On Wavelet Denoising and its Applications to Time Delay Estimation

P. C. Ching, H. C. So, and S. Q. Wu

***Abstract*— In this correspondence, the application of dyadic wavelet decomposition in the context of time delay estimation is investigated. We consider a model in which the source signal is deterministic and the received sensor outputs are corrupted by additive noises. Wavelet denoising is exploited to provide an effective solution for the problem. Denoising is first applied to preprocess the received signals from two spatially separated sensors with an attempt to remove the contamination, and the peak of their cross correlation function is then located from which the time delay between the two signals can be derived. A novel wavelet shrinkage/thresholding technique for denoising is introduced, and the performance of the algorithm is analyzed rigorously. It is proved that the proposed method achieves global convergence with a high probability. Simulation results also corroborate that the technique is efficient and performs significantly better than both the generalized cross correlator (GCC) and the direct cross correlator (CC).**

***Index Terms*—Denoising, time delay estimation, wavelet decomposition.**

I. Introduction

Extraction of the time difference of arrival (TDOA) between signals received at two spatially separated sensors has been widely applied to sonar and radar to find the position and speed of a target transmitter [1]. A recent application is found in a global positioning system where the location of a radiating object can be determined using differential satellite path delay measurements [2].

There are many existing time delay estimation (TDE) algorithms, such as the generalized cross correlation (GCC) [3], [4] and parameter estimation techniques [5], [6]. The transmitted signal $s(t)$ and additive noises at the two sensors are all taken to be jointly stationary, mutually uncorrelated, and zero mean processes. In the GCC system, each received signal is fed through a prefilter before taking cross correlation. It has been proved that the Cramér–Rao lower bound can be achieved if the prefilter is designed properly. This method, however, requires estimation of both the source and noise spectra, which often gives rise to a large delay variance, particularly for short data lengths. Furthermore, the GCC assumes that $s(t)$ is a Gaussian process, and thus, it is not appropriate for use in situations where the source signal is deterministic. Another approach tackles the TDE problem in a discrete-time domain and employs an FIR filter to model the TDOA whose estimated value is obtained by interpolating the filter coefficients. This method is capable of tracking nonstationary delays by simply making the filter adaptive [6], [7].

Wavelet analysis has recently been shown to be a useful mathematical tool for many practical applications [8]–[10]. The aim of this correspondence is to develop a novel technique based on wavelet denoising to provide accurate TDOA measurement when the source input is a deterministic signal, specifically for a radar signal that can be expressed in the form of

$$s(t) = a(t)\,\sin[w_o t + \theta(t)] \tag{1}$$

where $a(t)$ is the transmitting signal envelope, and $[w_o + \theta'(t)]$ is the radian frequency modulation function. In most radar applications, $a(t)$ is simply a rectangular pulse envelope function. However, the analysis developed here is not restricted to this particular case only. We shall show that the GCC approach and other conventional methods fail to perform well because of modeling deficiencies, whereas the proposed wavelet denoising-based algorithm can give a viable and effective solution.

The new method consists of two steps. Wavelet denoising is first applied to each received signal to recover the corresponding source waveform. The process essentially optimizes the mean-square error between the source signal $s(t)$ and its estimate $\hat{s}(t)$ restored from the noisy signal and subject to the condition that it is highly probable that $\hat{s}(t)$ is at least as smooth as $s(t)$. We then cross correlate the two restored signals, and the delay estimate is found by locating the peak of the correlation function. The method can be considered to be a special type of GCC method where its prefilter requires no *a priori* information of the signal and noise spectra. It has the advantage that accurate TDOA estimation can be obtained for deterministic input signals, and it also has great potential to operate in environments where the corrupting noises are correlated.

The corespondence is organized as follows. Section II briefly introduces the periodic wavelet transform (WT). Section III describes a novel wavelet denoising approach, whereas Section IV presents the proposed time delay estimation method and then verifies its global convergence. Finally, Section V gives the simulation results, and conclusions are stated in Section VI.

