WAVELET-BASED EMPIRICAL WIENER FILTERING

Jean-Paul G. Gallaire and Akbar M. Sayeed

Department of Electrical and Computer Engineering
University of Wisconsin–Madison
Madison, WI 53706
akbar@engr.wisc.edu gallaire@ese-metz.fr

ABSTRACT

Existing denoising scheme rarely use multiple-bases representations and if they do, they do not address the choice of the different bases. We present a new denoising scheme based on a multiple bases processing. The multiple bases used in the denoising algorithm are generated via unitary transforms. These unitary transforms also allow the construction of new wavelet families. In the resulting new domains, we apply a simple hard thresholding technique as well as a more complex Wiener filtering scheme. Preliminary results suggest that the resulting algorithms can deliver significantly improved performance over the undecimated wavelet transform without being computationnaly more expensive.

1. INTRODUCTION

In this research, we consider the problem of estimating a deterministic signal from noisy observations via wavelet-based processing. The ability of wavelets to concentrate the signal in relatively few large coefficients makes wavelet-based processing particularly attractive for signal estimation. Wavelet thresholding procedures exploit this energy compaction property for effective signal denoising, and are optimal in a minimax mean-square-error (MSE) sense for a variety of signal classes [1]. Nevertheless, for any given signal, Wiener filtering is optimal from an MSE viewpoint.

We approach the signal estimation problem from the perspective of designing the Wiener filter in the wavelet domain. Since the wavelet transform tends to decorrelate the data, in this study we restrict our attention to "diagonal" Wiener filters which process each coefficient independently. Wavelet thresholding techniques closely approximate the Wiener filter at large coefficients, whereas for smaller coefficients, whose magnitude is comparable to the noise standard deviation, thresholding procedures are far from the optimal Wiener weighting. Significant improvements in MSE are possible via weightings that are closer to the optimal Wiener filter [2].

The design of the Wiener filter in the wavelet domain requires the knowledge of noise-free signal coefficients which necessitates a data-adaptive or empirical approach that infers the filter directly from the noisy observations. The challenging part is the estimation of smaller signal coefficients. In [2], a curious design approach based on two wavelet transforms was proposed. The motivation for choosing two bases was to implicitly provide an estimate for the smaller coefficients from the larger ones. Large coefficients were estimated via wavelet thresholding. The thresholded signal was then transformed into a new basis to implicitly provide an estimate of the smaller coefficients via the smearing produced by the mismatch between the two bases [2]. The resulting signal estimates showed reduction in both bias and variance compared to thresholding methods, and produced MSE improvements up to a factor of 2. However, the critical issue of the choice of the second basis, given the first, was not addressed in [2].

In this paper, we extend the results of [2] along two directions. First, we provide a systematic approach to generating the second basis via a simple characterization of a class of unitary transforms. Second, we generalize the procedure to multiple bases which afford a richer diversity for signal estimation. Our approach is also related to the redundant wavelet transform denoising scheme proposed in [3], which can also be interpreted as a multiple-basis approach. Some other authors have also considered the idea of multiple-basis denoising recently [4].

In section 2, we present the basic ideas behind the use of multiple bases. Section 3 describes a systematic method for generating new bases, and new wavelet transforms. In section 4, we discuss possible improvements based on thresholding and Wiener filtering in multiple transform domains. Some experimental results are provided in Section 5. Section 6 concludes the paper with implications of this work for future research.

2. MULTIPLE-BASES DENOISING

Consider the problem of recovering samples of an unknown deterministic continuous-time signal $s(t)$, $t \in (0,1]$, from the set of noise-corrupted samples

$$x(i) \stackrel{def}{=} s(i/N) + n(i), \qquad i = 1, 2, \ldots, N, \qquad (1)$$

with $n(i)$ a zero-mean white Gaussian noise of variance σ^2. Let $\boldsymbol{x}$, $\boldsymbol{s}$, $\boldsymbol{n}$ denote $N \times 1$ column vectors containing the samples $x(i)$, $s(i/N)$, and $n(i)$, respectively, and let $\boldsymbol{W}$ denote an $N \times N$ orthonormal wavelet transform matrix. In the wavelet domain, (1) becomes

$$\boldsymbol{y} \stackrel{def}{=} \boldsymbol{\theta} + \boldsymbol{z}, \qquad (2)$$

