Wavelet Thresholding by Using Multiscale Correlation

Cvetko Mitrovski and Mitko Kostov[1]

Abstract – **In this paper we present a technique for thresholding wavelet coefficients based on multiscale product analysis. We apply experiments both on noise-free and noisy signals; noise is signal dependent. Experimental results show that the noise is effectively removed.**

Keywords – **Denoising, threshold, wavelet domain filtering, multiscale products, signal-dependent noise.**

I. Introduction

Lately, there are many developed methods for image noise filtration in a transformation domain. In the last decade the stress on researches in this field is put on the signal processing in the wavelet domain. Wavelet transforms are multiresolution representations of signals and images. They decompose signals and images into multiscale details.

The reason of using the wavelet transform for denoising purposes is that the basis functions used in wavelet transforms are locally supported; they are nonzero only over part of the domain represented. Hence, adequately chosen wavelet basis groups the coefficients in two groups – one with a few coefficients with high SNR, and other with a lot of coefficients with low SNR. In addition, the signal power at large scales corresponds to that at low frequencies in the Fourier transform; the power at small scales corresponds to that at high frequencies in the Fourier transform.

In case of white Gaussian noise, the noise level is same through whole signal and for all the wavelet coefficients, independently on the signal. Hence, a well-estimated global threshold shrinks all the coefficients for an equal portion and removes the noise. But, in some signals, like nuclear medicine (NM) images, the noise level is proportional to the local signal intensity. Obviously, denoising them by using a global threshold is not the best solution.

In this paper we propose removing signal noise by exploiting wavelet multilevel correlation: the small scale data is passed at positions where the correlation is large and suppressed if the correlation is small. The paper is organized as follows. Section II briefly outlines wavelet theory. Section III discusses how to analyze the signal through the scale. Section IV verifies the validity of our approach on deterministic signals contaminated with signal dependent noise. At the end, Section V concludes the paper.

[1]Mitko Kostov and Cvetko Mitrovski are with the Faculty of Technical Sciences, I.L.Ribar bb, 7000 Bitola, Macedonia, E-mails: mitko.kostov@uklo.edu.mk, cvetko.mitrovski@uklo.edu.mk

II. Discrete Wavelet Shrinkage Method

The Discrete Wavelet Transform (DWT) decomposes a signal into a set of orthogonal components describing the signal variation across the scale [1]. The orthogonal components are generated by dilations and translations of a prototype function ψ called mother wavelet.

In analogy with other function expansions, a function f may be written for each discrete coordinate t as a sum of a wavelet expansion up to certain scale J plus a residual term, that is:

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

where ψ_{jk} is component obtained by dilation and translation of the mother wavelet. The approximation coefficients c_{Jk} contain the signal identity while the detail coefficients d_{jk} likely contain noise and need to be processed in order to remove the noise.

The most popular form of conventional wavelet-based signal filtering [1], can be expressed by:

$$\begin{aligned}&\{\mathbf{A}^{(k)},\mathbf{D}^{(1)},\mathbf{D}^{(2)},\Lambda\ ,\mathbf{D}^{(k)}\} = \mathrm{DWT}(\mathbf{s}+\mathbf{n}),\\ &\mathbf{s}^{*} = \mathrm{IDWT}\Big(f\big(\mathbf{A}^{(k)},\mathbf{h}^{(1)}.\times\mathbf{D}^{(1)},\mathbf{h}^{(2)}.\times\mathbf{D}^{(2)},\Lambda\ ,\mathbf{h}^{(k)}.\times\mathbf{D}^{(k)}\big)\Big)\end{aligned} \tag{2}$$

where $\mathbf{s}$ is noise-free signal, $\mathbf{n}$ is noise, $\mathbf{s}^*$ is filtered signal, $\mathbf{A}^{(k)}$ and $\mathbf{D}^{(k)}$ are approximation and detail coefficients at level k, respectively, f is a function of the modified detail and approximation coefficients, .× is element-by-element multi-plying and

$$\mathbf{h}^{(k)} = \left[h_1^{(k)}, h_2^{(k)}, \Lambda\ , h_j^{(k)}\right]^T \tag{3}$$

are weighting coefficients of the corresponding detail coefficients at level k.