II. Wavelet Transform

Suppose the signal $s(t)$, as described by (1), is corrupted by white noise $\sigma e(t)$ with power σ^2 and a sensor outputs the noisy data sequence $x(t_0), x(t_1), \cdots, x(t_{n-1})$ within the time interval $[0, 1]$, i.e., $x(t_k) = s(t_k) + \sigma e(t_k)$, $t_k = k/n$, $n = 2^{J+1}$, and $J+1$ is a predetermined positive integer that describes the hardware capability. For notational convenience, let us write $x(k)$, $s(k)$, and $e(k)$ instead of $x(t_k)$, $s(t_k)$, and $e(t_k)$, as wel as $\mathbf{x} = (x(0), x(1), \cdots, x(n-1))^T$, $\mathbf{s} = (s(0), s(1), \cdots, s(n-1))^T$, and $\mathbf{e} = (e(0), e(1), \cdots, e(n-1))^T$.

It is well known that a periodic wavelet transform is a linear orthonormal transform; hence, there exists a $n \times n$ orthonormal matrix $\mathbf{W}$ formed by some QMF coefficients, say, $c_0, c_1, \cdots, c_{2N-1}$ [9]. This matrix transforms the sample vector $\mathbf{x}$ into

$$\mathbf{w} = \mathbf{W}\mathbf{x} \tag{2}$$

where $\mathbf{x}$ can be reconstructed by

$$\mathbf{x} = \mathbf{W}^T\mathbf{w} = \mathbf{W}^T\mathbf{w}_s + \sigma\mathbf{W}^T\mathbf{w}_e. \tag{3}$$

The component of $\mathbf{w}$ is usually indexed dyadically, i.e., $\mathbf{w} = (w(j, k), j = -1, 0, 1, \cdots, J; k = 0, 1, \cdots, 2^j - 1)^T$, $\mathbf{w}_s$ and $\mathbf{w}_e$ are the wavelet coefficient vectors of the digital source signal $\mathbf{s}$ and the white noise $\mathbf{e}$, respectively. Since $\mathbf{W}$ has a very special structure, the above transformation can be implemented by a pyramidal scheme with considerably less computation.

Manuscript received September 17, 1996; revised March 12, 1999. The associate editor coordinating the review of this paper and approving it for publication was Dr. Nurgun Erdol.

P. C. Ching is with the Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong.

H. C. So is with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong.

S. Q. Wu is with Nortel Networks, Ottawa, Ont., Canada, K1Y 4H7.

Publisher Item Identifier S 1053-587X(99)07651-5.

III. Wavelet Denoising

One of the recent breakthroughs in wavelet theory and its applications owes to Donoho and Johnstone (DJ) [10], who demonstrated that wavelet denoising is a powerful tool for removing the noisy component of a corrupted data sequence. In this section, we first describe the fundamentals of DJ's denoising method and then introduce a novel thresholding technique for denoising, which is proved to be efficacious.

A. DJ Soft and Hard Thresholding [10]

Define an estimator of $\mathbf{s}$ as $\bar{\mathbf{s}} = \mathbf{W}^T \overline{\mathbf{w}}_\mathbf{s}$. The component $\overline{w}_s(j, k)$ of $\overline{\mathbf{w}}_\mathbf{s}$ is given by

$$\overline{w}_s(j, k) = \begin{cases} w(j, k), & \text{if } 0 \le j \le j_0 \\ \hat{w}_s(j, k), & \text{if } j_0 < j \le J+1 \end{cases} \tag{4}$$

where j_0 is a critical separating point, and $\hat{w}_s(j, k)$ can be calculated either by the so-called soft threshold, shown in (5) at the bottom of the page, or the hard-threshold

$$\hat{w}_s(j, k) = \begin{cases} w(j, k), & \text{if } |w(j, k)| \ge \sigma\sqrt{2 \log n}, \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

Donoho and Johnstone [10] proved that these two simple coefficient selection rules actually give the best spatial adaptation, and the theoretical risk performance of this method can hardly be further improved.

B. A Novel Thresholding for Denoising

Generally speaking, DJ's soft threshold and hard threshold have similar characteristics. They suppress the wavelet coefficients of both the input source signal and additive noise in high scales simultaneously. If the source signal is smooth enough, this thresholding technique will perform reasonably well since the source signal will only contribute to a few wavelet coefficients, whereas the noisy components tend to distribute evenly to each and every wavelet coefficient. However, most practical source signals are actually not that smooth at some spatial points. Radar and ultrasound signals are typical examples. Hence, special attention is required for those wavelet coefficients associated with singularity points because they carry important information about the transmitted signal. Waveform restoration might suffer severe distortion if these terms are ignored. Moreover, the critical value of j_0 is difficult to select, which is particularly true as its corresponding wavelet spectrum is usually distributed in the middle scales. In our study, we apply the idea of Neyman and Pearson [11] and propose a new thresholding technique to resolve the problem. First of all, the following binary test is applied:

$$H_0\colon w(j, k) \sim \mathcal{N}(0, \sigma) \text{ versus } H_1\colon w(j, k) \sim \mathcal{N}(w_s(j, k), \sigma). \tag{7}$$

If $\mathcal{P}\{w(j, k)|H_1\}/\mathcal{P}\{w(j, k)|H_0\} > \mathcal{P}\{H_0\}/\mathcal{P}\{H_1\}$, H_1 is assured; otherwise, H_0 is taken. The motivation behind this test is that we want to keep $w(j, k)$ whenever its composition $w_s(j, k)$ is found to be significant. Since the periodic wavelet transform is an orthonormal transform, $w(j, k)$ will be independently distributed as $\mathcal{N}(w_s(j, k), \sigma)$. Therefore, if $w_s(j, k)$ is sufficiently small, $w(j, k)$ will be approximately i.i.d. as $\mathcal{N}(0, \sigma)$. Now, we choose the confidence interval as $[-\lambda, \lambda]$ with confidence level $\alpha = \mathcal{P}(|w(j, k)| \le \lambda)$. If $w(j, k) \in [-\lambda, \lambda]$, it will be regarded as total noise with a probability of $(1-\alpha)$ rejecting H_0, even though it is true. Otherwise, the noisy coefficients are shrunk by some soft thresholds. It is well known that for any given $\alpha \in (0, 1)$, the corresponding λ is optimized in the Neyman–Pearson sense.

Let $\alpha \in (0, 1)$ and $\lambda = \sqrt{2}\sigma \operatorname{erf}inv(\alpha)$ [where $\operatorname{erf}inv(*)$ is the inverse function of $\operatorname{erf}(y) = 2/(\sqrt{\pi}) \int_0^y \exp(-t^2)\, dt$], instead of using DJ's thresholds, we define three different thresholds/shrinkages as follows.

Threshold I:

$$\overline{w}_s(j, k) = \begin{cases} w(j, k) \exp\left(-\dfrac{\lambda}{|w(j, k)| - \lambda}\right), & \text{if } |w(j, k)| \ge \lambda \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

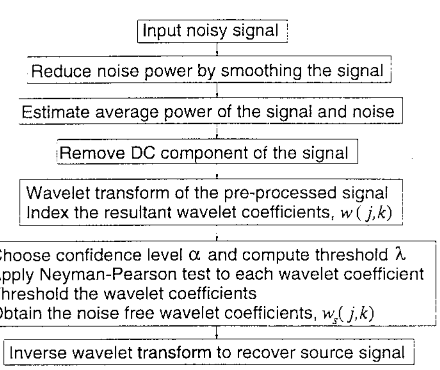
Threshold II:

$$\overline{w}_s(j, k) = \begin{cases} w(j, k), & \text{if } |w(j, k)| \ge \lambda \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