with $\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x}$, $\boldsymbol{\theta} = \boldsymbol{W}\boldsymbol{s}$, and $\boldsymbol{z} = \boldsymbol{W}\boldsymbol{n}$. The goal is to estimate the true signal wavelet coefficients $\boldsymbol{\theta}$ given the noisy observations $\boldsymbol{y}$. Note that an orthonormal wavelet transformation maps $\boldsymbol{n}$ to a $\boldsymbol{z}$ that is also zero-mean white Gaussian with variance σ^2, while compacting typical signals $\boldsymbol{s}$ into a small number of large wavelet coefficients in $\boldsymbol{\theta}$. Thus a reasonable approach to wavelet-based signal estimation is to remove the small entries of $\boldsymbol{y}$ while retaining the large entries of $\boldsymbol{y}$. The motivation for processing the coefficients individually stems from the fact that the wavelet transform tends to decorrelate the data. This is the scheme introduced by Donoho [1] in which the wavelet coefficients of the noisy signal are processed through hard thresholding. The wavelet thresholding operation can be viewed as a diagonal filtering operation in the wavelet domain. Representing the filter by

$$\boldsymbol{H} \stackrel{def}{=} \mathrm{diag}[h(1), h(2), \ldots, h(N)], \qquad (3)$$

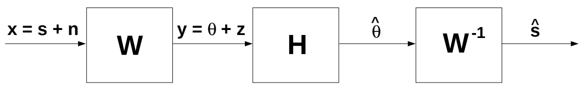

Figure 1. Wavelet-domain filtering using a diagonal weighting matrix $\boldsymbol{H}$.

we have the signal estimate

$$\widehat{\boldsymbol{s}} \stackrel{def}{=} \boldsymbol{W}^{-1}\boldsymbol{H}\,\boldsymbol{W}\,\boldsymbol{x}. \tag{4}$$

A block diagram for wavelet-domain filtering is given in Figure 1. The *hard threshold* wavelet filter $\boldsymbol{H}_h$ discards coefficients below a threshold value τ that is determined by the noise variance σ:

$$h_h(i) \stackrel{def}{=} \begin{cases} 1, & \text{if } |y(i)| > \tau \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

While hard thresholding optimal in a minimax mean-squared-error (MSE) sense over a class of signals, optimal minimum mean-squared-error (MMSE) processing for any given signal is achieved by the Wiener filter. Hard thresholding filter approximates the Wiener filter closely for large coefficients, but incurs errors at smaller coefficients. In [2], an implicit approach based on two wavelet bases was explored for designing the Wiener filter. The first basis was used to provide a thresholded estimate of the large coefficients, which were then fed to a slightly different second basis to estimate the smaller noisy coefficients from the reliable large coefficients. The inherent (slight) mismatch between the two slightly different bases spread the signal energy over a larger number of coefficients in the second basis, thereby implicitly yielding an estimate of the smaller coefficients in the second basis. Even though a mismatch is needed between the two bases, they both have to be sufficiently similar to provide compact representation of the signal at hand. However, the important issue of the design of the second basis was not addressed in [2].

In this paper, we present a systematic approach for generating the second basis which is sufficiently similar to the first one via a special class of unitary transforms. Furthermore, our approach facilitates the generation of an arbitrary number of closely related bases thereby enabling us to explore improved signal estimation by combining multiple wavelet domains. Similar techniques, based on hard thresholding in multiple wavelet domains, have been investigated in [4]. The basic motivation for multiple bases is that it can provide a richer or more diverse signal representation which may be leveraged into improved signal estimation. Finally we explore various filtering ideas in the multiple wavelet bases generated by our simple unitary transform scheme.

3. ROTATIONS AND MULTIPLE WAVELET BASES

In [2], the transform which takes the first estimate from the first wavelet domain to the second one is the composition of an inverse wavelet transform $\boldsymbol{W}_1^{-1}$ with another wavelet transform $\boldsymbol{W}_2$. The resulting transform $\boldsymbol{W}_1^{-1}\boldsymbol{W}_2$ is a unitary operator, or a rotation of coordinates. This suggests the generation of multiple bases via rotations.

