APPROXIMATE MAXIMUM-LIKELIHOOD DELAY ESTIMATION
VIA ORTHOGONAL WAVELET TRANSFORM

Y. T. Chan*, H. C. So† and P. C. Ching‡
* Dept. of Electrical and Computer Engg., Royal Military College of Canada, Kingston, Ontario, Canada, K7K 5L0
† Dept. of Electronic Engg., City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong
‡ Dept. of Electronic Engg., The Chinese University of Hong Kong, Shatin, Hong Kong
e-mail: chan-y@banyan.rmc.ca, hcso@ee.cityu.edu.hk, pcching@ee.cuhk.edu.hk

ABSTRACT
A novel approximate maximum likelihood algorithm is proposed for estimating the time difference of arrival between signals received at two spatially separated sensors. Prior to cross correlation, one of the channel outputs is optimally weighted at different frequency bands with the use of an orthogonal wavelet decomposition. It can be viewed as a time domain implementation of the generalized cross correlation method. However, it does not suffer from the performance degradation due to the errors inherent in spectral estimation obtained from finite length data and is computationally efficient. Simulation results show that the proposed method outperforms direct cross correlation particularly when the noise level is high.

1 INTRODUCTION
Time delay estimation between signals received at two spatially separated sensors in the presence of noise has important applications such as direction finding, source localization and velocity tracking [1]. The receiver outputs, denoted by \( r_1(k) \) and \( r_2(k) \), are

\[
r_1(k) = s(k) + n_1(k) \]
\[
r_2(k) = s(k - D) + n_2(k), \quad k = 0, 1, ..., T - 1
\]  

where \( s(k) \) is the unknown source signal, \( n_1(k) \) and \( n_2(k) \) are the additive noises at the respective sensors, \( D \) is the difference in arrival times at the two receivers and \( T \) is the number of samples collected at each channel. Without loss of generality, the sampling interval is assigned to be unity. It is assumed that the source signal is stationary Gaussian and bandlimited between -0.5 and 0.5, while the corrupting noises are uncorrelated white Gaussian processes which are independent of \( s(k) \).

Many of the methods devised to estimate the delay \( D \) are related through a generalized cross correlation (GCC) approach [2]-[4]. The system block diagram of a generalized cross correlator is shown in Figure 1. It consists of a pair of receiver prefilters, \( H_1(f) \) and \( H_2(f) \), followed by a cross correlator. In general, the role of the prefilters is to enhance the frequency bands where the signal is strong and to attenuate the bands where the noise is excessive. The output of the correlator, \( J(\tau) \), is given by

\[
J(\tau) = \int_{-0.5}^{0.5} H_1(f) H_2^*(f) \mathcal{G}_{r_1 r_2}(f) e^{j2\pi f \tau} df
\]  

where \( * \) stands for the complex conjugate and \( \mathcal{G}_{r_1 r_2}(f) \) is the estimated cross-power spectrum between the finite data sequences \( r_1(k) \) and \( r_2(k) \). The delay estimate \( \hat{D} \) is equal to the time argument at which \( J(\tau) \) achieves its maximum value. When the prefilters \( H_1(f) \) and \( H_2(f) \) is chosen as follows [2],

\[
H_1(f)H_2^*(f) = \frac{G_{ss}(f)}{1 + \frac{G_{n_1 n_1}(f)G_{n_2 n_2}(f)}{G_{ss}(f)} + \frac{G_{ss}(f)}{G_{n_1 n_1}(f) + G_{n_2 n_2}(f)}}
\]  

where \( G_{ss}(f) \), \( G_{n_1 n_1}(f) \) and \( G_{n_2 n_2}(f) \) denote the auto-power spectra of \( s(k) \), \( n_1(k) \) and \( n_2(k) \) respectively, the delay variance will attain the Cramér-Rao lower bound (CRLB). This choice of \( H_1(f)H_2^*(f) \) is known as maximum likelihood (ML) weighting. Since the prefilters are dependent on the signal and noise spectra which are generally unknown, they have to be estimated from \( r_1(k) \) and \( r_2(k) \). Due to inaccuracies associated with estimating spectrum for finite data length, and hence the optimal weights, the maximum likelihood delay estimator is difficult to achieve in practice.