Threshold III:

$$\overline{w}_s(j, k) = \begin{cases} \operatorname{sgn}(w(j, k))(|w(j, k)| - \lambda), & \text{if } |w(j, k)| \ge \lambda \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

These three thresholds, although similar to DJ's soft and hard thresholding rules, have some intrinsic differences. When λ is small, Threshold I behaves like a hard threshold but is more flexible. Specifically, when $|w(j, k)|$ is very close to λ, Threshold I will allow $\overline{w}_s(j, k)$ to take on an approximated value of $w_s(j, k)$ instead of setting it to zero. Hence, Threshold I essentially adds some elasticity to the soft and hard thresholds while suppressing the Gibbs phenomenon caused by alternating overshoot and undershoot of a specific region. Thresholds II and III are the same as hard and soft thresholds, except *a priori* information of noise distribution is used to define the shrinkage instead of simply choosing a constant threshold $\sigma\sqrt{2 \log n}$. The respective outcomes as a result of using any of these thresholds differ somewhat as well, but which particular one should be used is rather empirical and depends largely on individual application. It is noteworthy that the thresholding parameter λ can be made scale dependent in order to further improve the denoising performance, provided that knowledge of $s(k)$ is known in advance. In this correspondence, we only examine Threshold II since the theoretical analysis for Thresholds I and III is very complicated. The flow chart of the proposed denoising method can be summarized as follows.



$$\hat{w}_s(j, k) = \begin{cases} \operatorname{sgn}(w(j, k))(|w(j, k)| - \sigma\sqrt{2 \log n}), & \text{if } |w(j, k)| \ge \sigma\sqrt{2 \log n} \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

IV. Time Delay Estimation by Wavelet Denoising

Let $s(t)$, as defined in (1), be a signal emanating from a remote source that is received by two spatially separated sensors. The received signals are usually corrupted by additive noises. Suppose $x_1(t)$ and $x_2(t)$ are the sensors outputs in which they can be expressed as

$$x_1(t) = s(t) + n_1(t) \tag{11}$$

$$x_2(t) = \beta s(t-\tau) + n_2(t) \tag{12}$$

where $n_1(t)$ and $n_2(t)$ are the corrupting white Gaussian noises, $\beta \in (0, 1)$ is an attenuation constant, and the parameter τ is the TDOA that needs to be determined. It is obvious that if the cross-correlation function $f(u)$ of $s(t)$ and $\beta s(t-\tau)$ is known, τ can be derived from the unique peak of $f(u)$. However, it is difficult to obtain the exact cross-correlation function and to locate its peak because of contamination due to noise. Let us consider the discrete version of model (11) and (12), i.e.,

$$x_1(k) = s(k) + n_1(k), \qquad k = 0, 1, \cdots, n-1 \tag{13}$$

$$x_2(k) = \beta s(k-\tau) + n_2(k), \qquad k = 0, 1, \cdots, n-1. \tag{14}$$

For simplicity, we assume that $n_1(k)$ and $n_2(k)$ are i.i.d. $\mathcal{N}(0, \sigma_1)$ and i.i.d. $\mathcal{N}(0, \sigma_2)$, respectively. These assumptions enable mathematical analysis to be derived rigorously, but in theory, they can be relaxed to accommodate colored noises as well.

WT Time Delay Estimation Method: The time delay estimation process is described step by step as follows.

