients **D** contain signal details $\mathbf{D_s}$ around the positions 160 and 180 (jumps in *Fig. 2a*, i.e. peaks in *Fig. 2b*); while in the other regions in the interval 120-220 of the coefficients **D** there is noise. Also it can be noticed from *Fig. 2a* that the signal intensity in the interval 120-160 is higher than the signal intensity in the intervals 160-180 and 180-220, so the noise in *Fig. 2b* follows the signal level, too: it is higher in the interval 120-160, and lowest in the interval 180-220.

In addition, in *Fig. 2b* it can be seen that the detail coefficients |**D**| do not form monotonically increasing or decreasing vectors on a given interval and do not have local extrema (maximums/minimums) with values that vary very little from the near coefficients values. On the contrary, the coefficients **D** in *Fig. 2b* are like waves and very often change the sign of their values. Hence, some of the coefficients have values that are close to zero, and as a consequence they contain a lot of local extremes (maximums/minimums). The values of these local extremes can differ a lot from the near coefficients values. As a consequence of this, a mistake will be made if all detail coefficients (including those close to zero) take part in the normalization of the coefficients **A** (5-7), i.e. minimization of the error function (10). These coefficients (with values close to zero) should be omitted because they exist only as a result of the fast change of the sign of the coefficients **D**. Hence, it is better to form new signals $\mathbf{D_1}$ and $\mathbf{A_1}$ from the detail and approximation coefficients, **D** and **A**, respectively, consisted only from the local extremes. After that a threshold $\boldsymbol{\tau}$ should be found through:

1) Approximation coefficients normalization in (5) where the coefficient α is found as

$$\alpha = \sqrt{\frac{\sum_i D_1(i)^2}{\sum_i A_1(i)^2}} \quad \text{or through} \qquad (13)$$

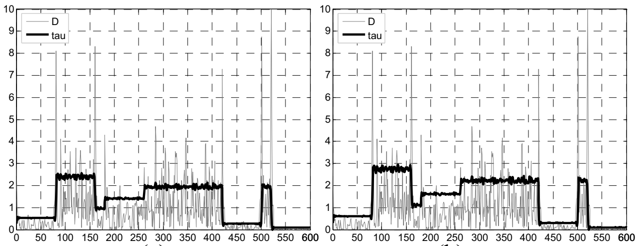
2) Minimization of the new function E_1:

$$E_1 = \frac{1}{2}\sum_i \left(\left|D_1(i)\right| - \left(\alpha_n A_1(i)^n + \alpha_{n-1}A_1(i)^{n-1} + \cdots + \alpha_1 A_1(i) + \alpha_0\right)\right)^2 \qquad (14)$$

The simplified form in (12) is:

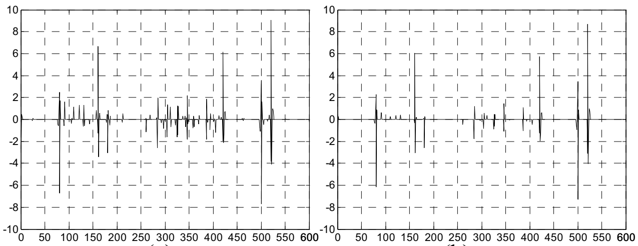
$$E_1 = \frac{1}{2}\sum_i \left(\left|D_1(i)\right| - \alpha A_1(i)\right)^2 . \qquad (15)$$

The coefficients $\mathbf{D_1}$ in (13) and (14) are selected according to the following mini-algorithm:



(a) (b)

Fig. 4. First level detail coefficients |**D**| and the threshold $\boldsymbol{\tau}$ when the parameter α is estimated through the normalization of the coefficients **A** by using: (a) all samples of **A** and **D** in (9); (b) $\mathbf{A_1}$ and $\mathbf{D_1}$ in (15).