By assuming \( G_{ss}(f) \) as a piecewise constant function, a new approximate maximum likelihood (AML) prefilter based on fast wavelet transform is proposed in this paper. After wavelet decomposition of one of the receiver outputs, each subband sequence is multiplied by a constant factor which corresponds to AML weighting in that frequency band. The weighted subband components are then combined using inverse wavelet transform to construct the AML prefiltered signal. As a result, estimation of the signal and noise spectra is not needed and less computation is required because wavelet decomposition/reconstruction is performed only in one of the two channels.
2 THE PROPOSED ALGORITHM

In this section, we first derive the wavelet-based AML prefilter for time delay estimation. In order to reduce computation, \( H_2(f) \) is arbitrarily set to 1 so that weighting is required only for \( r_1(k) \), and \( H_1(f) \) is given by the left hand side of (3). The prefiltering procedure, as depicted in Figure 2, consists of three steps, namely, discrete wavelet decomposition, scaling of each subband sequence, and inverse wavelet transform. Let \( T = 2^L \cdot I \) where \( L \) denotes the level of decomposition and \( I \) is a positive integer. In the first level of decomposition, the sequence \( r_1(k) \) is broken down into two subband components \( c_1(k) \) and \( d_1(k) \) with lowpass and highpass filters and decimation. Mathematically, \( c_1(k) \) and \( d_1(k) \) are calculated using fast orthogonal wavelet transform [9],

\[
c_1(k) = \sum_{l=2k}^{2(N+k)-1} \hat{r}_1(l) g_{l-2k}
\]

\[
d_1(k) = \sum_{l=2(k-N+1)}^{2k+1} \hat{r}_1(l) h_{l-2k}
\]

where \( k = 0, 1, \ldots, T/2 - 1 \). The notation \( \hat{x} \) denotes the periodic extension of \( x \), and \( (g_i), i = 0, 1, \ldots, 2N - 1 \), represent the \( 2N \)-tap quadrature mirror filter coefficients [6]. The corresponding highpass filter coefficients are given by \( h_i = (-1)^i g_{i+1} \) for \( i = -2N + 2, -2N + 3, \ldots, 1 \). Similarly, the lowpass data stream is decomposed \((L - 1)\) successive times until the wavelet coefficients \( d_1(k), d_2(k), \ldots, d_L(k) \) and the residue component \( c_L(k) \) are obtained. Notice that the length of each output sequence is halved after each decomposition and thus the sum of the total number of subband samples is equal to \( T \).

To derive the AML weights, it is assumed that the signal component in each subband sequence can be approximated by a flat spectrum. This implies that the auto-power spectrum of \( s(k) \) as well as the optimal weighting function are modeled by some piecewise constant functions. This assumption is not restrictive as it has been shown [7] that a Eckart prefilter, which aims to maximize the output signal-to-noise ratio (SNR) of \( J(\tau) \) at \( \tau = D \) in the limit of low input SNR, can be substituted by its piecewise constant versions with good performance. The AML filtering is performed by scaling each subband signal with a finite constant. As a result, spectral estimation of the signal and noise spectra which may introduce large delay variances is prevented. In fact, the weights are determined in the time domain of the multirate system level-by-level. Using (3) for piecewise constant \( G_{ss}(f) \), the weight \( w_{d_1} \) is computed as

\[
w_{d_1} = \frac{\hat{\sigma}_{a_1}^2}{\hat{\sigma}_{n_1}^2 + \hat{\sigma}_{a_1}^2 (\hat{\sigma}_{n_1}^2 + \hat{\sigma}_{n_2}^2)}
\]

where \( \hat{\sigma}_s^2 = \arg\max_{\tau} \hat{R}_{r_1r_2}(\tau), \hat{\sigma}_{a_1}^2 = 2\hat{\sigma}_s^2 + \hat{\sigma}_{n_1}^2 - \hat{R}_{c_1c_1}(0), \hat{\sigma}_{n_1}^2 = \hat{R}_{r_1r_1}(0) - \hat{\sigma}_s^2 \) and \( \hat{\sigma}_{n_2}^2 = \hat{R}_{r_2r_2}(0) - \hat{\sigma}_s^2 \).

The function \( \hat{R}_{xy}(\tau) \) denotes the estimated cross correlation of \( x(k) \) and \( y(k-\tau) \) using finite data, \( \hat{\sigma}_s^2 \) is the estimate of the total signal power, \( \hat{\sigma}_{a_1}^2 \) is the estimated power of \( d_1(k) \) and \( \hat{\sigma}_{n_1}^2 \) and \( \hat{\sigma}_{n_2}^2 \) represent the estimated noise powers. Similarly, the other weights can be sequentially calculated. Note that at each level, knowing only either the lowpass or highpass component power is sufficient because their sum is equal to twice the power of the lowpass sequence at the previous stage.