- Use the WT Denoising Method to process each of the data blocks $\{x_1(k)\}$ and $\{x_2(k)\}$. Put $\hat{\mathbf{s}}_1 = (\hat{s}_1(0), \hat{s}_1(1), \cdots, \hat{s}_1(n-1))^T$ and $\hat{\mathbf{s}}_2 = (\hat{s}_2(0), \hat{s}_2(1), \cdots, \hat{s}_2(n-1))^T$. That means $\hat{\mathbf{s}}_1$ is the estimate of the original source $\mathbf{s}$ while $\hat{\mathbf{s}}_2$ is the estimate of the delayed and attenuated version of $\mathbf{s}$, denoted by $\beta\mathbf{s}(\tau)$.
- Construct the estimated cross-correlation function $\hat{f}(u)$ of $f(u)$ by its estimated sampling values,

$$\hat{f}(m) = \frac{1}{n}\sum_{k=0}^{n-m-1} \hat{s}_1(k)\hat{s}_2(k+m),$$

$$m = 1-n, 2-n, \cdots, 0, 1, \cdots, n-1.$$

- Find the peak $\hat{\tau}$ of $\hat{f}(u)$.

From [10] and [12], we can write

$$E\|\hat{\mathbf{s}}_1 - \mathbf{s}\|^2 \le \rho_1(n) \tag{15}$$

and

$$E\|\hat{\mathbf{s}}_2 - \beta\mathbf{s}(\tau)\|^2 \le \rho_2(n) \tag{16}$$

where

$$\rho_1(n) = (2\log(n)+1)\left[\sigma_1^2 + \sum_{ji}\min\{\sigma_1^2, w_s^2(j,i)\}\right] \tag{17}$$

and

$$\rho_2(n) = (2\log(n)+1)\left[\sigma_2^2 + \sum_{ji}\min\{\sigma_2^2, \beta^2 w_s^2(j,i)\}\right]. \tag{18}$$

With more restrictive conditions, say, $s(t)$ and the wavelet function are both very smooth, it can be shown [12] that the error bound can be further reduced, viz. $\rho_1(n) = O(\log^2(n)) = \rho_2(n)$.

Using the definition of cross-correlation function, $\hat{f}(m)$ and $f(m)$ can be simply expressed as $\hat{f}(m) = (1/n)\hat{s}_{1m}^T\hat{s}_{2m}$ and $f(m) = (1/n)s_{1m}^T s_{2m}$, where $\hat{s}_{1m}$, $\hat{s}_{2m}$, s_{1m}, and s_{2m} are the corresponding data blocks of length $n-|m|$. Then, we have

$$\begin{aligned} E|\hat{f}(m) - f(m)|^2 &= \frac{E\|\hat{\mathbf{s}}_{1m}^T\hat{\mathbf{s}}_{2m} - \mathbf{s}_{1m}^T\mathbf{s}_{2m}\|^2}{n^2} \\ &= \frac{E\|(\hat{\mathbf{s}}_{1m} - \mathbf{s}_{1m})^T\hat{\mathbf{s}}_{2m} + \mathbf{s}_{1m}^T(\hat{\mathbf{s}}_{2m} - \mathbf{s}_{2m})\|^2}{n^2} \\ &\le \frac{2E\|\hat{\mathbf{s}}_{1m} - \mathbf{s}_{1m}\|^2\|\hat{\mathbf{s}}_{2m}\|^2 + 2\|\mathbf{s}_{1m}\|^2 E\|\hat{\mathbf{s}}_{2m} - \mathbf{s}_{2m}\|^2}{n^2} \\ &\le \frac{2E\|\hat{\mathbf{s}}_1 - \mathbf{s}\|^2\|\beta\mathbf{s}\|^2 + 2\|\mathbf{s}\|^2 E\|\hat{\mathbf{s}}_2 - \beta\mathbf{s}(\tau)\|^2}{n^2} \\ &\le \frac{2}{n}(\rho_1(n) + \rho_2(n))E_s \end{aligned} \tag{19}$$

where $E_s = \int |s(t)|^2\,dt \approx 1/n\|\mathbf{s}\|^2$ denotes the energy of $s(t)$. Therefore, when n or the observation interval tends to infinity, the mean square error $E|\hat{f}(m) - f(m)|^2$ will become null, which illustrates that the estimator $\hat{f}(u)$ is consistent. The above inequality also implies that $\hat{f}(m)$ uniformly converges to $f(m)$ with a high probability. This convergence property is particularly helpful for peak estimation. In fact, it can be proved that if the discrete peak m^* of $\{f(m)\}$ is close to the true peak τ of $f(u)$, then the peak $\hat{m}$ of $\{\hat{f}(m)\}$ will also be close to τ with a high probability. Therefore, the peak of an interpolation function of $\hat{f}(m)$ will be much more closer to the true peak τ.

