

# A noise reduction approach based on Stein's unbiased risk estimate

Qiang Guo<sup>a,b,\*</sup>, Caiming Zhang<sup>a,b,c</sup>

<sup>a</sup> School of Computer Science and Technology, Shandong University of Finance and Economics, Erhuan Dong Road, Jinan, 250014, China

<sup>b</sup> Shandong Provincial Key Laboratory of Digital Media Technology, Erhuan Dong Road, Jinan, 250014, China

<sup>c</sup> School of Computer Science and Technology, Shandong University, Shunhua Road, Jinan, 250101, China

\*Corresponding author, e-mail: qguo2010@126.com

Received 10 Jan 2012

Accepted 23 Apr 2012

**ABSTRACT:** This paper proposes a new wavelet-based shrinkage function for 1D signal noise reduction. This shrinkage function adopts the intrascale correlations between wavelet coefficients and exploits Stein's unbiased risk estimator to achieve the optimal parameter. Unlike the methods based upon Bayes estimators, the proposed method does not use any prior hypotheses on wavelet coefficients. Experiments performed on simulated signals clearly indicate that our method outperforms conventional noise reduction methods in the sense of the signal-to-noise ratio.

**KEYWORDS:** parameter estimation, stationary wavelet, wavelet shrinkage

## INTRODUCTION

Many real-world signals are often corrupted by additive noise due to imperfect acquisition and transmission. Hence, noise reduction is an essential step before signal analysis. The major goal of noise reduction is to restore the original signals from noisy measurements. The wavelet transform has been shown to be a powerful tool for noise reduction due to its capability of sparse representation<sup>1</sup>. A wavelet-based scheme via hard shrinkage and soft shrinkage was pioneered by Donoho and Johnstone<sup>2,3</sup>. Indeed, the wavelet transform concentrates the energies of most signals of interest in a few coefficients, while the power of the noise is uniformly spread throughout all the coefficients<sup>4</sup>. Wavelet shrinkage is therefore the most straightforward way to distinguish the signals of interest from noise. The basic idea of wavelet shrinkage is to perform wavelet decomposition on the noisy signal, then estimate the noise-free wavelet coefficients according to some shrinkage rule, and finally reconstruct the restored signal by performing the inverse wavelet transform<sup>5</sup>.

A great number of wavelet shrinkage methods based on Bayesian frameworks have been developed in the literature. Vidakovic presented a nonlinear shrinkage by coherent Bayesian inference in the wavelet domain<sup>6</sup>. Chang et al proposed an adaptive shrinkage method by using a generalized Gaussian

distribution as the prior model<sup>7</sup>. In Ref. 8, Sendur and Selesnick developed a bivariate shrinkage method which adopts a non-Gaussian bivariate distribution to model wavelet coefficients. A shrinkage method based on a Gaussian scale mixture model was proposed by Portilla et al<sup>9</sup>. Cho and Bui exploited a multivariate generalized Gaussian distribution to model the dependency between wavelet coefficients<sup>10</sup>.

The performance of the aforementioned shrinkage methods greatly depends on the effectiveness of the prior model of the wavelet coefficients. Recently, Luisier et al introduced an efficient method for noise reduction derived by the minimization of Stein's unbiased risk estimate (SURE)<sup>11</sup>. Unlike the methods based upon Bayes estimators, this method avoids any prior hypotheses on the noise-free signal. However, it does not consider the intrascale correlations between wavelet coefficients. In order to use the intrascale dependencies of wavelet coefficients to improve the performance of the noise reduction algorithm, we present an effective intrascale wavelet shrinkage method. Experimental results show that the restored signals by the proposed method can obtain a higher signal-to-noise ratio (SNR) than some existing methods.