In case of conventional hard threshold filtering the weighting coefficients are

$$h_j^{(k)}(hard) = \begin{cases} 1, & \text{if } \left|D_j^{(k)}\right| > \tau^{(k)} \\ 0, & \text{otherwise} \end{cases}, \tag{4}$$

while for the soft threshold filtering they are

$$h_j^{(k)}(soft) = \begin{cases} 1 - \dfrac{\tau^{(k)}\,\mathrm{sgn}\left(D_j^{(k)}\right)}{D_j^{(k)}}, & \text{if } \left|D_j^{(k)}\right| > \tau^{(k)} \\ 0, & \text{otherwise} \end{cases}, \tag{5}$$

where $\tau^{(k)}$ is user specified threshold for the k-th level details.

III. MULTISCALE CORRELATION

A. Algorithm

A noisy signal **x** can be expressed as a sum of noise-free signal **s** and noise **n**:

$$\mathbf{x} = \mathbf{s} + \mathbf{n}. \tag{6}$$

An orthogonal wavelet transformation of the noisy input yields an equivalent model in each wavelet sub-band:

$$\boldsymbol{\omega} = \boldsymbol{\beta} + \boldsymbol{\nu}. \tag{7}$$

where $\boldsymbol{\beta}$ are noise-free signal wavelet coefficients and $\boldsymbol{\nu}$ are noise wavelet coefficients.

For a given noise-free signal coefficient β_p (at position p) in a reference to a certain threshold τ_p, two hypotheses can be defined [2]:

$H_0: \ |\omega_p| \le \tau_p$, the signal of interest is absent (in the given coefficient);

$H_1: \ |\omega_p| > \tau_p$, the signal of interest is present (in the given coefficient).

We want to estimate the level of probability that the wavelet coefficient ω_p represents signal of interest. For this aim we want to exploit the correlation between the coefficients at the neighboring levels. This can be done by multiplying small scale coefficients and the coefficients derived from the larger scale, after the larger scale coefficients are shifted first. Hence, we define a shrinkage rule for estimating the coefficient $\hat{\beta}_p$:

$$\hat{\beta}_p = \omega_p \cdot P(H_1 \mid \omega_p)\,, \tag{8}$$

where

$$P = \begin{cases} 1 & \text{if } \left|\omega_p^{(k)}\right| \cdot \left|\lambda_p^{(k)}\right| \cdot \sqrt{\dfrac{\sum_j \left(\omega_j^{(k)}\right)^2}{\sum_j \left(\omega_j^{(k)} \lambda_j^{(k)}\right)^2}} > \left|\omega_p^{(k)}\right| \\ \left|\omega_p^{(k)}\right| \cdot \left|\lambda_p^{(k)}\right| \cdot \sqrt{\dfrac{\sum_j \left(\omega_j^{(k)}\right)^2}{\sum_j \left(\omega_j^{(k)} \lambda_j^{(k)}\right)^2}} & \text{otherwise} \end{cases} \tag{9}$$

$$\lambda_p^{(k)} = \max\left\{\Lambda\ \omega_{p-1}^{(k+1)}, \omega_p^{(k+1)}, \omega_{p+1}^{(k+1)}\ \Lambda\right\} \tag{10}$$

$\text{length}\{\boldsymbol{\lambda}\} = 4k$, k – decomposition level.

The coefficients $\boldsymbol{\lambda}^{(k)}$ in (10) are derived from the coarser level coefficients $\boldsymbol{\omega}^{(k+1)}$ and in this way we ensure that (shifted) maximums from the coarser level correspond to adequate spikes from the finer level. In addition, since the product coefficients $\boldsymbol{\omega}^{(k)}\boldsymbol{\lambda}^{(k)}$ in (9) have much bigger energy than the finer level coefficients $\boldsymbol{\omega}^{(k)}$, the coefficients $\boldsymbol{\omega}^{(k)}\boldsymbol{\lambda}^{(k)}$ are normalized. The criterion at (9) keeps a coefficient if it contains signal of interest or reduces it if not.