V. Simulation Results

Computer simulations had been carried out to evaluate the performance of the wavelet denoising-based cross correlation approach with the proposed new threshold (WD-CC-NT) for time delay estimation. Comparisons were made with the direct cross-correlator (CC), generalized cross-correlator with maximum likelihood prefilters (GCC-ML), and SCOT prefilters (GCC-SCOT). Cross correlation with wavelet denoising using DJ's soft threshold (WD-CC-DJ) was also examined in order to contrast the two thresholding rules. The source signal was given by (1) with $w_o = 160\pi$ and $\theta(t) = 0$. The envelope $a(t)$ had a value of 10 for $t \in [0, 0.2]$ and 0 for $t \in (0.2, 1]$. The actual delay was set to $0.292\,968\,75$ and $\beta = 1$, whereas the sampling interval was fixed at 2^{-11}. The corrupting noise sequences $n_1(k)$ and $n_2(k)$ were white Gaussian processes produced by a random number generator. The average power of $s(k)$ was 10, and different signal-to-noise ratios (SNR's) were obtained by proper scaling of the noise sequences. Prior to applying wavelet transform, the received signals were smoothed by taking averages of their adjacent samples. By so doing, the noise power was reduced, but the noise became nonwhite. For simplicity, we still applied Threshold II in the wavelet denoising procedure, although it could be modified for correlated noise [15]. Daubechies h2 wavelet (its four filter coefficients are 0.482 962 913 144 534 1, 0.836 516 303 737 807 7, 0.224 143 868 042 013 4, −0.129 409 522 551 260 3) was used while the threshold λ was calculated with the confidence level $\alpha = 0.95$. All results provided were averages of 200 independent tests.

Fig. 1 compares the mean square delay errors versus SNR of the five methods. As expected, the error of all these methods decreased as SNR increased. Although it has been proved [10] that soft thresholding has an optimum risk performance, the delay estimates obtained using this threshold were less accurate than the other four methods. This is probably because of the constant threshold $\sigma\sqrt{2\log(n)}$, which, in most cases, is too large to retain all pertinent signal information. Another reason might be due to the fact that

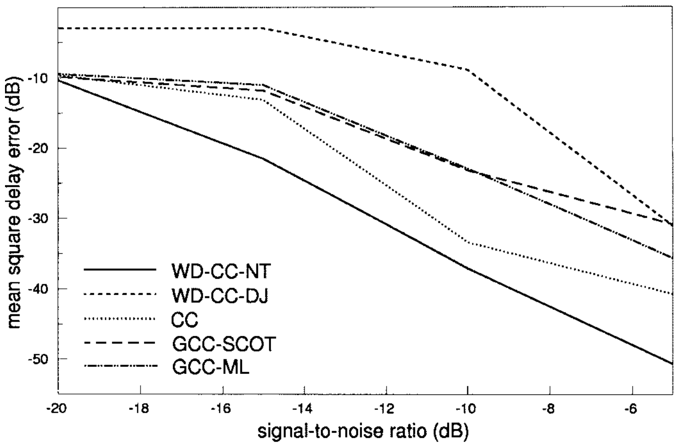


Fig. 1. Mean square delay error versus SNR.