## SHORT REVIEW OF WAVELET SHRINKAGE

Assume the noisy signal  $\mathbf{y} = (y_i)_{i \in [1, N]}$  is given by

$$\mathbf{y} = \mathbf{x} + \mathbf{n} \quad (1)$$

where  $\mathbf{x} = (x_i)_{i \in [1, N]}$  is the noise-free signal and  $\mathbf{n} = (n_i)_{i \in [1, N]}$  is Gaussian white noise with independent identical distributions  $N(0, \sigma_n^2)$ . In the wavelet domain, the noisy signal can be represented by

$$\mathbf{w} = \mathbf{u} + \mathbf{v} \tag{2}$$

where  $\mathbf{w} = (w_i)_{i \in [1, N]}$  is the noisy wavelet coefficients vector, and  $\mathbf{u} = (u_i)_{i \in [1, N]}$  represents the noise-free wavelet coefficients vector corrupted by the noise coefficients vector  $\mathbf{v} = (v_i)_{i \in [1, N]}$ . The goal of noise reduction is to obtain an estimate  $\hat{\mathbf{u}}$  of the noise-free wavelet coefficients vector  $\mathbf{u}$  by using a wavelet shrinkage function  $\theta(\cdot)$ .

Many shrinkage functions have been developed in the literature. The two most popular functions are the hard and the soft shrinkage functions<sup>3</sup>. The hard shrinkage function (HSF) is defined as

$$\theta_h(w_i) = \begin{cases} w_i, & \text{if } |w_i| \geq \lambda, \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

and the soft shrinkage function (SSF) is defined as

$$\theta_s(w_i) = \begin{cases} \text{sgn}(w_i)(|w_i| - \lambda), & \text{if } |w_i| \geq \lambda, \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

where  $\lambda = \sigma \sqrt{2 \log N}$ . Obviously, as  $N \rightarrow +\infty$ , the threshold  $\lambda$  increases, and as  $N \rightarrow 0$ , the threshold  $\lambda$  decreases, which is undesirable for the hard and the soft shrinkage functions. Luisier et al proposed an SURE-based shrinkage function<sup>11</sup> which is defined as

$$\theta_{\text{SURE}}(w_i) = \sum_{k=1}^K a_k \varphi_k(w_i) \tag{5}$$

where  $a_k (k \in [1, K])$  are the unknown parameters and  $\varphi_k(w_i) = w_i \exp(-(k-1)w_i^2/12\sigma_n^2)$ . However, the shrinkage function  $\theta_{\text{SURE}}(\cdot)$  does not consider the intrascale correlations between wavelet coefficients which can be used to further improve the performance of the noise reduction algorithm. Moreover, the computation time can be reduced by decreasing the number of parameters  $a_k$ . In the following section, we therefore propose a simple and effective SURE-based intrascale shrinkage method which exploits these correlations and decreases the number of unknown parameters.

**PROPOSED SURE-BASED INTRASCALE SHRINKAGE**

We use the averaged magnitude of the neighbouring coefficients to quantify the intrascale correlations be-

tween wavelet coefficients which can be expressed by

$$c_i = \frac{1}{2M} \sum_{j=i-M, j \neq i}^{i+M} |w_j| \tag{6}$$

where  $2M$  is the number of neighbouring coefficients. In practice, it is sufficient to set  $M = 2$ . In the following section, the impact of  $M$  on the performance of the proposed method is discussed. In order to use the intrascale correlation coefficient  $c_i$ , we introduce the following shrinkage function

$$\theta_{\text{new}}(w_i) = w_i + a\mu(c_i)\mu(w_i)w_i \tag{7}$$

where  $a$  is the unknown parameter and  $\mu(x) = \exp(-x^2/12\sigma_n^2)$ . This shrinkage function can be regarded as a modified version of  $\theta_{\text{SURE}}(\cdot)$  by decreasing the number of parameters  $a_k$  and using the intrascale correlations between wavelet coefficients. The function  $\mu(x)$  is a simplified version of  $\varphi_k(x)$ .

The (7) is our proposed intrascale shrinkage function which contains two unknown parameters  $\sigma_n$  and  $a$ . In Ref. 12, a robust estimate of  $\sigma_n$  is given by

$$\sigma_n = \frac{\text{MAD}}{0.6745} \sqrt{\frac{2p}{2p+1}} \tag{8}$$

where MAD is the median absolute deviation of the finest wavelet coefficients, and  $p$  is the number of vanishing moments of the wavelet transform. For fixed  $\sigma_n$ , we can obtain the optimal estimate of  $a$  by minimizing the mean squared error (MSE), i.e.,

$$\hat{a} = \arg \min_a \frac{1}{N} \sum_{i=1}^N |\theta_{\text{new}}(w_i) - u_i|^2. \tag{9}$$

In practice, however, we do not have access to the noise-free wavelet coefficient  $u_i$ . Fortunately, the SURE principle proposed by Stein overcomes this difficulty. This can be expressed by the following theorem.