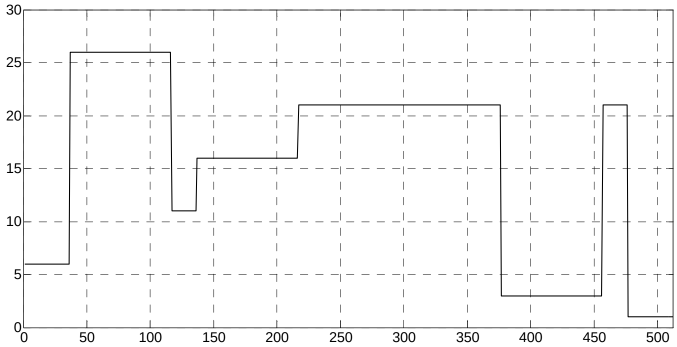

Fig. 1. Deterministic test signal.

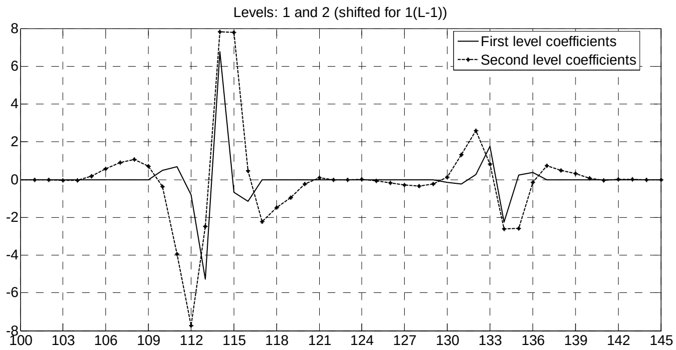

Fig. 2. Wavelet coefficients from 1st and 2nd level obtained by decomposing the signal with sym4 wavelet.

B. Analysis

If we consider signal decomposition in wavelet domain, starting from the finest level and moving through the scale, the signal energy increases and its shape slightly changes; the signal portions at the coarser scales contain more the signal identity than the portions at the finer scales. We want to use direct spatial correlations of the wavelet transform at different scales to identify the edges: the small scale data is passed at positions where the correlation is large and suppressed if the correlation is small. But, a small scale spike that appears at certain position stretches (dilates) in the next scale (coarser level), i.e. moves slightly around the considered positions in the next larger scale. This is shown by a wavelet decomposition of the deterministic noise-free signal shown in Fig. 1. The signal decompositions at different levels are carried out by using different wavelets and the results are shown in Figs. 2–5. Fig. 2 shows the signal decomposition carried out by wavelet sym4. The spikes at positions 77 and 78 at the first level are stretched in the second level, i.e. they are shifted to the positions 76 and 79. Similarly, the spikes at 497 and 498 (517 and 518) are shifted to positions 496 and 499 (516 and 519). Same results are obtained when other wavelets are used.

In addition, when wavelet coefficients from neighboring (k and k+1) levels are compared, the large level (k+1) coefficients are shifted for $k\times(L-1)$ samples, where L is the filter length.

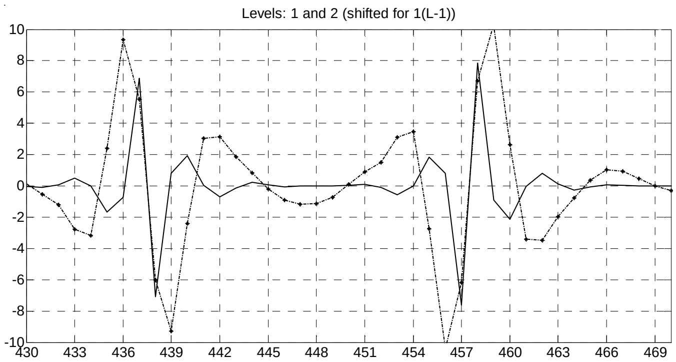

Fig. 3. Wavelet coefficients from 1st and 2nd level obtained by decomposing the signal with coif5 wavelet.