Let $\boldsymbol{\Omega}$ denotes an $N \times N$ rotation. With the notations of section 2, K different signal estimates are given by:

$$\widehat{\boldsymbol{s}}_i \stackrel{def}{=} \boldsymbol{\Omega}^{-i}\,\boldsymbol{W}^{-1}\boldsymbol{H}_i\,\boldsymbol{W}\,\boldsymbol{\Omega}^i\,\boldsymbol{x} \qquad i = 1, 2, \ldots, K, \tag{6}$$

where $\boldsymbol{H}_i$ denotes the filters in the different domains.

The global estimate is then given by a simple average of the K different estimates

$$\widehat{\boldsymbol{s}} \stackrel{def}{=} \frac{1}{K}\sum_{i=1}^{K} \widehat{\boldsymbol{s}}_i \tag{7}$$

However, for long signals, the implementation of that generation scheme using matrices for rotations is not efficient due to the dimensionality of those matrices. We thus consider a simple scheme to generate rotations via LTI filters.

Any real-valued rotation $\boldsymbol{\Omega}$ can be expressed as $\boldsymbol{\Omega} = e^{\boldsymbol{A}}$ with $\boldsymbol{A}$ a skew-symmetric matrix. For the purpose of getting a simpler generating scheme using LTI systems, we will restrict ourselves to Toeplitz skew-symmetric matrices for $\boldsymbol{A}$. Note that for infinite dimensions matrices, $e^{\boldsymbol{A}}$ is also a Toeplitz matrix and hence represents an LTI filter.

$$\boldsymbol{y}(n) = (\boldsymbol{\Omega}\,\boldsymbol{x})(n) = (\omega * \boldsymbol{x})(n) \tag{8}$$

where ω is one row of the matrix $\boldsymbol{\Omega}$. Since $\boldsymbol{\Omega}$ is a unitary operator, the filter ω is an all-pass filter.

Furthermore if A has only a few non-zero diagonals close to the main diagonal and such that $\| A \|$ is small, each row of $e^{\boldsymbol{A}}$ will have only a few non-zero coefficients. Thus the corresponding rotation can be implemented using a filter with only a few taps. In this study, it will be constructed from a Toeplitz skew-symmetric matrix with only two non-zero diagonals in position 1 and -1. Thus $\boldsymbol{A}$ has the form

$$\begin{pmatrix} 0 & \alpha & 0 & \cdots \\ -\alpha & \ddots & \ddots & \ddots \\ 0 & \ddots & \ddots & \alpha \\ \vdots & \ddots & -\alpha & 0 \end{pmatrix} \tag{9}$$

Such a scheme to generate the multiple bases also represents a generalization of the undecimated wavelet transform also known as the shift-invariant wavelet transform [3]. A shift is a special case of rotation, thus the undecimated wavelet transform which combines the wavelet transforms of all the different shifts of a signal, can be considered as a particular case of the scheme presented here.

Note that the different bases generated via the above construction are not necessary wavelet bases. In fact, from (6), we see that that generating scheme corresponds to apply a pre-processing to (4). However, it would be interesting to ensure that the new bases keep the properties of wavelet bases.

A wavelet transform is either defined by the scaling and wavelet functions or by the decomposition and reconstruction filters [5]. Let $\mathbf{l_d}$ and $\mathbf{h_d}$ denote respectively the decomposition low-pass and high-pass filters of $\boldsymbol{W}$, $\mathbf{l_r}$ and $\mathbf{h_r}$ respectively the reconstruction low-pass and high-pass filters of $\boldsymbol{W}$, and ω and ω^{-1} the rotation filter and its inverse. Given those filters, a new wavelet transform $\boldsymbol{W}'$ can be constructed as follow:

$$\begin{array}{rclrcl} \mathbf{l}'_\mathbf{d} & = & \omega * \mathbf{l_d} & \mathbf{h}'_\mathbf{d} = & & \omega * \mathbf{h_d} \\ \mathbf{l}'_\mathbf{r} & = & \omega^{-1} * \mathbf{l_r} & \mathbf{h}'_\mathbf{r} = & & \omega^{-1} * \mathbf{h_r} \end{array} \tag{10}$$