(a) (b)

Fig. 5. Filtrated detail coefficients by using the threshold from: (a) *Fig. 4a*; (b) *Fig. 4b*.

$$D'(i) = \begin{cases} D(i) & if \quad \begin{matrix} D(i) > \max\left(D(i-1), D(i+1)\right) \\ D(i) < \min\left(D(i-1), D(i+1)\right) \end{matrix} \\ 0 & \quad\quad \text{otherwise} \end{cases} \qquad (16)$$

$$\mathbf{D_1} = nonzero\_samples(\mathbf{D'}),$$

while the coefficients $\mathbf{A_1}$ are selected from the coefficients **A** for those indices i for which $D'(i) = D(i)$, i.e. when the particular coefficient $D(i)$ is a local extreme.

4. EXPERIMENTAL RESULTS

In this Section, our experimental results are explained. The experiments are made with the signal given in *Fig. 1b* and the image in *Fig. 6a*. In both experiments, the non-decimated wavelet transform is performed by using NPR-QMF prototype filter, instead of wavelet filters.

For the case with 1-D signal the threshold is calculated in two ways: 1) by normalization of the coefficients **A** using all the samples in **A** and **D**; 2) by normalization of the coefficients **A** using $\mathbf{A_1}$ and $\mathbf{D_1}$ (the local extremes in **A** and **D**). A comparison between these two thresholds is given, and also the filtered wavelet coefficients are compared with the coefficients obtained from the noise-free signal. In *Fig. 3* the detail coefficients from the noise-free and the noisy signal are given. *Fig. 4* shows the detail coefficients |**D**| and two thresholds obtained with normalization of the coefficients **A** in (5) by using the coefficients **A** and **D** in (7), i.e. $\mathbf{A_1}$ and $\mathbf{D_1}$ in (13). From both graphs in *Fig. 4* it can be noticed that where the noise level is higher, the threshold is higher and vice versa. Similar results for the threshold are obtained when it is obtained through minimization of the error function in (15). Moreover, adding more terms in (10) and minimizing the error function in (14) instead that in (15) does not change the threshold a lot. When the parameter α is estimated through the normalization of the coefficients **A** in (7), the detail coefficients |**D**| (grey line) and the threshold $\boldsymbol{\tau}$ (black line) are given in *Fig. 4a* (in this case α is estimated from all 600 samples of **D** and **A**). In the case when the detail coefficients with values close to zero are discarded, the coefficients |**D**| and the estimated threshold $\boldsymbol{\tau}$ through (13) are given in *Fig. 4b* (in this case α is estimated from 425 samples of **D** and **A**). By comparing these two graphs it can be seen that the threshold in *Fig. 4b* is extended and closer to the peaks of the coefficients |**D**|, which means it is better generated than the threshold in *Fig. 4a*. This can be better noticed from *Fig. 5*, where the coefficients filtered by using these two thresholds are given. By comparing the filtrated coefficients in *Fig. 5a* and *5b* with the noise-free coefficients in *Fig. 3a*, it can be seen that filtrated coefficients in *Fig. 5b* are closer to the coefficients in *Fig. 3a*. Moreover, the use of detail coefficients from *Fig. 5b* in the signal reconstruction will result in distortion at the signal jumps.

In order to quantitatively compare the proposed method to some known wavelet based methods, we use the energy of the remained noise in the filtrated signal $\mathbf{s_1}$ as a measure:

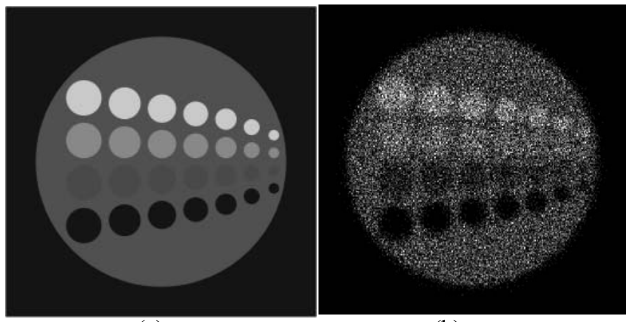
$$E_n = \sum_i \left(s(i) - s_1(i)\right)^2 \qquad (17)$$

Table 1 contains the results when the signal is reconstructed from the first level approximation and filtrated detail coefficients. It can be seen that when the proposed approach is applied, the noise energy is weaker compared to the other methods that use a global threshold. However, from *Fig. 1d, e* and *f* it can be noticed that reconstructed signals when a global threshold or multiscale product are used, suffer from distortion at the signal jumps positions, while there is no distortion at the signal filtrated with the proposed approach. This distortion appears as a result of removing signal information contained in the detail coefficients when a global threshold is used.