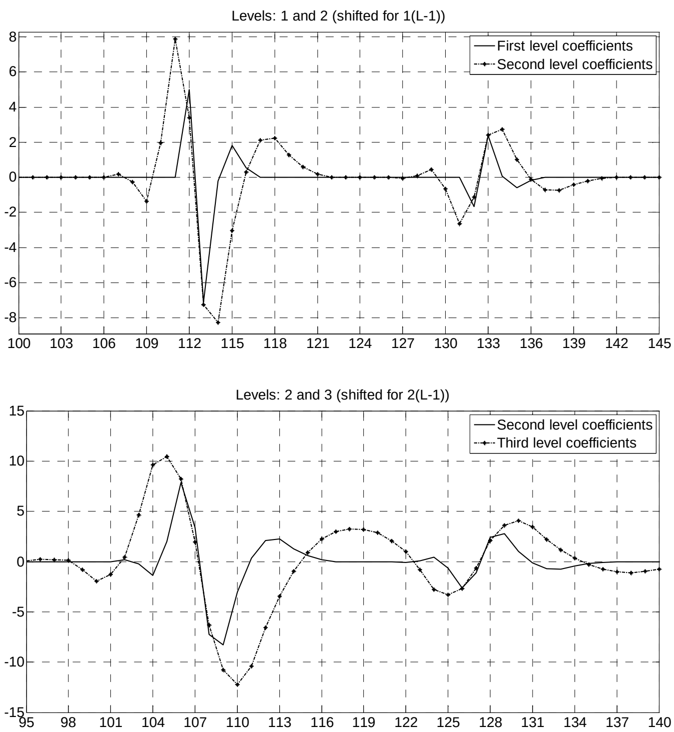

Fig. 4. Wavelet coefficients from 1st, 2nd and 3th level obtained by decomposing the signal with sym3 wavelet.

This correlation can be used to eliminate noise in the noisy signal. But, this filtering cannot be carried out by simple multiplying of the coefficients from the neighboring levels. We propose all the small scale coefficients to multiply by maximums in adequate large scale regions as it is shown in next section.

IV. EXPERIMENTS

In this Section, our experimental results are explained. The experiments are made with the noisy signal given in Fig. 6. The signal contains Poisson noise. It can be noticed in Fig 6 that there are signal information around the signal levels changes that should be preserved while filtering.

The results of filtering by using the proposed algorithm are shown in Figs. 7–9. It can be visually seen that the noise energy in the filtered signal is much lower than in the noisy signal. The filtered coefficients preserve the signal information (big spikes) and the small spikes are eliminated.

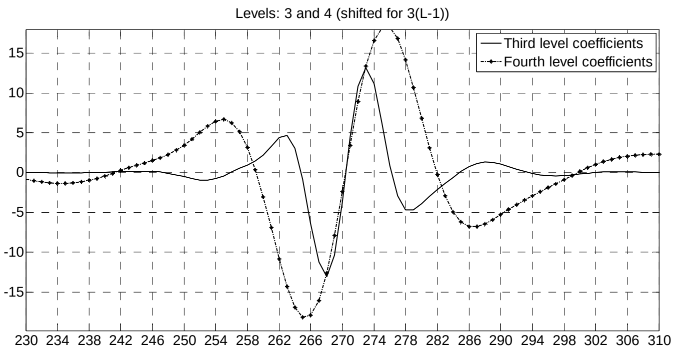

Fig. 5. Wavelet coefficients from 3th and 4th level obtained by decomposing the signal with coif5 wavelet.