**Theorem 1** Let  $\theta : \mathbb{R} \rightarrow \mathbb{R}$  be a (weakly) differentiable function that is finite at infinity. Then the random variable

$$\epsilon = \frac{1}{N} \sum_{i=1}^N (\theta(w_i)^2 - 2w_i\theta(w_i) + 2\sigma_n^2\theta'(w_i)) + \frac{1}{N} \sum_{i=1}^N u_i^2 \tag{10}$$

is an unbiased estimator of the MSE, i.e.,

$$E\{\epsilon\} = E \left\{ \frac{1}{N} \sum_{i=1}^N |\theta(w_i) - u_i|^2 \right\} \tag{11}$$

where  $E\{\cdot\}$  is the expectation operator.

Due to the fact that  $\theta_{\text{new}}(x)$  is differentiable, we can apply Theorem 1 to estimate  $a$  by performing differentiation over the parameter  $a$  in (7) as follows

$$\frac{\partial \epsilon}{\partial a} = \frac{2}{N} \sum_{i=1}^N \left( \theta_{\text{new}}(w_i) \mu(c_i) \mu(w_i) w_i + \sigma_n^2 \mu(c_i) \left( \mu(w_i) + w_i \mu'(w_i) \right) \right) = 0. \quad (12)$$

It is equivalent to

$$\sum_{i=1}^N \left( a \mu^2(c_i) \mu^2(w_i) w_i^2 + \sigma_n^2 \mu(c_i) \left( \mu(w_i) + w_i \mu'(w_i) \right) \right) = 0. \quad (13)$$

So the parameter  $a$  can be solved by

$$a = \frac{-\sigma_n^2 \sum_{i=1}^N (\mu(c_i) \mu(w_i) + w_i \mu(c_i) \mu'(w_i))}{\sum_{i=1}^N \mu^2(c_i) \mu^2(w_i) w_i^2}. \quad (14)$$

### EXPERIMENTAL RESULTS

In our simulations, we use the standard test signals Piece-Polynomial, Blocks, Bumps, and Doppler, which are generated by WAVELAB available at [www-stat.stanford.edu/wavelab](http://www-stat.stanford.edu/wavelab). The noisy signals are created by adding Gaussian white noise with different noise levels to the signals. The proposed method is evaluated by comparing it with some most commonly used shrinkage functions, namely, HSF (3), Luisier’s SURE (LSURE) (5), and Han’s SSF (HSSF)<sup>13</sup>. For the sake of evaluating the performance of these methods, the signal-to-noise ratio (SNR) is used as the quantitative criterion. It is defined by

$$\text{SNR} = 10 \log_{10} \left( \frac{\sum_{i=1}^N |x_i - \bar{x}|^2}{\sum_{i=1}^N |x_i - \hat{x}_i|^2} \right), \quad (15)$$

where  $\bar{x}$  and  $\hat{x}_i$  are the mean value and the estimate of the signal  $x_i$ , respectively. In all comparisons, we use the stationary Daubechies wavelet with four vanishing moments over three decomposition levels, which means  $p = 4$  in (8). All algorithms are fully implemented in Matlab 7.0, and in order to facilitate the reproducibility of our work, the source code of the proposed method is available on request.

Fig. 1 shows the visual quality of the various algorithms for the 1D signal Piece-Polynomial. As can be observed, the proposed method exhibits less distortion than the others. Table 1 displays the SNR results generated by all algorithms for the test signals. Our