Another experiment is made with the image in *Fig. 6a*. Instead of adding non-linear noise followed by denoising the noisy image, the image in *Fig. 6a* is used for generating the new image given in *Fig. 6b*. The new image is generated in a way that each non-zero pixel from the image in *Fig. 6a* generates in its neighbourhood *m* pixels, where *m* depends on the pixel intensity. This is very close to the way of generating NM images, because each point source of radionuclide contributes with signal and noise in its neighbourhood. Afterwards, an estimation of the true signal in *Fig. 6b* is made by using the proposed method and some known wavelet based denoising methods. After filtering the image, the estimated image is normalized in order to have the same energy as the original image in *Fig. 6a*. The results are shown in *Table 2*. SNR_1 is signal-to-noise ratio for the generated image in *Fig. 6b*, while $\Delta\mathrm{SNR}_2$ is the improved signal-to-noise ratio. From *Table 2* it can be noticed the advantage of the proposed method over the other methods.

Proposed	Known methods				
	wavelet	universal [1]	[4]	sq.root [1]	[2]
1586	sym4	2011	2032	2071	1566
	sym7	2041	2094	2149	1714
	coif3	1958	1964	2076	1753
	coif5	1923	1959	2059	1847
	db2	2017	2040	2091	1614
	db8	1789	1795	1838	1787

Table 1. Comparison of the proposed method with known methods in case of denoising the 1-D signal from *Fig. 1b*



(a) (b)

Fig. 6. (a) Test image; (b) Generated image.

SNR_1 generat.	$\Delta\mathrm{SNR}_2$ propos.	$\Delta\mathrm{SNR}_2$ Known methods				
		wavelet	universal [1]	[4]	sq.root [1]	[2]
4.323	+4.384	sym3	+3.975	+1.329	+3.462	+3.309
		sym5	+3.947	+1.326	+3.458	+2.911
		db3	+3.975	+1.329	+3.462	+3.260
		coif5	+3.751	+1.325	+3.308	+3.159

Table 2. Comparison of the proposed with known methods in case of true signal estimating in the image in *Fig. 6b*

5. CONCLUSION

In this paper we propose a wavelet based method for denoising signals that contain signal-dependent noise. Instead of using a global threshold, which is inappropriate for filtering wavelet coefficients with varying noise level, we propose an estimate of a non-uniform threshold directly from the noisy signal.

REFERENCES

[1] D. L. Donoho, "Wavelet Thresholding and W.V.D.: A 10-minute Tour", *Int. Conf. on Wavelets and Applications*, Toulouse, France, June 1992;

[2] Y. Xu, J. B. Weaver, D. M. Healy, Jr., and J. Lu "Wavelet Transform Domain Filters: A Spatially Selective Noise Filtration Technique", *IEEE Trans. on Image Processing*, vol. 3, no. 6, pp. 747-758, Nov. 1994;

[3] E. J. Balster, Y. F. Zheng and R. L. Ewing "Feature-Based Wavelet Shrinkage Algorithm for Image Denoising", *IEEE Trans. on Image Processing*, vol. 14, no. 12, pp. 2024-2039, Dec. 2005;

[4] R. D. Nowak, R. G. Baraniuk, "Wavelet-Domain Filtering for Photon Imaging Systems", *IEEE Trans. on Image Processing*, vol. 8, Iss. 5, p. 666-678, May 1999;

[5] S. Bogdanova, M. Kostov, and M. Bogdanov, "Design of QMF Banks with Reduced Number of Iterations", *IEEE Int. Conf. on Signal Processing, Application and Technology, ICSPAT '99,* Orlando, USA, Nov. 1999.