Using this scheme repeatedly, it is easy to generate a new family of N wavelet transforms $(\boldsymbol{W}_i)_{i=0,\ldots,N-1}$, such that $\boldsymbol{W}_i$ is constructed with ω^i as in (10). As ω is an all-pass filter, it is easy to show that $\boldsymbol{W}_i$ keeps all the properties of an orthonormal wavelet transform. Furthermore by changing the parameter α in the matrix $\mathbf{A}$, one can control the mismatch between $\boldsymbol{W}_i$ and $\boldsymbol{W}_{i+1}$. If $\alpha \approx 0$, $\boldsymbol{\Omega}$ will be close to **I** and the mismatch will be very small. It is interesting to note that, within one family of N wavelets, the different transforms span the mismatch between $\boldsymbol{W}$ and $\boldsymbol{W}_{N-1}$.

These new wavelet families can be used to compute signal estimates from the noisy data. The main difference with (6) is that instead of rotating the signal before applying the same wavelet transform, we used different wavelet transform generated via rotations for each estimate. Thus our K different signal estimates are given by:

$$\widehat{\boldsymbol{s}}_i \stackrel{def}{=} \boldsymbol{W}_i^{-1}\boldsymbol{H}_i\boldsymbol{W}_i\boldsymbol{x} \qquad i = 1, 2, \ldots, K, \qquad (11)$$

where $\boldsymbol{H}_i$ denotes the filtering operation in the different $\boldsymbol{W}_i$ domains. The final estimate is still computed as in (7).

4. FILTERING SCHEMES IN MULTIPLE BASES

We now discuss the issue of designing the filter $\boldsymbol{H}_i$ in (6) and in (11) for the two multiple bases schemes.

Many denoising schemes have successfully used a hard thresholding scheme to perform denoising in wavelet bases [1],[4]. This filtering operation defined in (5) will be applied in the different domains obtained from (6) via rotation preprocessing (*Multiple-Rotation-Hard threshoding* (MRH)). Analogous to [4], hard thresholding will also be applied in the different wavelet domains defined in (11) (*Multiple-Wavelet-Hard threshoding* (MWH)).

We also investigate a Wiener filtering approach with our new schemes using the idea developed in [2]. The reason for that approach is that Wiener filtering canyield substantially improved esimates with fewer number of bases. It is constructed as follows: an estimate of the noise-free signal is first obtained in the i^{th} domain via equation (6), then this estimate $\widehat{\boldsymbol{s}}_i$ is transformed into the $(i+1)^{th}$ domain obtained via the rotation $\boldsymbol{\Omega}^{i+1}$ and the wavelet transform $\boldsymbol{W}$. From this estimate, a Wiener filter is computed and applied in the same domain. This scheme can be applied for any i, thus resulting in several estimates which are averaged to compute the final estimate (*Multiple-Rotation-Wiener filtering* (MRW)). The new wavelet family can also be used instead of $\boldsymbol{W}$ and $\boldsymbol{\Omega}$ as in 11 and the Wiener filter computed and applied in different wavelet domains(*Multiple-Wavelet-Wiener filtering* (MWW)).

5. EXPERIMENTAL RESULTS

Figures 2 and 3 show the influence on the MSE of two key parameters: the rotation and the number of bases. The results are shown for the MWH scheme applied to the Doppler and Heavisine test signals. In our four algorithms, the rotation parameter α controls the incremental "mismatch" between two successive bases. With the total number K of bases , it fixes the maximum "mismatch" at which the last estimate is computed. This maximum "mismatch" seems to be a key parameter to achieve the best MSE in the denoising schemes of a given signal. Any single estimate $\widehat{\boldsymbol{s}}_i$ computed with a bigger (respectively smaller) "mismatch" than the optimal maximum "mismatch" and included in (7) will increase (respectively decrease) the MSE of the global estimate $\widehat{\boldsymbol{s}}$. Decreasing the rotation parameter α increases the number of bases within the interval yielded by the optimal maximum "mismatch". Thus, the optimal values for these two parameters, α and K need to strike a tradeoff between efficiency and complexity. For instance, $\alpha = .75$ seems the best choice for the Doppler signal as it achieves a small MSE with few bases, but it gives an incremental "mismatch" already bigger than the optimal maximum "mismatch" in the case of the Heavisine signal. From those results and other experiments conducted on the Bumps test signal, smooth signals like Doppler seems to have a bigger maximum "mismatch" than less smooth signals.

