Hindawi Publishing Corporation
EURASIP Journal on Advances in Signal Processing
Volume 2010, Article ID 641842, 14 pages
doi:10.1155/2010/641842

Research Article

Noiseless Codelength in Wavelet Denoising

Soosan Beheshti, Azadeh Fakhrzadeh, and Sridhar Krishnan

Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON, Canada M5B 2K3

Correspondence should be addressed to Soosan Beheshti, soosan@ee.ryerson.ca

Received 27 April 2009; Revised 14 October 2009; Accepted 4 January 2010

Academic Editor: Christoph Mecklenbräuker

We propose an adaptive, data-driven thresholding method based on a recently developed idea of Minimum Noiseless Description Length (MNDL). MNDL Subspace Selection (MNDL-SS) is a novel method of selecting an optimal subspace among the competing subspaces of the transformed noisy data. Here we extend the application of MNDL-SS for thresholding purposes. The approach searches for the optimum threshold for the data coefficients in an orthonormal basis. It is shown that the optimum threshold can be extracted from the noisy coefficients themselves. While the additive noise in the available data is assumed to be independent, the main challenge in MNDL thresholding is caused by the dependence of the additive noise in the sorted coefficients. The approach provides new hard and soft thresholds. Simulation results are presented for orthonormal wavelet transforms. While the method is comparable with the existing thresholding methods and in some cases outperforms them, the main advantage of the new approach is that it provides not only the optimum threshold but also an estimate of the associated mean-square error (MSE) for that threshold simultaneously.

1. Introduction

We can recognize different phenomena by collecting data from them. However, defective instruments, problems with the data acquisition process, and the interference of natural factors can all degrade the data of interest. Furthermore, noise can be introduced by transmission errors or compression. Thus, denoising is often a necessary step in data processing and various approaches have been introduced for this purpose. Some of these methods, such as Wiener filters, are grouped as linear techniques. While these techniques are easy to implement, their results are not always satisfactory. Over past decades, researchers have improved the performance of denoising methods by developing nonlinear approaches such as [1–6]. Although these approaches have succeeded in providing better results, they are usually computationally exhaustive, hard to implement, or use particular assumptions either on the noisy data or on the class of the data estimator.

Thresholding methods are alternative approaches to the denoising problem. The thresholding problem is first formulated in [7] where VisuShrink is introduced. This threshold is a nonadaptive universal threshold and depends only on the number of data points and noise variance. VisuShrink is a wavelet thresholding method which is both simple and effective in comparison with other denoising techniques. When an orthogonal wavelet basis is used, the coefficients with small absolute values tend to be attributed to the additive noise. Taking advantage of this property, finding a proper threshold, and setting all absolute values of coefficients smaller than the threshold to zero can suppress the noise. The main issue in such approaches is to find a proper threshold. In using the thresholding method for image denoising, the visual quality of the image is of great concern. An improper threshold may introduce artifacts and cause blurring of the image. One of the first soft thresholding methods is SureShrink [8] which has a better effect on the image than Visushrink in many cases. This method uses a hybrid of the universal threshold and the SURE (Stein's Unbiased Risk Estimator) threshold. The SURE threshold is chosen by minimizing Stein's estimate.

In this research we focus on the mean-square error (MSE) associated with the denoising process. The importance of this error in any signal reconstruction and estimation is inevitable [9, 10]. After all, in any estimation process, in this case reconstructing the denoised signal from the noisy one, the major goal is to achieve the original signal as much as possible, and the MSE is one of the most used criteria for evaluation purposes. To find the optimum threshold we estimate the MSE of a set of completing thresholds and choose the one that minimizes this error. The fundamentals of MSE estimation are similar to the method proposed in Minimum Noiseless Description Length (Codelength) Subspace Selection (MNDL-SS) [10]. Each subspace in this approach keeps a subset of the coefficients and discards the rest. Thresholding also produces a subspace that includes a set of coefficients that are being kept. On the other hand, it was shown in [10] that an estimate of noiseless description length (NDL) can be provided for each subspace by using the noisy data itself. We also show that in the process of estimating the NDL, the estimates of MSE is also provided. Furthermore comparison of the NDL of competing subspaces is equivalent to comparison of their MSEs. Therefore, the method presented in this paper is denoted by MNDL thresholding. The competing thresholds that are the sorted coefficients generate competing subspaces. For these subspaces the estimate of MSEs are provided and compared. The optimum threshold is associated with the subspace with minimum MSE (equivalently minimum NDL). Because of the particular choice of competing subspaces in MNDL thresholding, the effect of the additive noise is different from that in MNDL-SS. While the independence of the additive noise in MNDL-SS is the main advantage in estimating the desired NDL, in the case of thresholding, the additive noise is highly dependent. The main challenge in this work is to develop a method for NDL estimation acknowledging the presence of this noise dependence. We provide a threshold that is a function of the noise variance σ_w^2, the data length N, and the observed noisy data itself.

The paper is arranged as follows. Section 2 describes the considered thresholding problem. Section 3 briefly describes the fundamentals of the existing MNDL subspace selection. Section 4 introduces the MNDL thresholding approach. Hard and soft MNDL thresholdings are presented in Sections 5 and 6. Section 7 provides the simulation results and Section 8 is our conclusion.

2. Problem Statement

Noiseless data $\{\overline{y}(i),\ i = 1, \ldots, N\}$ of length N has been corrupted by an additive noise:

$$y(i) = \overline{y}(i) + w(i), \tag{1}$$

where $w(i)$ is an independent and identically distributed (i.i.d) Gaussian random process with zero mean and variance σ_w^2. (The method presented here is for real data. However, it can also be used for complex data.) In the considered denoising process, we project the noisy data into an orthogonal basis. The goal is to provide the optimum threshold for the resulting coefficients that minimizes the mean square error.

Assume that the noiseless data vector $\overline{\mathbf{y}}^N = [\overline{y}(1)\ \overline{y}(2) \cdots \overline{y}(N)]^T$ is generated by space S_N. The space S_N can be expanded by orthogonal basis vectors:

$$\left\langle \mathbf{s}_i, \mathbf{s}_j \right\rangle = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases} \tag{2}$$

where $\langle s_i, s_j \rangle$ is the inner product of vectors $\mathbf{s}_i$ and $\mathbf{s}_j$.

The data in (1) is represented in this basis as follows:

$$\mathbf{y}^N = \sum_{i=1}^{N} \theta(i)\mathbf{s}_i, \qquad \overline{\mathbf{y}}^N = \sum_{i=1}^{N} \overline{\theta}(i)\mathbf{s}_i, \qquad \mathbf{w}^N = \sum_{i=1}^{N} \nu(i)\mathbf{s}_i, \tag{3}$$

$$\theta(i) = \overline{\theta}(i) + \nu(i), \tag{4}$$

where $\overline{\theta}(i)$ is the ith coefficient of the noiseless data, $\theta(i)$ is the i-th coefficient of the noisy data, and $\nu(i)$ is the i-th coefficient of the additive noise. Note that since the basis vectors of S_N are orthogonal, $\nu(i)$ is also a sample of Gaussian distribution with zero mean and variance, σ_w^2.

The thresholding approach uses the available noisy coefficients, θ, to provide the best estimate of the noiseless coefficients denoted by $\hat{\theta}$. There are two general thresholding methods: hard and soft thresholdings. Hard thresholding eliminates or keeps the coefficients by comparing them with the threshold

$$\hat{\theta}(i) = \begin{cases} \theta(i), & \text{if } |\theta(i)| \geq T_h, \\ 0, & \text{otherwise}, \end{cases} \tag{5}$$

where T_h is the hard threshold. Soft thresholding eliminates the coefficients below T_s and reduces the absolute value of the rest of the coefficients:

$$\hat{\theta}(i) = \begin{cases} \operatorname{sgn}(\theta(i))(|\theta(i)| - T_s), & \text{if } |\theta(i)| \geq T_s, \\ 0, & \text{otherwise}. \end{cases} \tag{6}$$

There are different approaches for calculation of the proper T_h and T_s. In this paper, we expand the existing theory of the MNDL subspace selection method in [10] to provide new hard and soft thresholding methods.

Important Notation. In this paper a random variable is denoted by a capital letter, such as W and V, while a sample of that random variable is represented by the same letter in lower case such as w and v.

3. MNDL Subspace Selection (MNDL-SS)

The MNDL Subspace Selection (MNDL-SS) approach has been introduced in [10]. This approach addresses the problem of basis selection in the presence of a noisy data. In MNDL-SS, competing subspaces represent a projection of the noisy data on a complete orthogonal basis such as an orthogonal wavelet basis. each subspace contains a subset of the basis. The subspace keeps the coefficients of the noisy data in that subset and sets the rest of the coefficients to zero. Among competing subspaces, MNDL-SS chooses the subspace that minimizes the description length (codelength) of "noiseless" data. In this setting subspaces of the space S_N are chosen as follows: Each S_m is a subspace of S_N that is spanned by the first m elements of the bases. The estimate of the noiseless coefficients in subspace S_m is

$$\hat{\theta}_{S_m}(i) = \begin{cases} \theta(i), & \text{if } \mathbf{s}_i \in S_m, \\ 0, & \text{otherwise}, \end{cases} \quad (7)$$

and the estimate of the noiseless data in S_m, $\hat{\mathbf{y}}^N_{S_m}$, is

$$\hat{\mathbf{y}}^N_{S_m} = \sum_{i=1}^{N} \hat{\theta}_{s_m}(i)\mathbf{s}_i. \quad (8)$$

In each subspace the description length (codelength) of the noiseless data is defined as [10] (this criterion is different from MDL criterion and the explanation is provided in [10])

$$\mathrm{DL}\left(\overline{y}^N; \hat{\mathbf{y}}^N_{S_m}\right) = \log_2\sqrt{2\pi\sigma_w^2} + \frac{\log_2 e}{2\sigma_w^2} z_{S_m}, \quad (9)$$

where z_{S_m} is the reconstruction error (the equality of the error in the time domain and the error in coefficients of data is the result of the Parseval's Theorem) :

$$z_{S_m} = \frac{1}{N}\left\|\overline{\mathbf{y}}^N - \hat{y}^N_{S_m}\right\|_2^2 = \frac{1}{N}\left\|\overline{\theta}^N - \hat{\theta}^N_{S_m}\right\|_2^2, \quad (10)$$

which is a sample of random variable Z_{S_m}.

The optimum subspace $S_{m_{\text{opt}}}$ can be chosen by minimizing the average description length of noiseless data among the competing subspaces. Minimizing the average of noiseless data length in (9) is equivalent to minimizing the mean square error (MSE) in the form of $E(Z_{S_m})$ as the term $\log_2\sqrt{2\pi\sigma_w^2}$ is a constant and not a function of m:

$$m_{\text{opt}} = \arg\min_{S_m} E(Z_{S_m}). \quad (11)$$

MNDL-SS estimates the MSE for each subspace by using the available data error x_{S_m} in that subspace. The data error is defined in the following form:

$$x_{S_m} = \frac{1}{N}\left\|\mathbf{y}^N - \hat{\mathbf{y}}^N_{S_m}\right\|_2^2 = \frac{1}{N}\left\|\theta^N - \hat{\theta}^N_{S_m}\right\|_2^2, \quad (12)$$

which is a sample of random variable X_{S_m}. MNDL-SS studies the structure of the two random variables Z_{S_m} and X_{S_m} and uses the connection between these two random variables to provide an estimate of the desired criterion $E(Z_{S_m})$ for different m.

4. MNDL Thresholding

To use the ideas of subspace selection in thresholding, we first have to explain how a particular choice of subspaces serves the problem of thresholding. In MNDL-SS, the competing subspaces are chosen a priori and are not functions of the observed data. However, if the method is going to be used for thresholding, forming the competing subspaces is based on the observed data. In this case, we first sort the bases based on the absolute value of the observed noisy coefficients θ. This will dictate a particular indexing on the bases such that

$$|\theta(1)| \geq |\theta(2)| \geq \cdots \geq |\theta(N)|. \quad (13)$$

Therefore, the first subset represents the basis associated with the largest absolute value of the coefficients. The subset with two coefficients includes the two bases with the largest absolute value of the sorted coefficients and so on. The subspace S_m includes m of the basis and represents the first m largest absolute values of the coefficients and as a result this subspace represents thresholding with a threshold value of $\theta(m)$.

Back to the MNDL-SS, the subspace $S_{m_{\text{opt}}}$ that minimizes the average codelength of the noiseless data (equivalent to the subspace MSE in (11)) is the optimum subspace. Due to the indexing in the form of (13), the choice of this subspace results in the optimum threshold $\theta(m_{\text{opt}})$.

To estimate the MSE of the subspaces, we follow the fundamentals of the MNDL-SS method. Due to the random choice of subspaces in MNDL-SS, the random variables $V(i)$s that represent the additive noise of the coefficients in (4) are *independent* Gaussian random variables. In MNDL thresholding, the additive noise of coefficients is still Gaussian. However, due to the particular choice of the index for the coefficients, $V(i)$s are no longer independent. This will cause a major challenge in estimating the MSE of subspaces and is the main focus of this paper.

5. MNDL Hard Thresholding

In [10] it is shown that the expected value of the reconstruction error and that of the data error in subspace S_m can be written in the form of

$$E(Z_{S_m}) = E(A_{Z_{sm}}) + \frac{1}{N}\|\Delta_{S_m}\|_2^2, \quad (14)$$

$$E(X_{S_m}) = E(A_{X_{sm}}) + \frac{1}{N}\|\Delta_{S_m}\|_2^2, \quad (15)$$

where $\|\Delta_{S_m}\|_2$ is the l_2-norm of the discarded coefficient vector in subspace S_m:

$$\frac{1}{N}\|\Delta_{S_m}\|_2^2 = \frac{1}{N}\sum_{i=m+1}^{N} \overline{\theta}^2(i), \quad (16)$$

and the noise parts are

$$E(A_{Z_{sm}}) = \frac{1}{N}\sum_{i=1}^{m} E\left(V(i)^2\right), \quad (17)$$

$$E(A_{X_{sm}}) = \frac{1}{N}\sum_{i=m+1}^{N} E\left(V(i)^2\right), \quad (18)$$

where $V(i)$ is the Gaussian random variable with samples defined in (3).

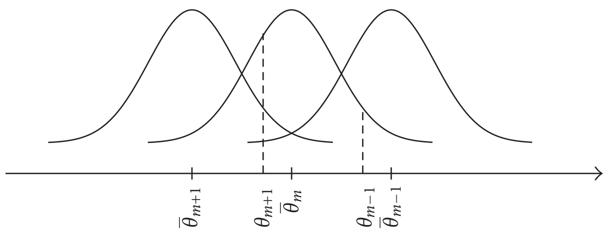

Figure 1: Distributions of $\theta(m+1)$, $\theta(m)$, and $\theta(m-1)$.

If the noise parts $E(A_{Z_{sm}})$ and $E(A_{X_{sm}})$ are available, then by estimating the expected value of the data error with the available sample x_{S_m} we have

$$\frac{1}{N}||\Delta_{S_m}||_2^2 \approx x_{S_m} - E(A_{X_{sm}}). \tag{19}$$

and then the estimate of the desired MSE in (14) is

$$E(Z_{S_m}) \approx x_{S_m} - E(A_{X_{sm}}) + E(A_{Z_{sm}}). \tag{20}$$

The main challenge in MNDL thresholding is in calculating $E(A_{X_{sm}})$ and $E(A_{Z_{sm}})$ in (18) and (17). In MNDL-SS, due to the independence of $V(i)$s in (4), random variables $A_{X_{sm}}$ and $A_{Z_{sm}}$ are Chi-square random variables and calculation of the expected values of these terms is straight forward. However, in MDNL thresholding, since the $V(i)$s are not independent, calculation of these expected values is not easy. In the following section we focus on estimating these desired expected values for the case of thresholding.

5.1. Estimate of MSE in MNDL Hard Thresholding. In order to calculate the expected value of the additive noise in (17) and (18), we need to study the noise effects that are associated with the sorted noisy coefficients. Each $\theta(i)$ in (4) has a Gaussian distribution with mean $\overline{\theta}(i)$ and variance σ_w^2. Figure 1 shows the distribution of $\theta(m+1)$, $\theta(m)$, and $\theta(m-1)$. The expected value of the noise part of $\theta(m)$ under the condition $\theta(m+1) < \theta(m) < \theta(m-1)$ is as follows:

$$E[V^2(m)] = \sigma_w^2 + \varepsilon_h(d_1(m), d_2(m)), \tag{21}$$

where

$$d_1(m) = \theta(m+1) - \overline{\theta}(m), \tag{22}$$

$$d_2(m) = \theta(m-1) - \overline{\theta}(m), \tag{23}$$

and $\varepsilon_h(d_1(m), d_2(m))$ is

$$\varepsilon_h(d_1(m), d_2(m)) = \frac{f_1(m) - f_2(m)}{Q(d_1(m)) - Q(d_2(m))}, \tag{24}$$

where $Q(x) = 1/2\pi\int_x^\infty \exp(-t^2/2)dt$ and $f_1(m)$ and $f_2(m)$ are defined as

$$f_1(m) = \frac{\sigma_w}{\sqrt{2\pi}} d_1(m) e^{-d_1^2(m)/2\sigma_w^2}, \tag{25}$$

$$f_2(m) = \frac{\sigma_w}{\sqrt{2\pi}} d_2(m) e^{-d_2^2(m)/2\sigma_w^2}. \tag{26}$$

Details of the calculationare provided in Appendix B.

Therefore, for the desired noise parts we have

$$E\left(n_{z_{S_m}}\right) = \frac{1}{N}\sum_{i=1}^{m} E(V^2(i)) \tag{27}$$

$$= \frac{m}{N}\sigma_w^2 + \frac{1}{N}\sum_{i=1}^{m}\varepsilon_h(d_1(i), d_2(i)). \tag{28}$$

Similarly we have

$$E\left(n_{x_{S_m}}\right) = \frac{1}{N}\sum_{i=m+1}^{N} E(V^2(i)) \tag{29}$$

$$= \frac{N-(m+1)}{N}\sigma_w^2 + \frac{1}{N}\sum_{i=m+1}^{N}\varepsilon_h(d_1(i), d_2(i)). \tag{30}$$

The noise part of the MSE in (27) and the noise part of the expected value of data error in (29) are dependent on d_1, d_2 that are not available. We suggest estimating them by using the available data as follows [11].

(i) Generate g_i Gaussian vectors of data length with variance of the additive noise.

(ii) Sort the absolute value of the associated noise coefficients, g_i^{sort}.

(iii) Find the estimate of the expected value of this vector $E[G^{\text{sort}}]$ by averaging over 50 samples of these vectors:

$$E[G^{\text{sort}}(m)] \approx \frac{1}{50}\sum_{i=1}^{50} g_i^{\text{sort}}(m). \tag{31}$$

(iv) Estimate $\overline{\theta}$ as follows:

$$\hat{\theta}(m) = \theta(m) - E[G^{\text{sort}}(m)]. \tag{32}$$

(i) Estimate d_1 and d_2 by replacing $\overline{\theta}$ with $\hat{\theta}$ in (22) and (23):

$$\hat{d}_1(m) = \theta(m+1) - \hat{\theta}(m), \tag{33}$$

$$\hat{d}_2(m) = \theta(m-1) - \hat{\theta}(m). \tag{34}$$

The estimates of the noise parts in (27) and (29) are

$$\hat{E}\left(n_{z_{S_m}}\right) = \frac{m}{N}\sigma_w^2 + \frac{1}{N}\sum_{i=1}^{m}\varepsilon_h\left(\hat{d}_1(i), \hat{d}_2(i)\right), \tag{35}$$

$$\hat{E}\left(n_{x_{S_m}}\right) = \frac{1}{N}\sum_{i=m+1}^{N}\varepsilon_h\left(\hat{d}_1(i), \hat{d}_2(i)\right) + \frac{N-(m+1)}{N}\sigma_w^2. \tag{36}$$

5.2. Calculating the Threshold. By using the provided noise part estimates in (35) and (36) we can estimate the desired MSE for subsets of different order with the following steps.

(i) Estimate the noise parts $E(n_{x_{S_m}})$ and $E(n_{z_{S_m}})$ using (36) and (35).

(ii) Estimate the MSE in (20) as follows:

$$\widehat{E}(Z_{S_m}) = x_{S_m} - \widehat{E}(A_{X_{sm}}) + \widehat{E}(A_{Z_{sm}}). \quad (37)$$

In MNDL thresholding the goal is to find m_{opt} by minimizing the MSE in (11). Here we provide an estimate of m_{opt} using the MSE estimate:

$$\widehat{m}_{\text{opt}} = \arg\min_{S_m} \widehat{E}(Z_{S_m}), \qquad T_h = \theta_{\widehat{m}_{\text{opt}}}. \quad (38)$$

6. MNDL Soft Thresholding

In some applications, such as image denoising, soft thresholding generally performs better and provides a smaller MSE than hard thresholding [12]. In soft thresholding, not only are the values smaller than the threshold set to zero, but also the value of coefficients larger than the threshold is also reduced by the amount of the threshold. Thus, we need to take into account this changing level of coefficients in MSE estimation. For MNDL soft thresholding we follow the same procedure as in MNDL hard thresholding. Here, the MSE in subspace S_m is

$$E(Z_{s_m}) = \frac{1}{N} E\left(\sum_{i=1}^{m} (V(i) - T_m)^2\right) + \frac{1}{N} \|\Delta_{S_m}\|_2^2, \quad (39)$$

where T_m is the smallest coefficient in subspace S_m (which is θ_m).

6.1. Noise Effects in the MSE. The noise part of the MSE in (39) is

$$E\left(n_{z_{s_m}}\right) = \frac{1}{N} \sum_{i=1}^{m} E(V(i) - T_m)^2, \quad (40)$$

where $V(i)$s are the associated noise parts of coefficients $\theta(i)$s in subspace S_m. The expected value of the noise part of $\theta(i)$ under the condition $\theta(i+1) < \theta(i) < \theta(i-1)$ is

$$E\left[(V(i) - T_m)^2\right] = T_m^2 + \sigma_w^2 + \varepsilon_s(d_1(i), d_2(i), T_m), \quad (41)$$

where $d_1(i)$ and $d_2(i)$ are defined similar to those in (22) and (23), and $\varepsilon_s(d_1(i), d_2(i), T_m)$ is defined as

$$\varepsilon_s(d_1(i), d_2(i), T_m) = \frac{j_1(i, T_m) + j_2(i, T_m)}{Q(d_1(i)) - Q(d_2(i))}, \quad (42)$$

where $j_1(i, m)$ and $j_2(i, m)$ are defined as

$$j_1(i, T_m) = \frac{\sigma_w}{\sqrt{2\pi}} e^{-d_1^2(i)/2\sigma_w^2} (d_1(i) - 2T_m), \quad (43)$$

$$j_2(i, T_m) = \frac{\sigma_w}{\sqrt{2\pi}} e^{-d_2^2(m)/2\sigma_w^2} (2T_m - d_2(i)). \quad (44)$$

Details of this calculation are provided in Appendix C.

Using the estimates of d_1 and d_2 from (33) and (34), the estimate of MSE's noise part in (40) is

$$\widehat{E}\left(n_{z_{S_m}}\right) = \frac{m}{N}(T_m^2 + \sigma_w^2) + \frac{1}{N} \sum_{i=1}^{m} \varepsilon_s \widehat{d}_1\left((i), \widehat{d}_2(i), T_m\right). \quad (45)$$

6.2. Estimate of the Noiseless Part of MSE. To complete the estimation of MSE in (39), we need also to estimate the noiseless part using the data error. The expected value of the data error in the case of soft thresholding is

$$E(X_{S_m}) = \frac{m}{N} T_m + \frac{1}{N} \|\Delta_{S_m}\|_2^2 + \frac{1}{N} \sum_{i=m+1}^{N} E(V^2(i)) \quad (46)$$

(Details are provided in Appendix A). The last component is the same as noise part in MNDL hard thresholding in (29) and can be estimated by using (36). Therefore, by estimating $E(X_{S_m})$ with its available sample x_{S_m}, from (46) we have

$$\frac{1}{N} \left\|\widehat{\Delta}_{S_m}\right\|^2 = x_{S_m} - \frac{m}{N} T_m - \frac{N - (m+1)}{N} \sigma_w^2 - \frac{1}{N} \sum_{i=m+1}^{N} \varepsilon_h\left(\widehat{d}_1(i), \widehat{d}_2(i), T_m\right). \quad (47)$$

where ε_h is defined in (24) and $\widehat{d}_1$ and $\widehat{d}_2$ are defined in (33) and (34).

6.3. Calculating the Threshold. The two components of MSE in (39) were estimated in previous sections. Therefore, the MSE can be estimated as follows.

(i) The noise part is estimated by using (45).

(ii) The noiseless part is estimated by using (47).

(iii) The estimate of the MSE in (39) is the sum of (45) and (47):

$$\widehat{E}(Z_{S_m}) = \frac{1}{N} \left\|\widehat{\Delta}_{S_m}\right\|^2 + \widehat{E}(A_{Z_{sm}}). \quad (48)$$

Similar to MNDL hard thresholding, the optimum subspace is the one for which the estimate of the MSE is minimized:

$$\widehat{m}_{\text{opt}} = \arg\min_{S_m} \widehat{E}(Z_{S_m}), \qquad T_s = \theta_{\widehat{m}_{\text{opt}}}. \quad (49)$$

6.4. Subband-Dependent MNDL Soft Thresholding. In image denoising, soft thresholding methods outperform hard thresholding methods in terms of MSE value and visual quality. In addition, it has been shown that subband-dependent thresholding performs better than universal thresholding methods [8]. In the subband-dependent method, a different threshold is provided for every subband of the wavelet transform. Here, we present the subband-dependent MNDL soft thresholding method. In every subband the MSE is estimated as a function of its subspaces. The subspace and its equivalent threshold that minimizes the MSE are chosen. The process of subband dependent MNDL thresholding with wavelet thresholds is as follows.

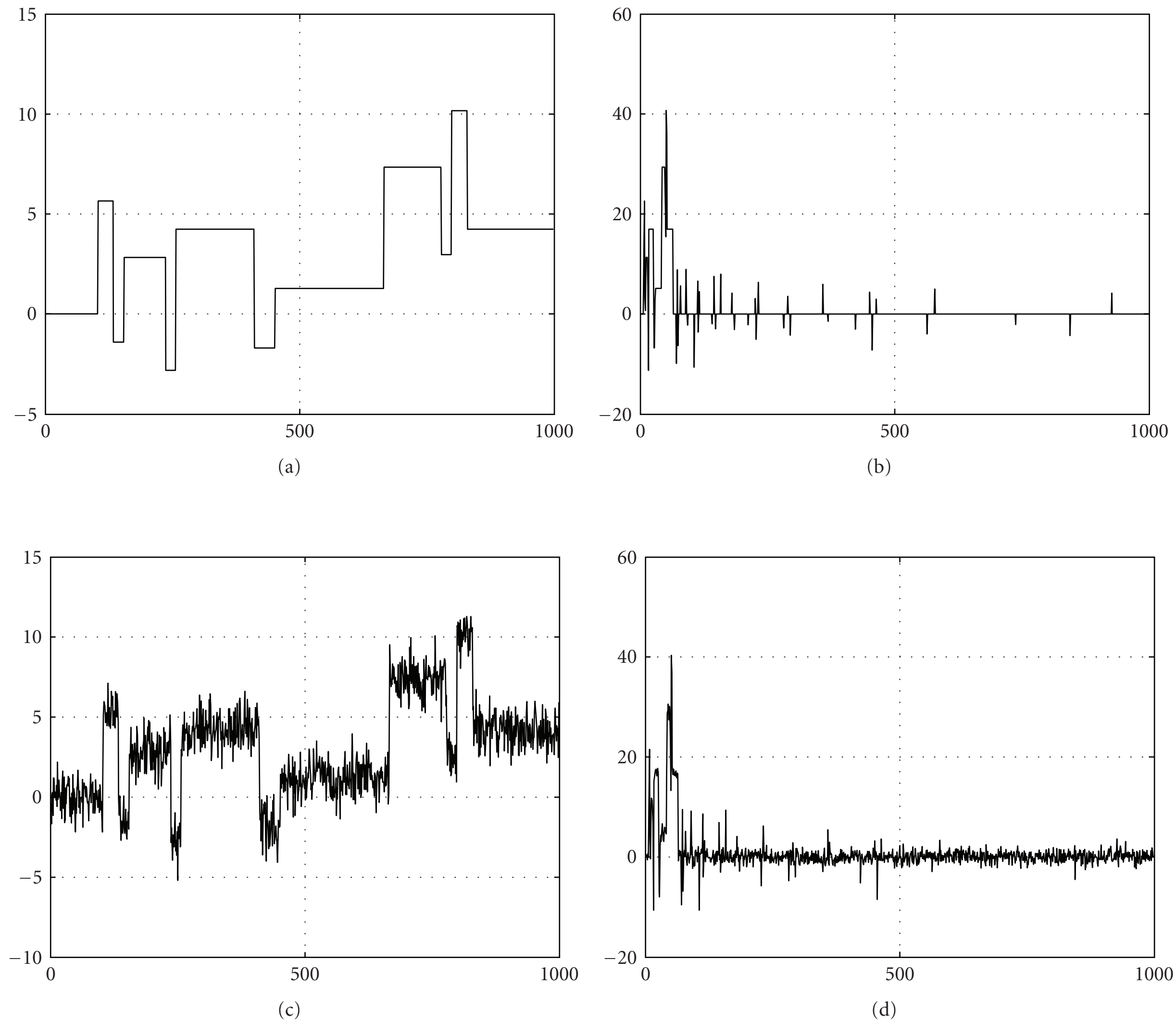

FIGURE 2: (a) Blocks signal with length 1024, (b) wavelet coefficients, (c) noisy blocks signal with additive white noise, $\sigma_w = 1$, and (d) noisy coefficients.

(i) The discrete wavelet transform of the image is taken.

(ii) In every subband the MSE, in (39), is estimated as a function of the S_m: the noise part is estimated using (45), and the noiseless part is estimated using (47). MSE estimate $\widehat{E}(Z_{S_m})$ is the sum of the noiseless part and the noise part estimates.

(iii) In each subband, the MSE is minimized over values of m, and $\widehat{m}_{\text{opt}}$ in (38) is chosen, where ($m \in \{1, 2, \ldots, N\}$), and N is the number of coefficients in the subband.

(iv) In each subband, the $\widehat{m}_{\text{opt}}$th largest absolute value of the coefficients is the optimum threshold.

(v) The image is denoised using the subband thresholds.

(vi) The inverse discrete wavelet transform is taken.

The unknown noise variance is estimated by the median estimator, $\widehat{\sigma}_n = \text{MAD}/0.675$, where MAD is the median of absolute value of the wavelet coefficients at the finest decomposition level (the diagonal direction of decomposition level one).

In the following section we provide simulation results of the method. The low complexity of this algorithm is an additional strength of the method. Our future plan is to utilize the approach for potential applications in areas such as biomedical engineering [13].

7. Simulation Results

We first demonstrate the performance of MNDL hard and soft thresholding by using two well-known examples in wavelet denoising. The first signal is the Blocks signal of length 1024 with few nonzero coefficients. The signal along with its wavelet coefficients is shown in Figure 2. Wavelet transform employs Daubechies's wavelet with eight vanishing moments with four scales of orthogonal decomposition

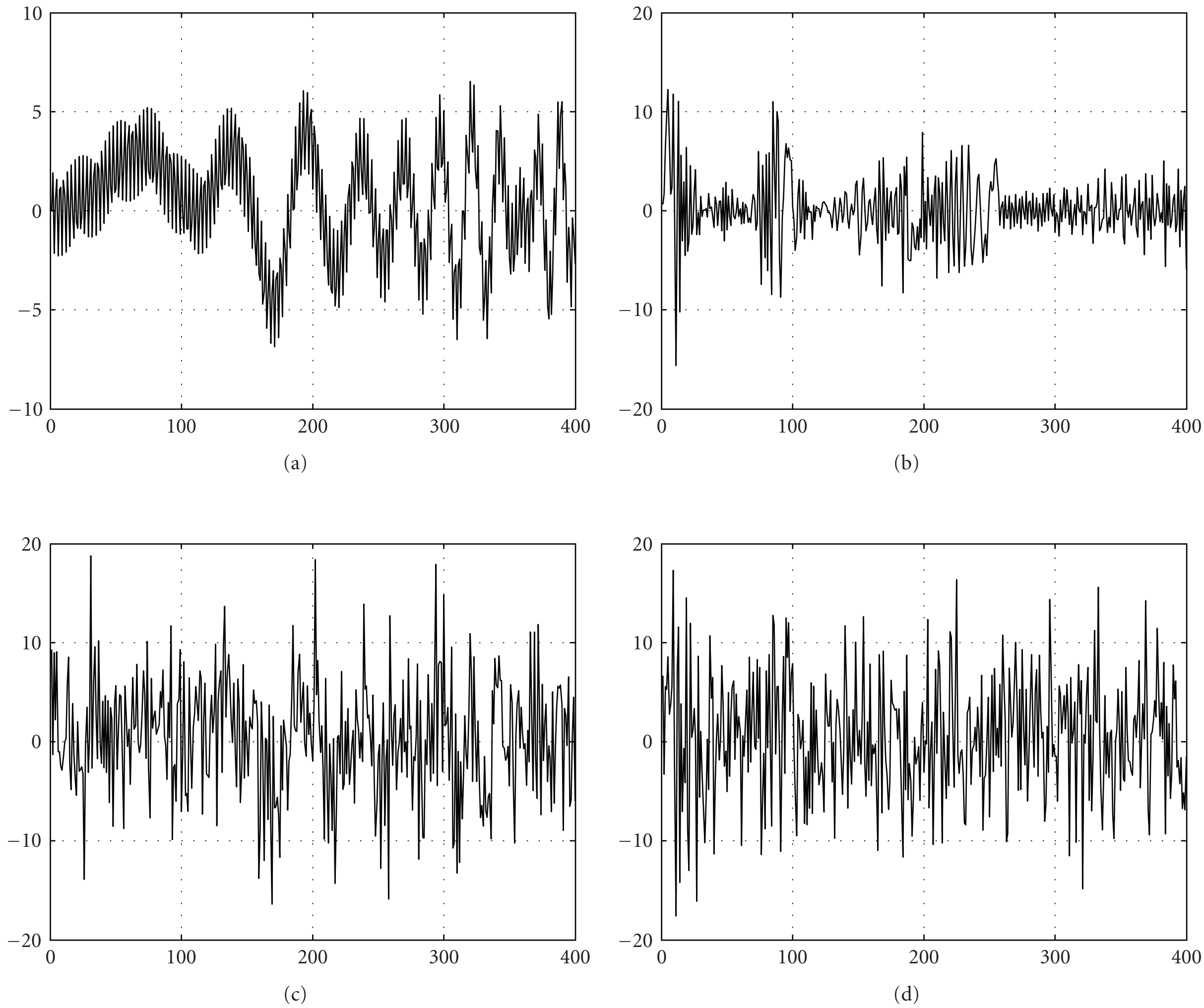

Figure 3: (a) Mishmash signal with length 1024, (b) wavelet coefficients, (c) noisy Mishmash signal with additive white noise, $\sigma_w = 5$, and (d) noisy coefficients.

Table 1: Comparing m_{opt} and its estimates using MNDL hard thresholding and MNDL-SS methods for the Blocks signal.

	m_{opt} Optimum order	$\widehat{m}_{\text{opt}}$ Hard thresholding	$\widehat{m}_{\text{opt}}$ MNDL-SS
$\sigma_w = 1$	72	74	91
$\sigma_w = 3$	34	32	40
$\sigma_w = 5$	22	19	37

[14]. The other signal is the Mishmash signal of length 1024 with no nonzero coefficients in Figure 3.

The MSE and its estimates using the existing MNDL-SS method [10] and the developed MNDL hard thresholding method are shown in Figure 4. The m_{opt} that minimizes the unavailable MSE and its estimate, $\widehat{m}_{\text{opt}}$, with these approaches are provided in Table 1 for different noise variances. The results in this table and the rest of the results in this section are averages of five runs. As the figure and the table show, as was expected MNDL thresholding outperforms the MNDL-SS approach.

Table 2 compares the MSE of the proposed hard thresholding method with that of two hard thresholding methods, VisuShrink and MDL. The comparison includes the optimum hard MSE, which represents the minimum MSE when the noisy coefficients are used as hard thresholds, along with the resulting MSE of different approaches. As the table shows, in most cases, MNDL hard thresholding provides the minimum MSE among the approaches.

The MSE and its estimate with MNDL soft thresholding are shown in Figure 5. The results in this figure are for Blocks and Mishmash signals and for two different levels of the additive noise. As the figure shows, MSE estimates are very close to the MSE itself.

The MSE results for soft threshoding are compared in Table 3. The table provides optimum MSE with both optimum thresholding and optimum subband thresholding along with the results for Sureshrink and MNDL soft thresholding. As the table shows for Blocks the MNDL

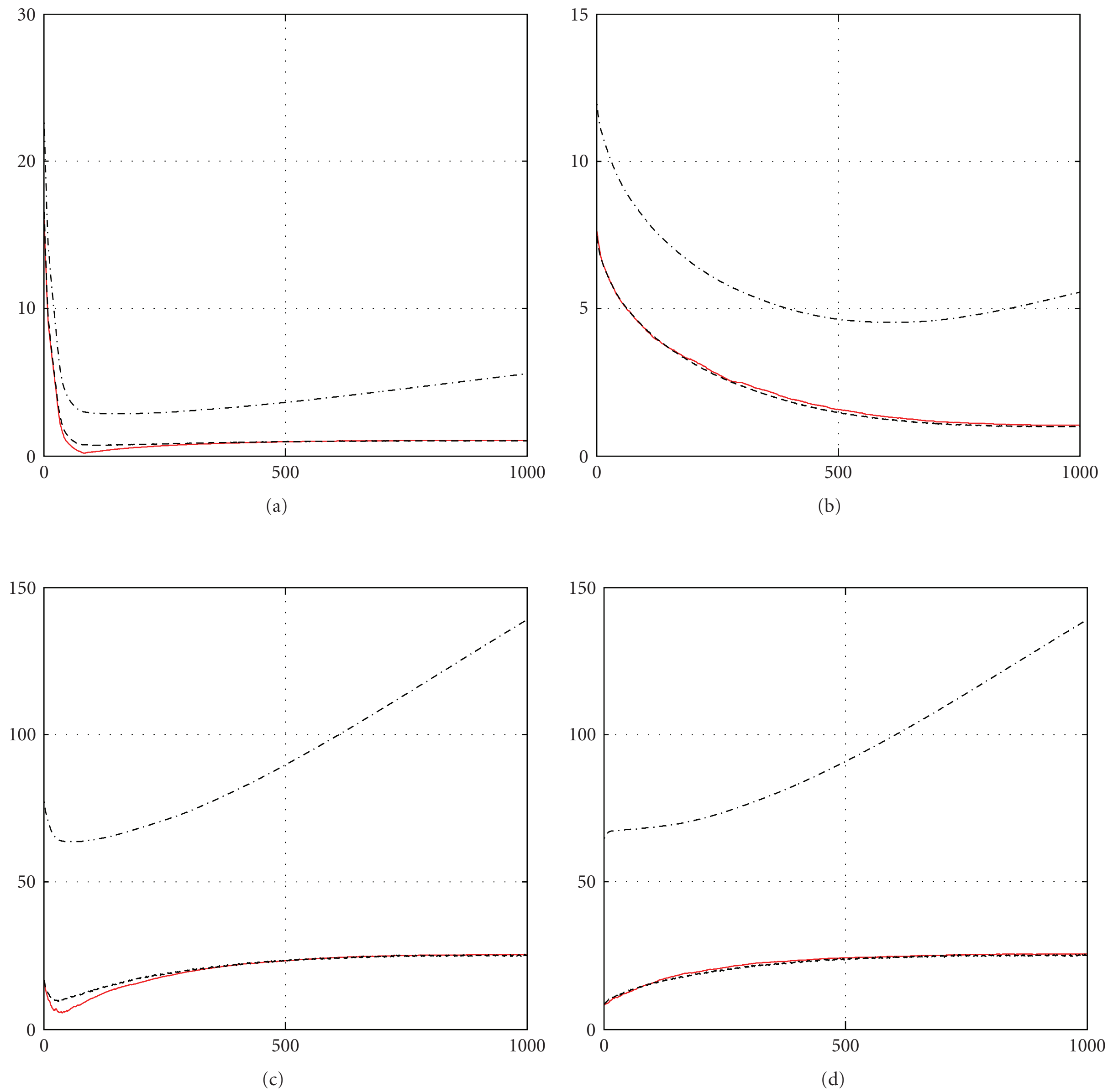

Figure 4: Desired unavailable MSE (solid red line), and its estimate using MNDL hard thresholding (- -) and MNDL-SS (. -) as a function of m. (a) Blocks signal with $\sigma_w = 1$, (b) Blocks signal with $\sigma_w = 5$, (c) Mishmash signal with $\sigma_w = 1$, and (d) Mishmash signal with $\sigma_w = 5$.

subband-dependent has the best results while for Mishmash MNDL soft thresholding outperforms the other approaches in almost all cases. While here we have shown the simulation results for two of six test signals in [8], the results for the other two signals (Heavy sine, Doppler) are similar to those of the provided signals.

7.1. Image Denoising. There are many image denoising approaches, such as recent work in [15, 16]. These approaches have succeeded in providing good results. They usually use a particular assumption either on the noisy image and/or on the class of the data estimator. A well-known image denoising thresholding approach is BayesShrink. BayesShrink [12] is a thresholding method that is also widely used for image denoising. This method attempts to minimize the Bayes' Risk Estimator function assuming a prior Generalized Gaussian Distribution (GGD) and thus yields a data adaptive threshold [17]. Note that our method does not make any particular assumption on the data and is not especially proposed for image denoising. Here we use the approach for images as an example of a class of two-dimensional data.

To explore the application of MNDL soft thresholding in image denoising we use four images: Cameraman (a sample of a soft image), Barbara (a sample of a highly detailed image), Lena, and Peppers. These images are shown

Table 2: MSE comparison of different thresholding methods with the MNDL hard thresholding approach.

Blocks	Optimum hard MSE	MNDL-SS	MNDL hard thresholding	MDL	VisuShrink
$\sigma_w = 1$	0.13	0.3	**0.18**	0.2	0.2
$\sigma_w = 3$	1.8	2.8	**1.5**	2.1	2.2
$\sigma_w = 6$	7.9	9.3	**9**	**9**	9.2
$\sigma_w = 10$	14	17.6	17.2	17.7	**15.4**
Bumps					
$\sigma_w = 1$	0.3	0.33	**0.32**	0.33	0.37
$\sigma_w = 3$	1.32	1.54	**1.4**	1.68	1.5
$\sigma_w = 6$	4.16	6.1	5	5.2	**4.9**
$\sigma_w = 10$	9.3	11.3	**10.59**	15.7	10.63
Quadchirp					
$\sigma_w = 1$	0.96	1	**.98**	1.4	2.29
$\sigma_w = 3$	6.62	4.7	**6.75**	7	6.9
$\sigma_w = 6$	7.7	8.9	**7.8**	10.51	8.01
$\sigma_w = 10$	7.86	12	**7.86**	15.9	8.5
Mishmash					
$\sigma_w = 1$	0.9	1.2	**0.9**	2	3.3
$\sigma_w = 3$	7.1	7.8	7.5	7.4	**7.3**
$\sigma_w = 6$	7.8	8	**7.8**	10	7.9
$\sigma_w = 10$	7.86	8.2	**7.86**	16	8