The modified subband sequences \( d_1^1, d_2^1, \ldots, d_L^1 \) and \( c_L^1 \) are then combined to form the AML prefiltered signal \( r_1^\tau(k) \). It is the inverse process of the filter bank decomposition and starts from the formation of new
lowpass sequence $c_{L-1}^n(k)$ which is given by [5]

$$c_{L-1}^n(k) = \sum_{l=-\frac{1}{2} - N+1}^{\frac{N}{2} - 1} c_{L}^n(l)g_{k-2l} + \sum_{l=-\frac{1}{2}}^{\frac{N}{2} - 1} d_{L}^n(l)h_{k-2l}$$  \hspace{1cm} (6)

$$c_{L-1}^n(k) = \sum_{l=\frac{1}{2} - N}^{\frac{N}{2} - 1} c_{L}^n(l)g_{k-2l} + \sum_{l=\frac{1}{2}}^{\frac{N}{2} - 1} d_{L}^n(l)h_{k-2l}$$

where the first formula is for even $k$ while the second is for odd $k$. The sequence $r_1^2(k)$, which is equivalent to $c_0^n(k)$, is obtained by using (7) $L$ times. Finally, the AML delay estimate is given by the peak of the cross correlation function of $r_1^2(k)$ and $r_2(k)$.

3 SIMULATIONS & DISCUSSIONS

Simulation tests had been carried out to evaluate the performance of the proposed algorithm by comparing the mean square delay errors with those of direct cross correlation method and the CRLB. Four experiments with different signal spectra described in Figure 3 were presented. One of the spectra corresponded to the white noise case while the others were nonwhite but with piecewise constant values. To obtain the two sensor outputs, three uncorrelated white Gaussian noise sequences were first generated. Two of these sequences were used to represent the additive noises while we passed the remaining one through specific FIR filters to produce the desired spectra of $s(k)$. The source signal has unity power and different SNRs are obtained by proper scaling of the random noise sequences. For simplicity but without loss of generality, the additive noises were assigned to have identical power. The time delay parameter $D$ was set to 1.0 and the data length $T$ was 4096. Moreover, 16-tap Daubechies wavelet filter coefficients were used in order to provide sharp frequency responses for each subband decomposition. The mean square delay errors obtained were based on 200 independent runs.

Spectrum A: One-level and two-level wavelet decompositions were examined for the new prefilter. It can be seen from Figure 4 that all methods attained the CRLB for a large range of SNR. The two versions of the proposed method had very similar performance with the direct cross correlator which gave the ML estimates for white source signal, except that 2-level decomposition case advanced the threshold effect by 1 dB. It is because at the lower levels of decomposition, the smaller number of subband components results in larger variances in the power estimates.

Spectrum B: The signal power was now concentrated between 0 and 0.25. From Figure 5, we see that the proposed method with one decomposition was the best. Similar to the previous test, its delay variances were very close to the one using 3 subbands apart from the threshold effect. Furthermore, it had a few dB improvement over the cross correlator for small SNR even if the latter attained the CRLB for high SNR.

Spectrum C: The auto-power spectrum of $s(k)$ consisted of 3 piecewise constant functions and thus only the 2-level decomposition was used. It is observed from Figure 6 that the new AML method outperformed the direct cross correlator for whole range of SNR, although the improvement was much apparent for the low SNR conditions.

Spectrum D: The power spectrum of the source signal had the peak value of 25/7 from 0.125 to 0.25. Figure 7 shows that the proposed method provided significant improvement over the cross correlation method when the SNR was small, even though they both had threshold effects at the SNR of -10 dB.

By using fast orthogonal wavelet transform, a novel AML delay estimator based on the GCC approach is derived. Unlike conventional GCC methods, errors introduced in spectral estimates using finite sensor outputs are avoided in the proposed method. By considering the optimum weighting function as piecewise con-
constant, optimum prefiltering can be achieved by scaling the wavelet coefficients with some constant values. Extensive computer simulation with different signal spectra, including the above four experiments, show that the algorithm is stable and gives good time delay resolution. In general, this estimator has significant improvement over the direct cross correlation method particularly for small SNR and its variance is close to the CRLB. It is also worthy to note that it should perform much better than the AML delay estimator using spectral estimation which has been shown to have similar performance with standard cross correlation [8].

Since the proposed method may give earlier threshold effect when the number of decomposition levels is greater than required, our future research will concentrate on automatically determining the optimum number of wavelet decomposition. We will consider the cases when the corrupting noises are nonwhite where in these cases the noise powers may have to be determined band-by-band instead of from $r_1(k)$ and $r_2(k)$. Biorthogonal filter banks [9], which have more design flexibility, will also be investigated in order to reduce the number of filter coefficients and/or improve the sharpness of the frequency responses.

**ACKNOWLEDGEMENT**

This work is partially supported by a research grant awarded by the Hong Kong Research Grant Council.

**References**