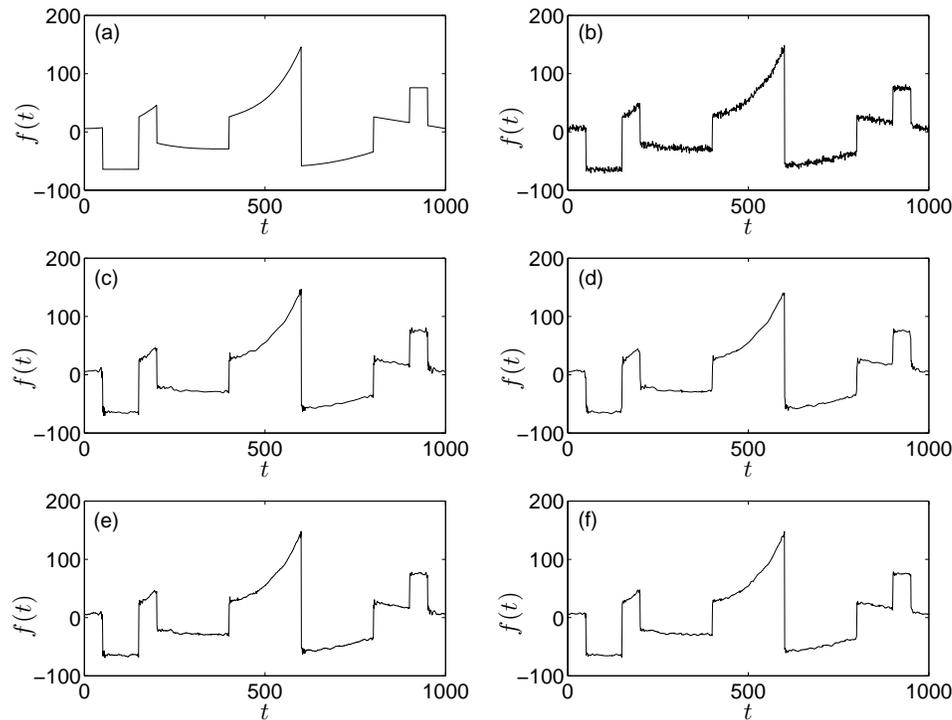
**Table 1** Comparison of simulation results on test signals in terms of SNR (dB).

Signal	Noisy	HSF	HSSF	LSURE	Ours
Piece-Polynomial	34.05	38.11	31.24	38.85	40.68
	28.03	31.85	28.49	33.06	34.87
	24.51	28.52	26.45	29.41	31.40
	14.05	17.30	18.89	18.97	20.81
	8.03	12.72	13.48	13.70	15.32
Blocks	27.80	29.91	19.57	31.45	33.92
	19.84	22.40	17.23	24.11	26.29
	13.82	17.05	15.51	18.59	20.19
	10.30	14.32	14.26	15.44	16.80
	7.80	11.42	13.46	13.03	14.73
Bumps	37.42	40.58	16.50	40.75	43.09
	31.40	34.09	14.80	35.45	37.49
	27.88	31.25	13.82	32.29	34.23
	17.42	19.60	10.82	21.83	24.38
	7.88	11.50	7.42	12.75	15.37
Doppler	29.33	33.05	17.34	34.53	36.20
	23.31	27.92	15.58	28.86	30.12
	19.79	24.57	14.72	25.65	26.77
	13.77	19.84	13.51	20.83	21.32
	9.33	15.97	12.47	17.09	17.20

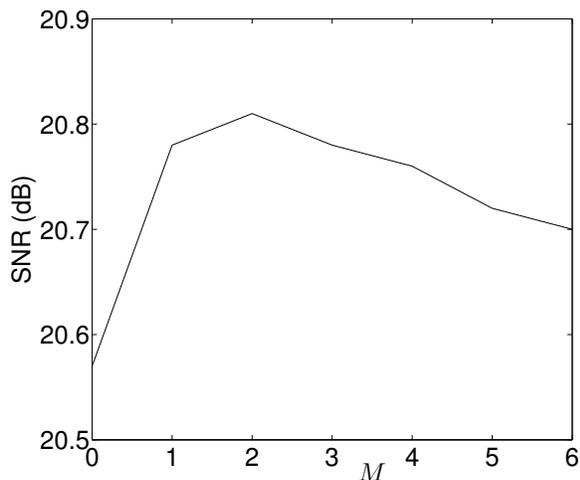
method clearly outperforms HSF, HSSF and LSURE in terms of SNR, and the average gains are nearly 3.0 dB, 10.3 dB and 1.7 dB, respectively. In order to illustrate the impact of the parameter  $M$ , we perform our method with different values of  $M$  for the signal Piece-Polynomial. Fig. 2 shows the SNR curve as the function of the parameter  $M$ . We can see that the optimal SNR value is achieved in  $M = 2$ . Therefore, to get the best performance, the parameter  $M$  is set as 2.

