High Resolution Feature Extraction from Reverberation Data

Robert B. MacLeod*, Donald W. Tufts**

*Naval Undersea Warfare Center
Newport R.I. 02841 USA
macleod@fs811.npt.nuwc.navy.mil

**University of Rhode Island
Kingston R.I. 02881 USA
tufts@ele.uri.edu

Abstract

A new method of signal approximation is presented which has the capability to model complex active returns due to reverberation. The method explored herein is data adaptive in the sense that high resolution processing is attempted where necessary, while fast algorithms are used to estimate the remaining components. The basic goal is to achieve accurate and unbiased estimation of the delays and time scalings of a group of closely spaced signal returns. A secondary goal is to achieve this estimation performance without sacrificing the computational efficiency of the fast algorithms available.

1. Introduction

We will be treating the problem of feature extraction from high frequency sonar data, generated by an active transmission. Our viewpoint will be that the portion of returning signal due to backscattering in the environment can be adequately approximated as the sum of frequency shifted (time scaled) and delayed replicas of the transmit. This is to be contrasted with the statistical approach taken in references [1, 2]. We will present means to make signal approximations with these assumptions, motivating alternative post-processing algorithms which use these features vice the original data sequence.

This type of approximation becomes more valid for the higher acoustic frequencies of the problem considered here. In particular, for the frequency range of interest, 15 kHz to 25 kHz, propagation can be approximated by ray models, and propagation through the bottom is of diminished importance. Interaction of a ray with the surface or bottom can be approximated as an attenuation of the path, along with the direction change and possible signal inversion. Target returns, under these conditions, will arrive at the receiver with the same temporal relationships within the return as were generated at the target. This will also hold true for returns due to prominent features in the scattering environment, due to structured discontinuities on the bottom or surface. Wavelengths are on the order of magnitude of 1/4 foot, which in combination with the wider-bandwidth signals being transmitted, result in resolution of target features down to 5 feet with traditional processing methods.

1.1. Signal Model

With these considerations in mind, the returned signal will be assumed to be a sampled, vectorized representation of the signal

$$\begin{aligned} x &= \sum_{i=1}^{K_S} \tilde{A}_i s(a_i t - \tau_i) + \sum_{i=K_S+1}^{K_S+K_I} \tilde{A}_i s(a_i t - \tau_i) + n \\ &= x_S + x_I + n \ . \end{aligned} \tag{1}$$

The signal x_S, and the interference x_I, are both modeled as linear combinations of time delayed (τ_i) and time scaled (a_i) replicas of the transmit, s. The signals modeled will be taken as a complex valued (in general), as is typical for sonar systems processing in a single sideband. The signal and interference amplitudes, delays, and time scaling factors will be considered to be deterministic but unknown parameters, the amplitudes $\tilde{A}_i$ complex valued. The noise will be assumed to be complex Gaussian, i.e.

$$n \sim \mathcal{N}\left(0, (\sigma^2/2)I\right) + j\mathcal{N}\left(0, (\sigma^2/2)I\right) \ . \tag{2}$$

1.2. Parameter Estimation

For the case of the measurement noise of equation 2, the maximum likelihood estimate of the signal parameters can be made by maximizing the compressed likelihood function (CLF)

$$\begin{aligned} \mathsf{L}_{\mathrm{clf}} &= x^t * \mathbf{H}(\mathbf{H}^t\mathbf{H})^{-1}\mathbf{H}^t * x \\ &= \|\mathbf{P}_{\mathbf{H}}x\|^2 \ , \end{aligned} \tag{3}$$

searching over the dimension $2(K_S + K_I)$ space of the vector model parameters a and τ. The compressed likelihood is a compact notation describing the likelihood that given the data sequence x, parameters $[a_i, \tau_i]$ generated the sequence. Maximizing the compressed likelihood function results in the maximum likelihood estimates for the time scale and delay of the components. The amplitudes are functionally related to the time scale and delay estimates, in fact, given the maximum likelihood estimates of $[a_i, \tau_i]$, the maximum likelihood est imates of the amplitudes (vector notation) are

$$[\ \tilde{A}_1 \quad \tilde{A}_2 \quad \ldots \quad \tilde{A}_{K_S+K_I}\ ]' = (\mathbf{H}^t\mathbf{H})^{-1}\mathbf{H}^t * x \ . \tag{4}$$

A complete treatment of the topic is offered in reference [3]. Here, $\mathbf{H}$ is the combined signal and interference model matrix, formed by stacking together in columns time delayed and scaled transmit replicas.

$$\mathbf{H} = [s(a_1 t - \tau_1) \ldots s(a_{K_S} t - \tau_{K_S}) \ldots s(a_{K_S+K_I} t - \tau_{K_S+K_I})] \tag{5}$$

Since both the signal and the interference are modeled in exactly the same way, it will not be necessary to distinguish the components, and we will refer to the model matrix as simply $\mathbf{H}$.

1.3. Parameter Estimation as Feature Extraction

We will make the somewhat unorthodox assumption that all signal components in $\mathbf{H}$, due either to target returns or environmental scattering, are high signal to noise components. A more typical approach would be to focus on the strength of target components with respect to some measure of the strength of the reverberant field. In our work, we seek to estimate all components due either to reverberation or target scattering, and as such, they compete only against the measurement noise (and each other). The noise will be simple measurement noise, such as is generated by analog preamplification and 16 bit sampling. With this viewpoint, we refer to components due to the reverberant field as 'features' of the field, in the sense that any post processing desired can use the extracted features in lieu of the data itself.

2. