Denoising using higher-order statistics

Samuel Peter Kozaitis
Sunghee Kim

Denoising using higher-order statistics

Samuel P. Kozaitis, and Sunghee Kim
Florida Institute of Technology
Department of Electrical and Computer Engineering
150 W. University Blvd.
Melbourne, FL 32901

ABSTRACT

We used a higher-order correlation-based method for signal denoising. In our approach, we determined which wavelet coefficients contained mostly noise, or signal, based on higher-order statistics. Because the higher that second-order moments of the Gaussian probability function are zero, the third-order correlation coefficient will not have a statistical contribution from Gaussian noise. We obtained results for both 1-D signals and images. In all cases, our approach showed improved results when compared to a more popular denoising method.

Keywords: denoising, wavelet transform, higher-order statistics

1. INTRODUCTION

Removing noise from a signal often degrades the signal in the process. Therefore, various types of filters have been developed to reduce noise while retaining signal characteristics. Nonlinear filtering techniques have been introduced that are based on the theory of robust estimation[1]. They have showed success in removing noise from an image while preserving details. Many of these approaches are based on the median filter. For example, in the relaxed median filter, the lower and upper order statistics are compared to the center sample of the filter window[2]. These type of filters seem to work best with impulsive type noise. Under the assumption of Gaussian noise, there have been several-wavelet based approaches[3-7]. The basic idea is that the wavelet transform compacts a signal's energy more efficiently than noise in the wavelet domain.

We also used a wavelet-based approach to remove noise from signals. We retained or set to zero wavelet coefficients based on higher-order statistics. Higher-order correlations are an extension of the more familiar second-order cross-correlation function, and have the advantage of being theoretically shown to eliminate noise of unknown spectral density under certain conditions. In our approach, we applied a third-order correlation technique to identify wavelet coefficients that contained mostly signal. Because the higher that second-order moments of the Gaussian probability function are zero, a third-order correlation coefficient will statistically not have a contribution from Gaussian noise. Therefore, in our approach, we examined correlation coefficients in an environment where the noise had been reduced.

2 DENOISING

A straightforward approach to denoising involves setting wavelet coefficients that are mostly noise to zero. This typically involves scaling an image, taking its wavelet transform, thresholding the wavelet coefficients, then taking the inverse wavelet transform. After calculating the wavelet transform of the scaled image, the noisy wavelet

Independent Component Analyses, Wavelets, and Neural Networks, Anthony J. Bell,
Mladen V. Wickerhauser, Harold H. Szu, Editors, Proceedings of SPIE
Vol. 5102 (2003) © 2003 SPIE · 0277-786X/03/$15.00

transform coefficients are usually subjected to a soft or hard threshold. A hard threshold indicates that wavelet transform coefficients are retained only if their absolute value is greater than or equal to a threshold *t*.

There are different ways to choose a threshold. One threshold that is easy to implement, has been shown to work well for sparse data, and produces a good visual quality of a denoised signal is described as[3]

$$t = \sigma(2\log(N))^{1/2}\ , \tag{1}$$

where N is the number of wavelet coefficients to be thresholded, and σ is the standard deviation of the noise. The noise variance may be estimated by taking the median value of the wavelet coefficients at the smallest scale. After thresholding, the inverse wavelet transform is performed. This thresholding method removes essentially all of the empirical wavelet coefficients that could (statistically) be attributed solely to noise.

3. HIGHER-ORDER CORRELATIONS

The correlation between two functions has been often used as a measure of their similarity. The conventional correlation function is a second-order correlation, and is a special case higher-order correlations.[8,9] Although higher-order correlations have been used for many years, there use has been limited. The nth-order correlation of the signal $f(x)$ is defined as

$$f_n(\tau_1,\ \tau_2,\ .\ .\ .\ \tau_{n-1}) \equiv \sum_{k=0}^{N-1} f\ (x)\, f\ (\tau_1+x)\, f\ (\tau_2+x)\ .\ .\ .\ f\ (\tau_{n-1}+x) \tag{2}$$

where the nth-order correlation is a function of n - 1 independent variables. For n = 2, Eq. (2) becomes the second-order correlation of $f(x)$ which is the familiar autocorrelation function.

We primarily considered the third-order or triple correlation because it has the same advantages for our purpose, and is easier to calculate than other higher-order correlations. The third-order correlation, n = 3, of a one-dimensional function is a function of two variables. From Eq. (2) the third-order correlation of $f(x)$ is

$$f_3(\tau_1,\ \tau_2) = \sum_{k=0}^{N-1} f\ (x)\, f\ (\tau_1+x)\, f\ (\tau_2+x) \tag{3}$$

where $f_3(t_1,t_2)$ is symmetric with respect to its variables τ_1, and τ_2.

The third-order correlation coefficient $f_3(0,0)$, can be found by sampling the triple correlation $f_3(t_1,t_2)$, at zero displacement where $\tau_1 = \tau_2 = 0$. From Eq. (3), the third-order correlation coefficient becomes

$$f_3(0,\ 0) = \sum_{k=0}^{N-1} f^3(x) \tag{4}$$

which shows that the third-order correlation coefficient $f_3(0, 0)$, of $f(x)$, can be calculated directly as the sum of the cubes of $f(x)$ from $k = 0 - N$-1.

4. HIGHER-ORDER STATISTICS FOR SIGNAL DENOISING

In this section we described a method for denoising using third-order correlations. We considered a noisy signal that contained zero-mean noise of unknown spectral density. Our goal was to select wavelet coefficients that were due primarily to the signal, based on third-order correlations. We thought of wavelet coefficients as correlations between a signal and a wavelet at a particular scale and translation. The noisy input signal was described as $s(x) = f(x) + n(x)$ where $f(x)$ represented the signal and $n(x)$ the noise .

We first calculated the cross-correlation function between the noisy input signal and each scaled and translated wavelet. The resulting correlation function was labeled as $b_{jk}(\tau)$. Then, we calculated the third-order autocorrelation coefficient of $b_{jk}(\tau)$. Using Eq. (4), the third-order autocorrelation coefficient of the second-order correlation result was described as

$$b3_{jk}(0,0) = \sum_{\tau=1}^{2m=1} \left(b_{jk}(\tau)\right)^3 , \tag{5}$$

where m is the length of the wavelet, and the summation is performed only on that portion of the signal. Because the noisy input signal could be considered as a signal plus noise, we considered the output of the second-order correlation as consisting of two parts. One part was the correlation between the input signal and a wavelet, and the second part was the correlation between the noise and the wavelet. The second-order correlation result was written as

$$b_{jk}(\tau) = fb_{jk}(\tau) + nb_{jk}(\tau) , \tag{6}$$

where $fb_{jk}(\tau)$ and $nb_{jk}(\tau)$ represented the signal-wavelet correlation, and the noise-wavelet correlation respectively. Substituting Eq. (6) into Eq. (6), and rearranging, the expression for the third-order autocorrelation coefficient can be written as[10]

$$b3_{jk}(0,0) = \sum_{\tau=1}^{2m-1} \left(fb_{jk}(\tau)\right)^3 + 3\sum_{\tau=1}^{2m-1} \left(fb_{jk}(\tau)\right)^2 nb_{jk}(\tau) + \sum_{\tau=1}^{2m-1} fb_{jk}(\tau)\left(nb_{jk}(\tau)\right)^2 + \sum_{\tau=1}^{2m-1} \left(nb_{jk}(\tau)\right)^3 . \tag{7}$$

If we assume zero-mean noise, the second term in Eq. (7) will approach zero in the limit. If the noise is Gaussian or has a symmetric distribution, then the last term will also approach zero. The third term is related to the product of signal and the noise power. It can be minimized if the mean of the signal is set to zero.

The performance of the method can be assessed from the distribution of the decision statistic, $|b^3_{jk}(0,0)|$. To simplify matters, we considered only the first and last terms in Eq. (7). The last term can be shown to be approximately normally distributed with zero mean[11]. The first term is related to the distribution of the values of the wavelet coefficients. In signal compression this is often modeled as being Gaussian with zero mean. Under these conditions, the optimal solution in terms of minimizing the probability of error has been derived and can be used to determine whether a measurement indicates noise or signal[12]. The solution can be written as a threshold so that a wavelet coefficient is retained when,

$$\left|b3_{jk}(0,0)\right| > \sqrt{\frac{\sigma_{3n}^2 \sigma_{3f}^2 \ln\left(\frac{\sigma_{3f}^2}{\sigma_{3n}^2}\right)}{\sigma_{3f}^2 - \sigma_{3n}^2}} \quad , \tag{8}$$

where σ_{3f}^2 and σ_{3n}^2 are the variances of the first and fourth terms of Eq. (7) respectively. If we assume that the wavelet coefficients at the finest scale are mainly due to noise, then the variance of $b^3_{jk}(0,0)$ at that scale can be used to estimate σ_{3n}^2. The value of σ_{3f}^2 can be estimated from training samples and is related to the signal strength. The decision statistic is related to the ratio of the variances used in Eq. (8). If $\sigma_{3f}^2 = x\sigma_{3n}^2$, where $x > 1$, then Eq. (8) can be written as

$$\left|b3_{jk}(0,0)\right| > \sqrt{\frac{x\sigma_{3n}^2 \ln(x)}{(x-1)}} \; . \tag{9}$$

5. RESULTS

We compared our method to the more conventional approach described in section 2. We added Gaussian noise to a signal, and then attempted to remove the noise using the denoising algorithms. In both methods we estimated the noise power by examining the wavelet transform of a signal at the finest scale which was assumed to be mostly noise. In all cases, we used two levels of the wavelet transform and used Daubechies minimum-phase orthogonal 4-tap wavelets with 2 vanishing moments. The threshold for the conventional denoising approach was that of Eq. (4). For the third-order approach we used $t = 2\sigma_{3n}^2$. Figure 1 shows a sample 1-D signal. Figure 2 shows the mean-squared error (MSE) as a function of signal-to-noise ratio. Each point was an average of 20 trials due to independent Gaussian noise samples. The results showed that the third-order method consistently showed improved results when compared to the conventional method.

We also performed experiments on images. The image in Fig. 3 was used in experiments. Noise was added as before. Due to computational burden of the third-order approach, we denoised each row separately, then denoised each column of the original noisy separately, then, averaged the two results. The results comparing the second- and third-order approach are shown in Fig. 4. The figure shows that the third-order method consistently showed improved results when compared to the conventional method.

6 CONCLUSION

We used a higher-order correlation-based method for signal denoising. In our approach, we determined that a wavelet coefficient contained mostly noise, or signal, based on higher-order statistics. The method should work best for noise distributions that are symmetric, because their 3^{rd}-order correlation coefficient as described here will be statistically zero. Spatially adapting the method is expected to improve the results.

REFRENCES

1 A. B. Hamza, and H. Krim, "Image denoising: a nonlinear robust statistical approach," *IEEE Trans. on Signal Processing* **49**(12), 3045-3054 (2001)

2 A. B. Hamza, P. Luque, J. Martinez, and R. Roman, "Removing noise and preserving details with relaxed median filters," *J. Math. Imag. Vision.* **11**(2) 161-177 (1999)

3 D. L. Donoho, "De-noising via soft thresholding," *IEEE Trans. Inf. Theory*, (May 1995)

4 D. L. Donoho, and I. M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, 425-455 (1994)

5 M. K. Mihcak, I. Kozintsev, K. Ramchandran, and P. Moulin, " Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Letters* **6**(12), 300-303 (1999)

6 M. Shao, and C. L. Nikias, "Signal processing with fractional lower order moments: stable processes and their applications," *Proc. IEEE* **81**, 986-1010 (1993)

7 H. Krim, D. Tucker, S. Mallat, and D. Donoho, "On denoising and best signal representation," *IEEE Trans. on Inform. Theory* **45**, 2225-2238 (1999)

8 C. L. Nikias, and A. P. Peptropulu, *Higher-Order Spectral Analysis: A Nonlinear Processing Framework,* Prentice-Hall: New York (1993)

9 J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proc. IEEE* **79(3)**, 278-305 (1991)

10 G. B. Giannakis, and M. K. Tsatsanis, *IEEE Trans. on Acoustics, Speech, and Signal Procecssing*, **38**, 1284-1296 (1990)

11 K. S. Lii, and M. Rosenblatt, "Deconvolution and estimation of transfer function phase coefficients for non-Gaussian linear processes," *Ann. Stat.*, 1195-1208 (1982)

12 H. L. Van Trees, "*Detection, Modulation, and Estimation Theory*," Wiley:New York, vol. 1 (1969)

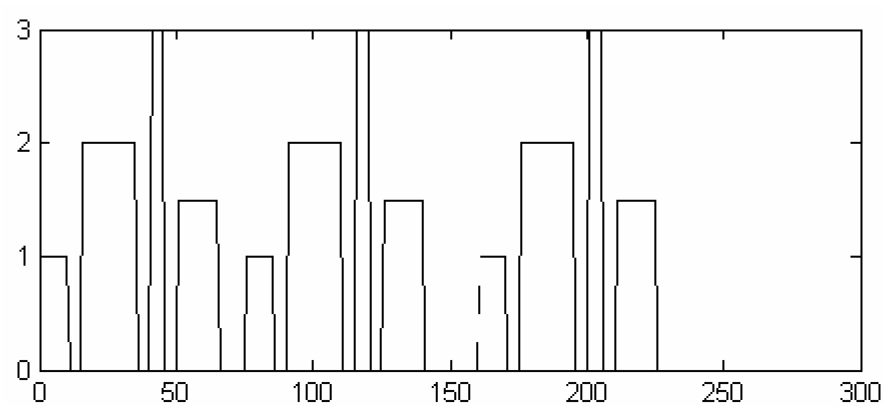

Figure 1 1–D signal used in denoising experiments.

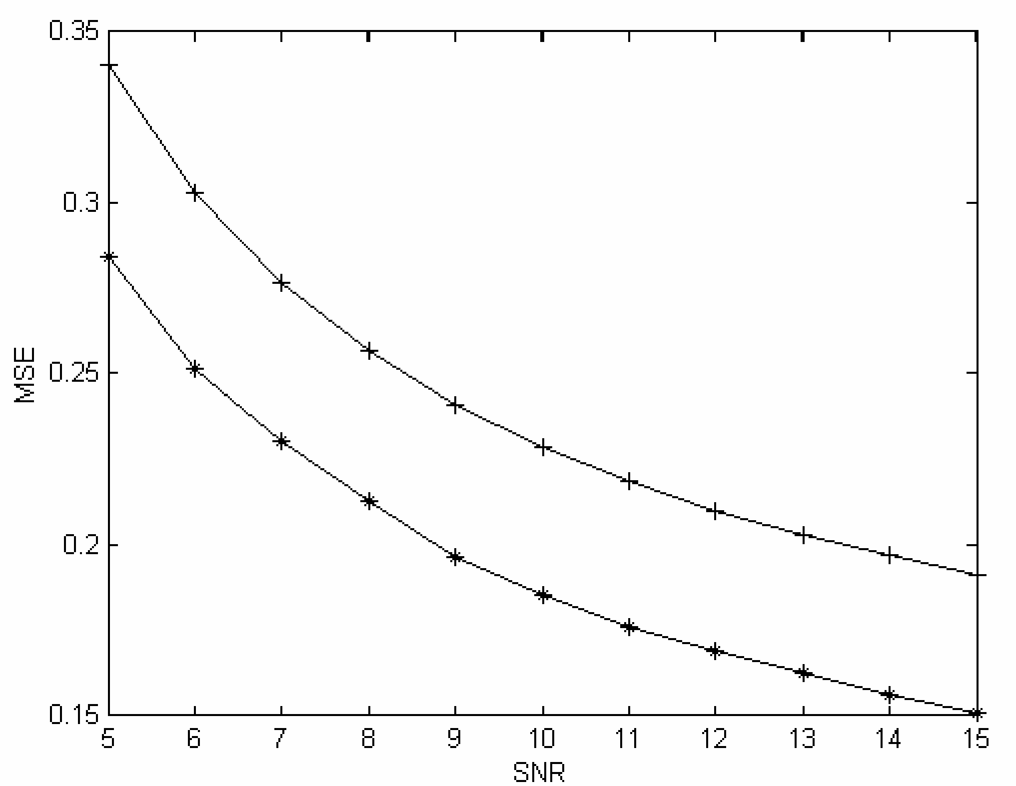

Figure 2 Comparison of 2nd (+), and 3rd (*), -order denoising methods for signal in Fig. 1.

Figure 3 Image used in denoising experiments.

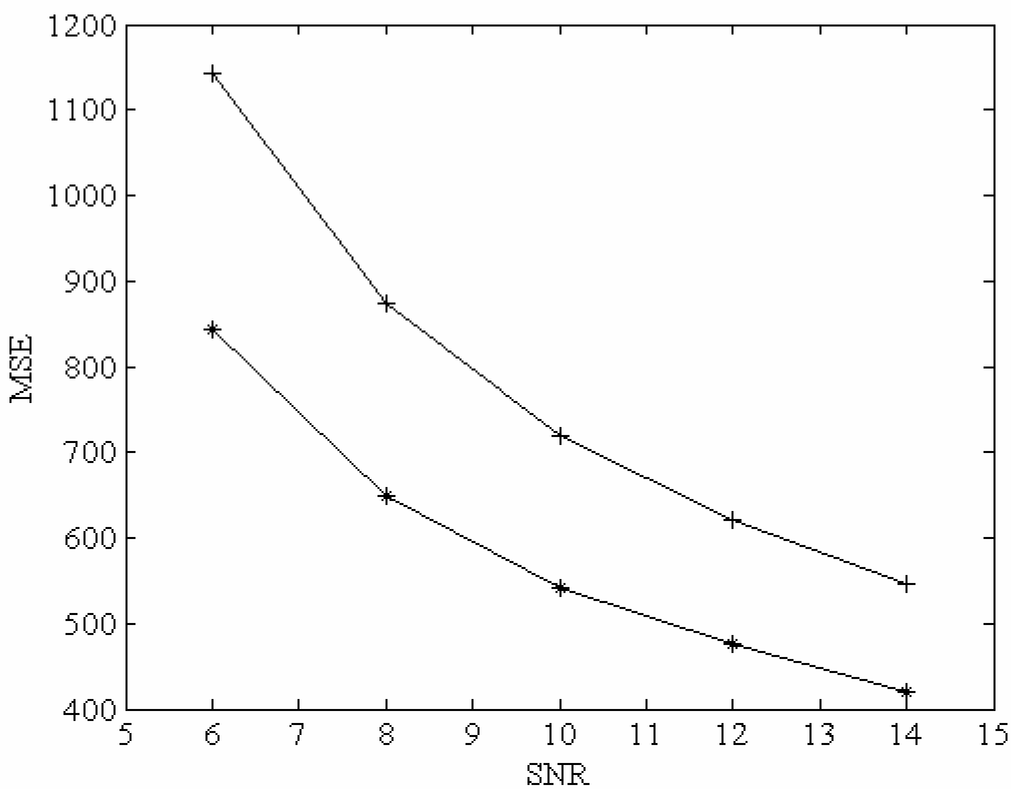

Figure 4 Comparison of 2nd (+), and 3rd (*), -order denoising methods for image in Fig. 3.