As seen from the Tables 1 to 3, our schemes achieve at least comparable performance in denoising than the undecimated wavelet-transform scheme. Using the appropriate parameters, our schemes can achieve about the same results with less than 10 bases, offering computational complexity comparable to the redundant scheme (N Log N). The algorithm combining multiple bases and Wiener filtering (MWW) can further reduce the number of bases needed to achieve the optimal MSE, as seen from tables 4 and 5.

Finally, the first 256 points of the estimates resulting from different denoising schemes applied on a noisy Doppler signal (Figure 4) are shown in Figures 5 to 7. The real difference between the algorithms is their ability to restore the small high frequency coefficients. The MWW scheme (Figure 7) gives not only the best result for MSE but also provides estimates which are visually attractive.

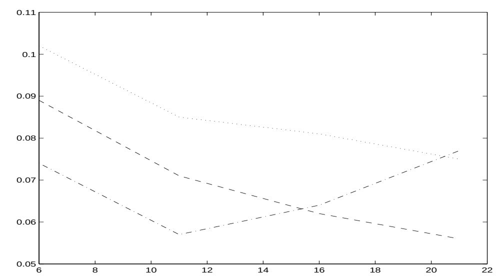

Figure 2. Influence of the Rotation ($\ldots \alpha = .1, -- \alpha = .25, -.- \alpha = .75$) and the number of bases on the MSE of the Doppler estimate.

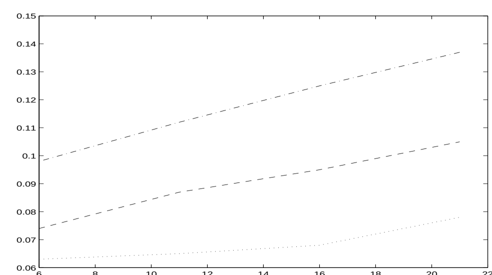

Figure 3. Influence of the Rotation ($\ldots \alpha = .1, -- \alpha = .25, -.- \alpha = .75$) and the number of bases on the MSE of the Heavisine estimate.

Table 1. Hard thresholding performance on "Doppler" test signal, $\boldsymbol{W}$ = symlet 4.

Estimator	MSE	$Bias^2$	Variance
Hard threshold	0.199	0.056	0.143
MRH(11 bases, $\alpha = 0.75$)	0.091	0.035	0.056
MRH(21 bases, $\alpha = 0.75$)	0.081	0.028	0.053
UndecimatedHard	0.104	0.042	0.062
MWH(6 bases, $\alpha = 0.75$)	0.074	0.032	0.042
MWH(11 bases, $\alpha = 0.75$)	0.058	0.019	0.039
Optimal Wiener Filter	0.069	0.014	0.055

6. CONCLUSION

We have presented new denoising algorithms based on the generation of multiple bases and wavelet transforms via unitary transforms. Given the appropriate parameters,

Table 2. **Hard thresholding performance on "Heavisine" test signal, W = symlet 4.**

Estimator	MSE	$Bias^2$	Variance
Hard threshold	0.132	0.035	0.097
MRH(11 bases, $\alpha = 0.75$)	0.055	0.027	0.028
MRH(21 bases, $\alpha = 0.75$)	0.056	0.029	0.027
UndecimatedHard	0.050	0.023	0.027
MWH(6 bases, $\alpha = 0.1$)	0.063	0.032	0.031
MWH(11 bases, $\alpha = 0.1$)	0.065	0.035	0.030
Optimal Wiener Filter	0.038	0.009	0.029

Table 3. Hard thresholding performance on "Bumps" test signal, W = Daubechies 4.

Estimator	MSE	$Bias^2$	Variance
Hard threshold	0.428	0.156	0.272
MRH(6 bases, $\alpha = 0.25$)	0.258	0.075	0.183
MRH(11 bases, $\alpha = 0.25$)	0.223	0.074	0.149
UndecimatedHard	0.218	0.045	0.173
MWH(11 bases, $\alpha = 0.1$)	0.222	0.074	0.148
MWH(16 bases, $\alpha = 0.1$)	0.204	0.057	0.147
Optimal Wiener Filter	0.179	0.020	0.159

these schemes can give better results than the undecimated wavelet transform denoising scheme with a comparable computational complexity. The automatic estimation of the parameters from the data is thus a key issue which will be investigated in the future. Another interesting research direction is to explore the properties of the wavelet-families presented in this paper and to construct rotation invariant wavelet transforms using cyclic rotations to generalize the shif-invariant wavelet transform.