### DISCUSSION

In this paper, we propose a new noise reduction method based on the wavelet transform which exploits the SURE principle to estimate the noise-free wavelet coefficients and avoids any prior statistical model of wavelet coefficients. The proposed method gives better noise reduction than LSURE. Experimental results show that our method is superior to conventional methods. Since the proposed shrinkage function  $\theta_{\text{new}}(\cdot)$  is a point processing operator, its computational complexity is  $O(1)$ . Hence the complexity of the proposed method is dominated by the wavelet transform. In experiments we use the stationary wavelet which simply eliminates the down-sampling process in the filter-bank implementation of a wavelet transform<sup>14</sup>. For a length- $N$  signal, the stationary wavelet transform requires  $O(N \log N)$  operations<sup>15</sup>. Thus the computational complexity of the proposed



**Fig. 1** Experimental results. (a) original signal; (b) noisy signal, SNR = 24.51 dB; (c) HSF, SNR = 28.52 dB; (d) HSSF, SNR = 26.45 dB; (e) LSURE, SNR = 29.41 dB; (f) the proposed method, SNR = 31.40 dB.



**Fig. 2** SNR value as a function of the number  $M$  of neighbouring coefficients.

method is  $O(N \log N)$ .

**Acknowledgements:** This work was supported by the National Natural Science Foundation of China (Grant No. 61020106001), Shandong Province Natural Science Foundation (Grant No. ZR2011FL028), Shandong Province Higher Educational Science and Technology Programme

(Grant No. J11LG77) and Doctoral Foundation of Shandong University of Finance and Economics.

## REFERENCES

1. Guo Q, Yu S (2010) Image denoising using a multivariate shrinkage function in the curvelet domain. *IEICE Electron Express* **7**, 126–31.
2. Donoho DL (1995) De-noising by soft-thresholding. *IEEE Trans Inform Theor* **41**, 613–27.
3. Donoho DL, Johnstone IM (1994) Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **81**, 425–55.
4. Donoho DL (1993) Unconditional bases are optimal bases for data compression and for statistical estimation. *Appl Comput Harmon A* **1**, 100–15.
5. Tian J, Chen L, Ma L (2010) A wavelet-domain non-parametric statistical approach for image denoising. *IEICE Electron Express* **7**, 1409–15.
6. Vidakovic B (1998) Nonlinear wavelet shrinkage with Bayes rule and Bayes factor. *J Am Stat Assoc* **93**, 173–9.
7. Chang SG, Yu B, Vetterli M (2000) Adaptive wavelet thresholding for image denoising and compression. *IEEE Trans Image Process* **9**, 1532–46.
8. Sendur L, Selesnick IW (2002) Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency. *IEEE Trans Signal Process* **50**, 2744–56.

9. Portilla J, Strela V, Wainwright MJ, Simoncelli EP (2003) Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Trans Image Process* **12**, 1338–51.
10. Cho D, Bui TD (2005) Multivariate statistical modeling for image denoising using wavelet transforms. *Signal Process Image Comm* **20**, 77–89.
11. Luisier F, Blu T, Unser M (2007) A new SURE approach to image denoising: interscale orthonormal wavelet thresholding. *IEEE Trans Image Process* **16**, 593–606.
12. Jansen M, Bultheel A (2001) Asymptotic behavior of the minimum mean squared error threshold for noisy wavelet coefficients of piecewise smooth signals. *IEEE Trans Signal Process* **49**, 1113–8.
13. Han M, Liu Y, Xi J, Guo W (2007) Noise smoothing for nonlinear time series using wavelet soft threshold. *IEEE Signal Process Lett* **14**, 62–5.
14. Lang M, Guo H, Odegard JE, Burrus CS, Wells RO (1996) Noise reduction using an undecimated discrete wavelet transform. *IEEE Signal Process Lett* **3**, 10–2.
15. Wei D, Rajashekar U, Bovik AC (2005) Wavelet denoising for image enhancement. In: Bovik AC (ed) *Handbook of Image and Video Processing*, 2nd edn, Elsevier, pp 157–65.