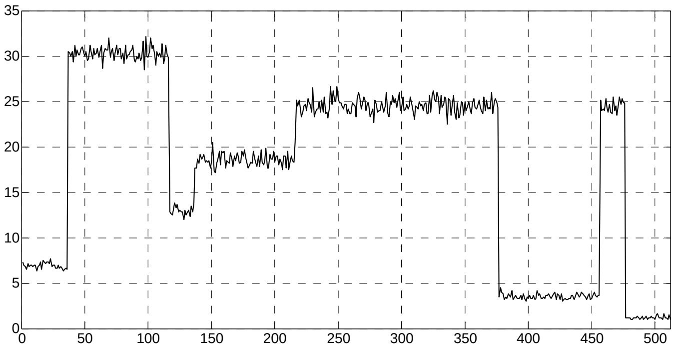

Fig. 6. Signal with noise.

MSE (Fig. 6)						
Wavelet	coif5		db4		sym3	
Level	1	2	1	2	1	2
Proposed	0.1446	0.1154	0.1419	0.1135	0.1382	0.1078
Multiscale	0.2703	0.3579	0.2041	0.2985	0.2411	0.4004
Universal	0.9077	1.8941	1.5997	2.0458	1.1990	4.1028

Table 1. Comparison of the proposed algorithm with other methods.

In order to quantitatively compare the proposed method to some known wavelet based methods, we use the mean squared error of the remained noise in the filtrated signal $\mathbf{s}_1$ (normalized to the energy of the noise free signal $\mathbf{s}$) as a measure:

$$MSE = \frac{1}{N}\sum_i \left(s(i) - s_1(i)\right)^2 \tag{11}$$

where N is the signal length.

The results of filtering of signal are shown in Table 1. The signal is reconstructed from the first and second level filtrated detail coefficients and second level approximation coefficients. The table shows that when the proposed approach is applied, the noise energy is weaker compared to filtering by Donoho's universal threshold [3, 4] and multiscale product algorithm [5].

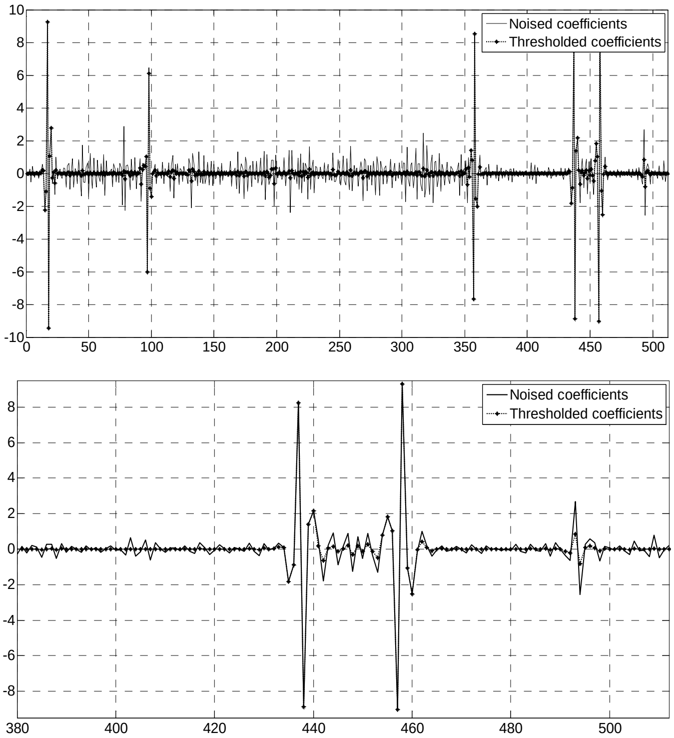

Fig. 7. Wavelet noised coefficients and thresholded coefficients at first level obtained with coif5.