REFERENCES

[1] D. Donoho and I. Johnstone, "Adapting to unknown smoothnes via wavelet shrinkage," *J. Amer. Stat. Assoc.*, vol. 90, pp. 1200–1224, Dec. 1995.

[2] S. P. Ghael, A. M. Sayeed, and R. G. Baraniuk, "Improved wavelet denoising via empirical Wiener filtering," in *Proc. SPIE 3169*, 1997.

[3] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells, "Nonlinear processing of a shift invariant dwt for noise reduction," in *Proc. SPIE Symp. on OE/Aerospace Sensing and Dual Use Photonics.*

[4] P. Ishwar, K. Ratakonda, P. Moulin, and N. Ahuja, "Image denoising using multiple compactions domains," in *Proc. ICASSP*, 1998.

[5] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, vol. XLI, pp. 909–996, 1988.

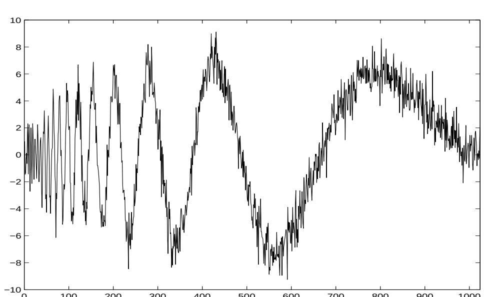

Figure 4. Doppler signal plus noise with $\sigma^2 = 1$

Table 4. Wiener filtering algorithms performance on "Doppler" test signal, W = symlet 4.

Estimator	MSE	$Bias^2$	Variance
WienerShrink from [2]	0.133	0.041	0.092
MWW(5 bases, $\alpha = .75$)	0.065	0.013	0.052
MWW(10 bases, $\alpha = .75$)	0.058	0.009	0.049
Optimal Wiener Filter	0.069	0.014	0.055

Table 5. Wiener filtering performance on "Heavisine" test signal, W = symlet 4.

Estimator	MSE	$Bias^2$	Variance
WienerShrink from [2]	0.075	0.034	0.041
MWW(5 bases, $\alpha = .25$)	0.054	0.018	0.036
MWW(10 bases, $\alpha = .25$)	0.055	0.019	0.036
Optimal Wiener Filter	0.038	0.009	0.029

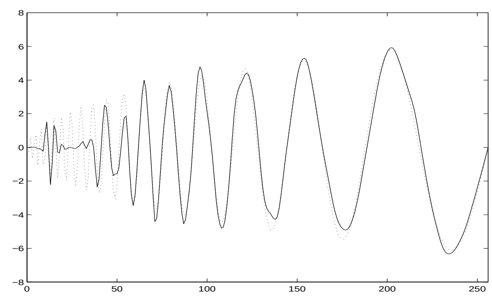

Figure 5. Result of the WienerShrink denoising scheme from [2] (Detail)

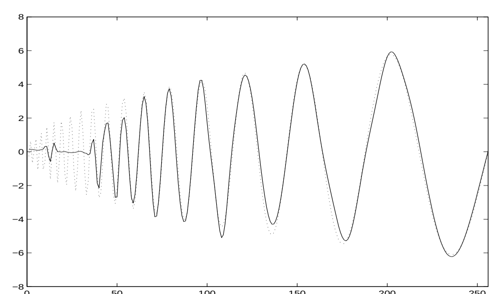

Figure 6. Result of the redundant wavelet transform denoising scheme (Detail)

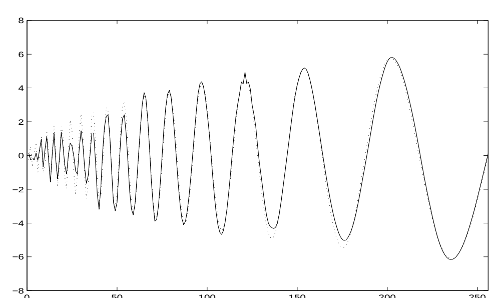

Figure 7. Result of the MultiWaveWiener denoising scheme (Detail)