DJ's analysis is actually based on worst-case measurement. It can be seen that the performance of the two generalized cross correlators was worse than that of the direct cross-correlation method. This is because larger variances had arisen when prefilter coefficients were computed using the spectral estimates of the received signals. This finding agrees with previous studies in [13] and [14]. On the other hand, the WD-CC-NT method has attained the best performance for all SNR conditions. Actually, for a wide range of values of $\alpha, \alpha \in (0.7, 0.997)$, WD-CC-NT gave very similar results. This means that the choice of the confidence level is not that critical, and nearly minimum delay estimation error can be acquired if α is selected as described previously. When SNR is less than $-$20 dB, the performance of WD-CC-NT seems to deteriorate. It is because in this case, the threshold, which is chosen according to (9), becomes fairly large. Hence, thresholding of wavelet coefficients essentially removes a considerable amount of corrupting noises, but at the same time, it also suppresses the source signal. However, this phenomenon will disappear if we use a smaller threshold. Therefore, there is a great potential in exploiting the denoising method for many applications if we can classify the input signal and optimize the threshold accordingly.

It is perhaps worthy of note that should *a priori* information of the source signal be available, it is possible to derive a scale-dependent threshold that could provide better denoising performance. Preliminary results have confirmed this observation, and further details will be reported elsewhere.

VI. Conclusions

This correspondence considers the time delay estimation problem under a noisy environment. The source signal is assumed to be deterministic rather than stationary Gaussian process. A new method for time delay estimation that makes use of wavelet denoising is proposed. The new method is verified to be globally convergent with a high probability and is computationally efficient. Simulation results show its superiority to other conventional methods, including the commonly used direct cross correlation method.

Acknowledgment

The authors wish to thank the anonymous reviewers for their careful reading and constructive suggestions that improved the quality of this correspondence.

References

[1] G. C. Carter, *Coherence and Time Delay Estimation: An Applied Tutorial for Research, Development, Test, and Evaluation Engineers.* New York: IEEE, 1993.

[2] W. W. Smith and P. G. Steffes, "Time-delay techniques for satellite interference location system," *IEEE Trans. Aerosp. Electron. Syst.,* vol. 25, pp. 224–230, Mar. 1989.

[3] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-24, pp. 320–327, 1976.

[4] A. O. Hero and S. C. Schwartz, "A new generalized cross correlator," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-33, pp. 38–45, 1985.

[5] Y. T. Chan, J. M. F. Riley, and J. B. Plant, "A parameter estimation approach to time-delay estimation and signal detection," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-28, pp. 8–15, Feb. 1980.

[6] K. C. Ho and P. C. Ching, "A new constrained least mean square time-delay estimation system," *IEEE Trans. Circuits Syst.,* vol. 37, pp. 1060–1064, 1990.

[7] P. L. Feintuch, N. J. Bershad, and F. A. Reed, "Time delay estimation using the LMS adaptive filter—Dynamic behavior," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-29, pp. 571–576, June 1981.

[8] Y. T. Chan, *Wavelet Basics.* Boston, MA: Kluwer, 1995.

[9] I. Daubechies, *Ten Lectures on Wavelets.* Philadelphia, PA: SIAM, 1992.

[10] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika,* vol. 81, no. 3, 1994.

[11] L. L. Scharf, *Statistical Signal Processing.* New York: Addison-Wesley, 1991.

[12] S. Q. Wu and P. C. Ching, "Signal waveform restoration by wavelet denoising," in *Proc. Int. Symp. Time-Freq. Time-Scale Anal.,* Paris, France, 1996, pp. 89–92.

[13] K. Scarbrough, N. Ahmed, and G. C. Carter, "On the simulation of a class of time delay estimation algorithms," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-29, pp. 534–540, 1981.

[14] J. P. Ianniello, "Time delay estimation via cross-correlation in the presence of large estimation errors," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-30, pp. 998–1003, 1982.

[15] I. M. Johnstone and B. W. Silverman, "Wavelet threshold estimators for data with correlated noise," Tech. Rep., Dept. Statist., Stanford Univ., Stanford, CA, 1994.