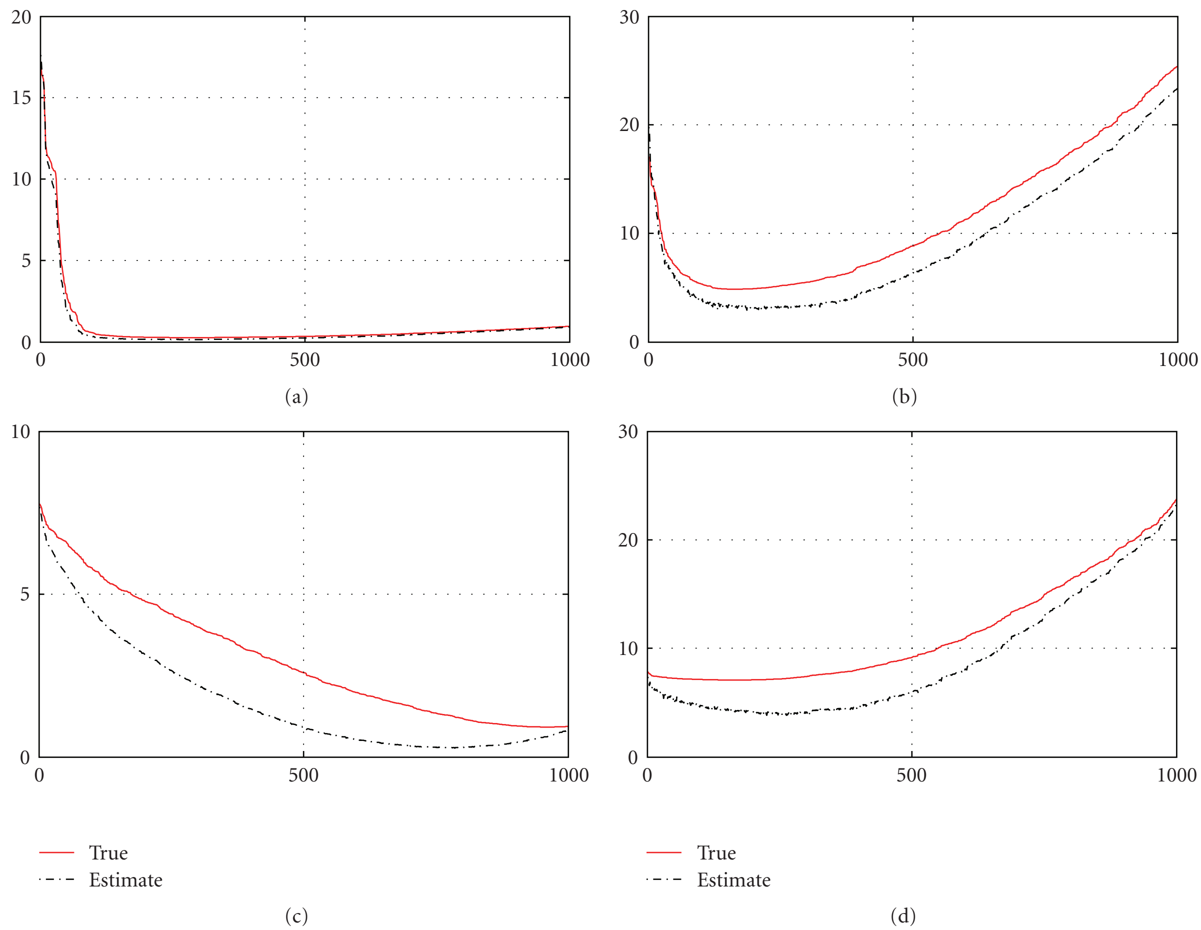

Figure 5: The MSE and its estimate in MNDL soft thresholding as a function of m. (a) Blocks signal, $\sigma_w = 1$, (b) Blocks signal with $\sigma_w = 5$, (c) Mishmash signal with $\sigma_w = 1$, and (d) Mishmash signal with $\sigma_w = 5$.

Table 3: MSE for (1) Optimum soft thresholding; (2) Optimum Subband-dependent soft thresholding; (3) MNDL soft thresholding, (4) MNDL Subband-dependent soft thresholding, (5) Sureshrink.

Blocks	1	2	3	4	5
$\sigma_w = 1$	0.3	0.18	0.3	**0.21**	0.27
$\sigma_w = 3$	2.1	1.13	2.2	**1.23**	1.35
$\sigma_w = 6$	6.5	3.2	6.9	3.8	**3.3**
$\sigma_w = 10$	12.1	7.9	12.5	**8**	9.7
Bumps					
$\sigma_w = 1$	0.3	0.23	0.3	**0.26**	0.3
$\sigma_w = 3$	1.6	1.13	1.6	**1.32**	1.5
$\sigma_w = 6$	4.2	3.2	4.3	**3.8**	3.9
$\sigma_w = 10$	8	7.7	**8.3**	9.9	8.6
Quadchirp					
$\sigma_w = 1$	0.82	0.83	1	**0.9**	0.91
$\sigma_w = 3$	4.5	4.7	**4.6**	4.8	6.5
$\sigma_w = 6$	7.2	8.1	**7.5**	8.9	8.5
$\sigma_w = 10$	7.8	13.19	**9.03**	15	13.5
Mishmash					
$\sigma_w = 1$	0.9	0.95	**1.2**	1.3	**1.2**
$\sigma_w = 3$	4.8	5	**4.9**	5.1	7
$\sigma_w = 6$	7.3	8.5	**7.5**	9.1	9
$\sigma_w = 10$	7.7	12.7	**8.8**	14.5	12.8

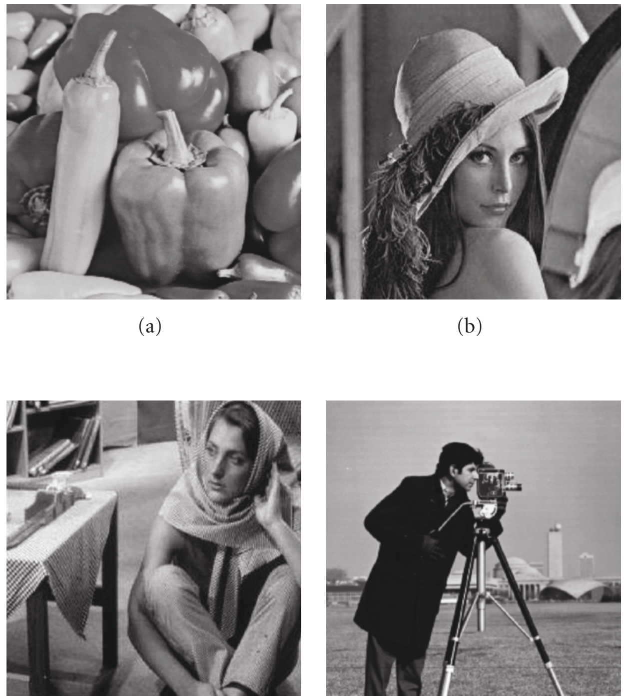
(a) (b) (c) (d)

Figure 6: Test images. From top left, clockwise: Peppers, Lena, Cameraman, and Barbara.

in Figure 6 and with size 512×512. The wavelet transform employs Daubechies's wavelet with eight vanishing moments and with four scales of orthogonal decomposition.

In Table 4, we compare the MSE of the MNDL with two well-known thresholding methods: BayesShrink [12] and SureShrink [8]. All these thresholds are soft subband-dependent. As the table shows, MNDL thresholding performs better than SureShrink in most cases and is comparable with the BayesShrink. The MNDL soft thresholding is compared visually with BayesShrink and SureShrink in Figure 7. As the figure shows, the ringing effect at the edges of the image with the MNDL soft thresholding is less than that with the BayesShrink approach. The importance of the new approach is that it can provide an estimate of MSE simultaneously. Note that it can also provide estimate of MSE for other thresholding methods as follows. Find the closest absolute value of the coefficients to the given threshold and use the index m of that coefficient and check the estimate of MSE for the associated S_m. Table 5 shows the optimum subband threshold for Cameraman and Table 6 shows the thresholds for BayesShrink and MNDL. As the tables show, the thresholds of MNDL are slightly larger than the optimum ones. On the other hand the Bayes thresholds are smaller than the optimum ones, especially for the coarsest level. While the MSE at this noise level is almost the same for these methods, the thresholds indicate that MNDL keeps fewer coefficients compared to BayesShrink and its threshold is much closer to the optimum one.

8. Conclusion

We proposed thresholding method based on the MNDL-SS approach. This approach uses the available data error to provide an estimate of the desired noiseless codelength for comparison of competing subspaces. In this approach, the statistics of the data error plays an important role. Unlike MNDL-SS, in MNDL thresholding the involved additive noises of the error are highly dependent. The main challenge of this work was to estimate the desired criterion in the presence of such dependence. We developed a method to estimate the desired criterion for the purpose of thresholding.

Experimental results show that the proposed MNDL hard thresholding method outperforms VisuShrink most of the time and the proposed MNDL soft thresholding method outperforms SureShrink most of the time. Application of MNDL soft thresholding for image denoising is also explored and the method proves comparable with the BayesShrink approach. Unlike the image denoising approaches, with MNDL thresholding no assumption on the structure of the signal is necessary. The main advantage of the method is in estimating both the desired noiseless description length and the mean square error (MSE). The calculated MSE estimate provides a quantitative quality measure for the proposed threshold simultaneously. An additional strength of the approach is its ability to estimate the MSE for any given threshold and it can be used for quality evaluation of any thresholding method.