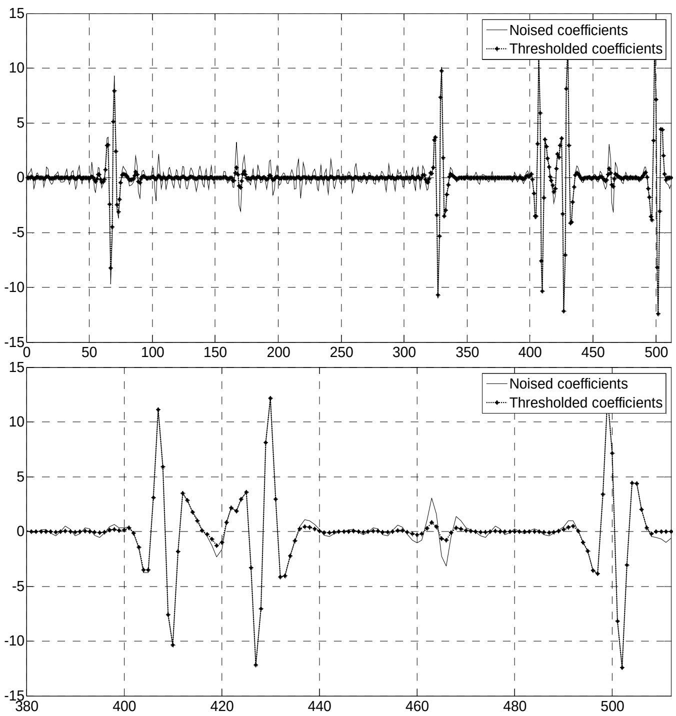

Fig. 8. Wavelet noised coefficients and thresholded coefficients at second level obtained with coif5.

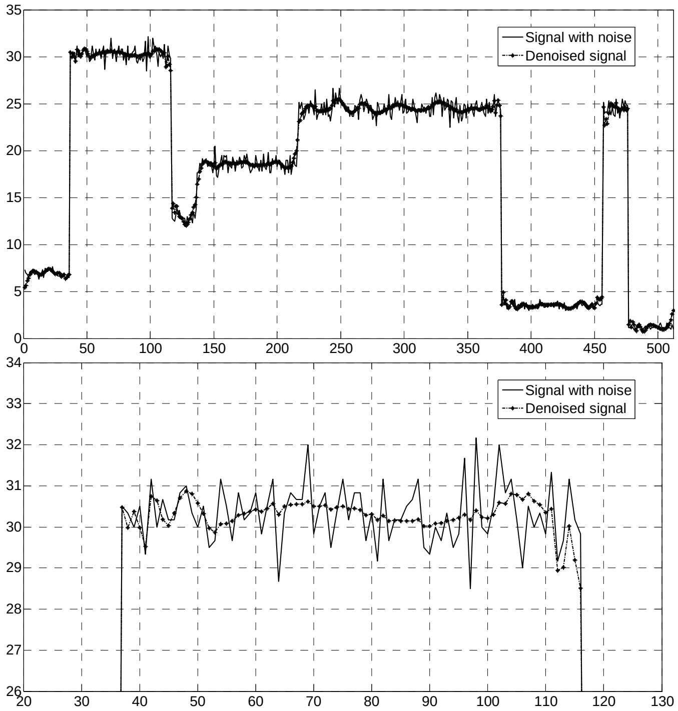

Fig. 9. Noise-free signal and denoised signal by using wavelet decomposition in two levels with coif5.

V. Conclusion

In this paper we propose removing signal noise by exploiting wavelet multilevel correlation: the small scale data is passed at positions where the correlation is large and suppressed if the correlation is small. Experimental results show that the noise is effectively removed.

References

[1] G. Strang, T. Nguyen, Wavelets and Filter Banks, Wellesley-Cambridge Press, 1996.

[2] A. Pižurica, W. Philips, “Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising”, *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 645–665, Mar. 2006.

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

[4] D.L. Donoho, I.M. Johnstone, "Ideal spatial adaptation via wavelet thresholding", *Biometrika*, vol. 81, pp. 425-455, 1994.

[5] Y. Xu, J.B. Weaver, D.M. Healy, Jr., 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.