Table 4: MSE for various images with (1) Optimum soft MSE, (2) MNDL soft thresholding, (3) BayesShrink, and (4) SureShrink. Averaged over five runs.

Cameraman	Optimum soft MSE	MNDL thresholding	BayesShrink	SureShrink
$\sigma_w = 5$	11.3	**12**	13.3	12.6
$\sigma_w = 10$	32.5	**34.6**	35.7	62
$\sigma_w = 15$	58	**61**	62	85
$\sigma_w = 20$	83.4	**87**	89.2	102
Barbara				
$\sigma_w = 5$	14.5	16.8	**15.6**	21
$\sigma_w = 10$	42	**46.4**	51	56
$\sigma_w = 15$	74	80	**79**	86
$\sigma_w = 20$	109.4	**115.2**	121.6	121.2
Lena				
$\sigma_w = 5$	18.2	20	**18.7**	19.3
$\sigma_w = 10$	28	**30**	30.6	**30**
$\sigma_w = 15$	51	52.9	**51.5**	54.7
$\sigma_w = 20$	54.8	**58**	61.6	64
Peppers				
$\sigma_w = 5$	16.4	18.1	17.85	**17.45**
$\sigma_w = 10$	39.4	**42**	44.5	43.8
$\sigma_w = 15$	61.4	**64.95**	67	75.95
$\sigma_w = 20$	82.65	**87**	87.7	94.9

Table 5: Optimum threshold of subbands for Cameraman and $\sigma_w = 10$.

Level	LH	HH	HL
1 (finest)	10.4	13.9	9.9
2	6.4	8.7	6.9
3	5	5.7	6.6
4	3	4.6	4

Table 6: (a): MNDL threshold, (b): BayesShrink threshold of subbands for Cameraman and $\sigma_w = 10$.

(a)

Level	LH	HH	HL
1 (finest)	12.5	13.9	11.8
2	10.8	11.7	11.2
3	9.3	9.9	9.4
4	5	5.3	4.7

(b)

Level	LH	HH	HL
1 (finest)	8.2	13.8	6
2	3.8	6	2.5
3	1.7	3.1	1.3
4	1	1.5	0.6

Appendices

A. X_{S_m} in Hard and Soft Thresholding

The expected value of the data error in hard thresholding, in (12), is

$$E(X_{S_m}) = \frac{1}{N} E\left[\sum_{i=m+1}^{N} \left(\overline{\theta}(i) + V(i)\right)^2\right] \tag{A.1}$$

$$= \frac{1}{N}\sum_{i=m+1}^{N} \overline{\theta}^2(i) + \frac{1}{N}\sum_{i=m+1}^{N} E(V^2(i)) + \frac{2}{N}\sum_{i=m+1}^{N} E\left(\overline{\theta}(i)V(i)\right). \tag{A.2}$$

Since the noiseless coefficients $\overline{\theta}(i)$s are independent of the noise part $V(i)$s, the third term becomes zero and we conclude with (15).

The expected value of the data error in soft thresholding is

$$E(X_{S_m}) = \frac{1}{N}\left(\sum_{i=1}^{m} E(T_m^2) + \sum_{i=m+1}^{N} E\left(\left(\overline{\theta}(i) + V(i)\right)^2\right)\right) \tag{A.3}$$

$$= \frac{m}{N} T_m^2 + \frac{1}{N}\sum_{i=m+1}^{N} E\left(\left(\overline{\theta}(i) + V(i)\right)^2\right). \tag{A.4}$$

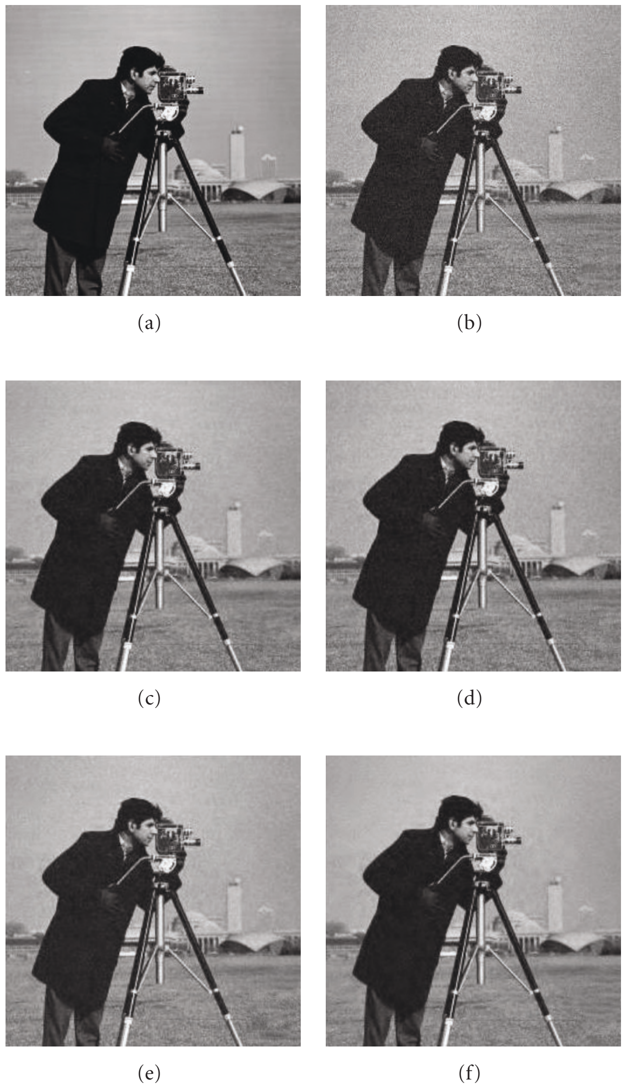

Figure 7: (a) Noiseless image, (b) Noisy image with $\sigma_w = 15$, (c) Optimum soft thresholding, (d) Bayeshrink, (e) Sureshrink, (f) MNDL soft thresholding.

The second part of the expected value of the data error in (A.4) is the same as the expected value of the data error in the hard thresholding in (A.1). Therefore, (A.4) can be written in the form of (46).

B. Calculating $E[V^2(m)]$ in MNDL Hard Thresholding

The distribution of noise coefficients is a Gaussian one and we have $V(m) \sim (0, \sigma_w^2)$. On the other hand, for a sorted version of coefficients we have $\theta(m+1) < \theta(m) < \theta(m-1)$ while $\theta(m) = \nu(m) + \overline{\theta}(m)$. Therefore, the following extra condition holds on the noise coefficients:

$$\theta(m+1) - \overline{\theta}(m) < \nu(m) < \theta(m-1) - \overline{\theta}(m). \tag{B.1}$$

Under the above condition, the desired conditional expected value is

$$\begin{aligned} &E\left[V^2(m) \mid \theta(m+1) - \overline{\theta}(m) < V(m) < \theta(m-1) - \overline{\theta}(m)\right] \\ &\quad = \frac{\tau}{\Pr\left(\theta(m+1) - \overline{\theta}(m) < V(m) < \theta(m-1) - \overline{\theta}(m)\right)}, \end{aligned} \tag{B.2}$$

where the numerator τ is

$$\begin{aligned} \tau &= \int_{\theta(m+1)-\overline{\theta}(m)}^{\theta(m-1)-\overline{\theta}(m)} y^2 \frac{1}{\sqrt{2\pi}\sigma_w} e^{-y^2/2\sigma_w^2} dy \\ &= \frac{\sigma_w}{\sqrt{2\pi}}\left[\theta(m-1) - \overline{\theta}(m)\right] e^{-\frac{\left[\theta(m-1)-\overline{\theta}(m)\right]^2}{2\sigma_w^2}} \\ &\quad - \frac{\sigma_w}{\sqrt{2\pi}}\left[\theta(m+1) - \overline{\theta}(m)\right] e^{-\frac{\left[\theta(m+1)-\overline{\theta}(m)\right]^2}{2\sigma_w^2}} \\ &\quad + \sigma_w^2\left[Q\left(\frac{\left[\theta(m-1)-\overline{\theta}(m)\right]}{\sigma_w}\right) \right. \\ &\qquad \left. - Q\left(\frac{\left[\theta(m+1)-\overline{\theta}(m)\right]}{\sigma_w}\right)\right], \end{aligned} \tag{B.3}$$

and the denominator is

$$\begin{aligned} &\Pr\left(\theta(m+1) - \overline{\theta}(m) < V(m) < \theta(m-1) - \overline{\theta}(m)\right) \\ &\quad = \kappa_1 - \kappa_2, \end{aligned} \tag{B.4}$$

where κ_1 and κ_2 are

$$\begin{aligned} \kappa_1 &= \int_{\theta(m+1)-\overline{\theta}(m)}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_w} e^{-y^2/2\sigma_w^2} dy \\ &= Q\left(\frac{\left[\theta(m+1)-\overline{\theta}(m)\right]}{\sigma_w}\right), \end{aligned} \tag{B.5}$$

$$\begin{aligned} \kappa_2 &= \int_{\theta(m-1)-\overline{\theta}(m)}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_w} e^{-y^2/2\sigma_w^2} dy \\ &= Q\left(\frac{\left[\theta(m-1)-\overline{\theta}(m)\right]}{\sigma_w}\right). \end{aligned} \tag{B.6}$$

The conditional expectation of $E[(V^2(m))]$ in (B.2) can be simplified to

$$E\left[V^2(m)\right] = \sigma_w^2 + \varepsilon_h(d_1(m), d_2(m)), \tag{B.7}$$

where $d_1(m)$, $d_2(m)$, and $\varepsilon_h(d_1(m), d_2(m))$ are defined in (22), (23), and (24).

Any $E[V^2(m)]$ in the paper is the conditional expectation in (B.2). For notation simplicity, the condition of the expectation is eliminated throughout the paper.

C. Calculating $E[(V(i) - T_m)^2]$ in MNDL Soft Thresholding

Similar to calculating the conditional expected value of $E(V^2(i))$ for MNDL hard thresholding, under the same condition $\theta(i+1) - \overline{\theta}(i) < v(i) < \theta(i-1) - \overline{\theta}(i)$, we have

$$\begin{aligned} E\Big[(V(i) - T_m)^2 \mid \theta(i+1) & \\ -\overline{\theta}(i) < V(i) < \theta(i-1) - \overline{\theta}(i)\Big] & \\ = \frac{\mu}{\Pr\Big(\theta(i+1) - \overline{\theta}(i) < V(i) < \theta(i-1) - \overline{\theta}(i)\Big)}, & \end{aligned} \tag{C.1}$$

where the denominator is provided in (B.4) and the numerator is

$$\begin{aligned} \mu &= \int_{\theta(i+1)-\overline{\theta}(i)}^{\theta(i-1)-\overline{\theta}(i)} (\gamma - T_m)^2 \frac{1}{\sqrt{2\pi}\sigma_w} e^{-\gamma^2/2\sigma_w^2} d\gamma \\ &= \int_{\theta(i+1)-\overline{\theta}(i)}^{\theta(i-1)-\overline{\theta}(i)} \gamma^2 \frac{1}{\sqrt{2\pi}\sigma_w} e^{-\gamma^2/2\sigma_w^2} d\gamma \\ &\quad + \int_{\theta(i+1)-\overline{\theta}(i)}^{\theta(i-1)-\overline{\theta}(i)} T_m^2 \frac{1}{\sqrt{2\pi}\sigma_w} e^{-\gamma^2/2\sigma_w^2} d\gamma \\ &\quad - \int_{\theta(i+1)-\overline{\theta}(i)}^{\theta(i-1)-\overline{\theta}(i)} 2\gamma T_m \frac{1}{\sqrt{2\pi}\sigma_w} e^{-\gamma^2/2\sigma_w^2} d\gamma. \end{aligned} \tag{C.2}$$

There are three integrals in μ. The first integral is

$$\begin{aligned} &\int_{\theta(i+1)-\overline{\theta}(i)}^{\theta(i-1)-\overline{\theta}(i)} \gamma^2 \frac{1}{\sqrt{2\pi}\sigma_w} e^{-\gamma^2/2\sigma_w^2} d\gamma \\ &= \frac{\sigma_w}{\sqrt{2\pi}} \left[\theta(i-1) - \overline{\theta}(i)\right] e^{-\frac{\left[\theta(i-1) - \overline{\theta}(i)\right]^2}{2\sigma_w^2}} \\ &- \frac{\sigma_w}{\sqrt{2\pi}} \left[\theta(i+1) - \overline{\theta}(i)\right] e^{-\frac{\left[\theta(i+1) - \overline{\theta}(i)\right]^2}{2\sigma_w^2}} \\ &+ \sigma_w^2 \left[Q\left(\frac{\left[\theta(i-1) - \overline{\theta}(i)\right]}{\sigma_w}\right) - Q\left(\frac{\left[\theta(i+1) - \overline{\theta}(i)\right]}{\sigma_w}\right)\right]. \end{aligned} \tag{C.3}$$

The second integral is

$$\begin{aligned} &\int_{\theta(i+1)-\overline{\theta}(i)}^{\theta(i-1)-\overline{\theta}(i)} T_m^2 \frac{1}{\sqrt{2\pi}\sigma_w} e^{-\gamma^2/2\sigma_w^2} d\gamma \\ &= T_m^2 \left[Q\left(\frac{\left[\theta(i+1) - \overline{\theta}(i)\right]}{\sigma_w}\right) - Q\left(\frac{\left[\theta(i-1) - \overline{\theta}(i)\right]}{\sigma_w}\right)\right], \end{aligned} \tag{C.4}$$

and the third integral is

$$\begin{aligned} &\int_{\theta(i+1)-\overline{\theta}(i)}^{\theta(i-1)-\overline{\theta}(i)} 2\gamma T_m \frac{1}{\sqrt{2\pi}\sigma_w} e^{-\gamma^2/2\sigma_w^2} d\gamma \\ &= \frac{2T_m\sigma_w}{\sqrt{2\pi}} \left[e^{-\frac{\left[\theta(i+1) - \overline{\theta}(i)\right]^2}{2\sigma_w^2}} - e^{-\frac{\left[\theta(i-1) - \overline{\theta}(i)\right]^2}{2\sigma_w^2}} \right]. \end{aligned} \tag{C.5}$$

The numerator of (C.1) is calculated by adding up (C.3), (C.4) and (C.5). Therefore, the simplified version of $E[(V(i) - T_m)^2]$ in (C.1) is

$$E\Big[(V(i) - T_m)^2\Big] = T_m^2 + \sigma_w^2 + \varepsilon_s(d_1(i), d_2(i), T_m), \tag{C.6}$$

where $d_1(m)$, $d_2(m)$, and $\varepsilon_s(d_1(i), d_2(i), T_m)$ are defined in (22), (23), and (42).

References

[1] R. Yang, L. Yin, M. Gabbouj, J. Astola, and Y. Neuvo, "Optimal weighted median filtering under structural constraints," *IEEE Transactions on Signal Processing*, vol. 43, no. 3, pp. 591–604, 1995.

[2] F. Abramovich, T. Sapatinas, and B. W. Silverman, "Wavelet thresholding via a Bayesian approach," *Journal of the Royal Statistical Society. Series B*, vol. 60, no. 4, pp. 725–749, 1998.

[3] H. A. Chipman, E. D. Kolaczyk, and R. E. McCulloch, "Adaptive Bayesian wavelet shrinkage," *Journal of the American Statistical Association*, vol. 92, no. 440, pp. 1413–1421, 1997.

[4] M. Clyde and E. I. George, "Flexible empirical Bayes estimation for wavelets," *Journal of the Royal Statistical Society. Series B*, vol. 62, no. 4, pp. 681–698, 2000.

[5] S. Sardy, "Minimax threshold for denoising complex signals with waveshrink," *IEEE Transactions on Signal Processing*, vol. 48, no. 4, pp. 1023–1028, 2000.

[6] A. Ben Hamza, P. L. Luque-Escamilla, J. Martínez-Aroza, and R. Román-Roldán, "Removing noise and preserving details with relaxed median filters," *Journal of Mathematical Imaging and Vision*, vol. 11, no. 2, pp. 161–177, 1999.

[7] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaption via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.

[8] D. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Association*, vol. 90, pp. 1200–1224, 1995.

[9] H. Krim, D. Tucker, S. Mallat, and D. Donoho, "On denoising and best signal representation," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2225–2238, 1999.

[10] S. Beheshti and M. A. Dahleh, "A new information-theoretic approach to signal denoising and best basis selection," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3613–3624, 2005.

[11] A. Fakhrzadeh and S. Beheshti, "Minimum noiseless description length (MNDL) thresholding," in *Proceedings of the IEEE Symposium on Computational Intelligence in Image and Signal Processing (CIISP '07)*, pp. 146–150, 2007.

[12] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532–1546, 2000.

[13] S. Krishnan and R. M. Rangayyan, "Automatic de-noising of knee-joint vibration signals using adaptive time-frequency representations," *Medical and Biological Engineering and Computing*, vol. 38, no. 1, pp. 2–8, 2000.

[14] I. Daubechies, *Ten Lectures on Wavelets*, vol. 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*, SIAM, Philadelphia, Pa, USA, 1992.

[15] F. Luisier, T. Blu, and M. Unser, "A new sure approach to image denoising: interscale orthonormal wavelet thresholding," *IEEE Transactions on Image Processing*, vol. 16, no. 3, pp. 593–606, 2007.

[16] M. Zhang and B. K. Gunturk, "Multiresolution bilateral filtering for image denoising," *IEEE Transactions on Image Processing*, vol. 17, no. 12, pp. 2324–2333, 2008.

[17] S. G. Mallat, "Theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.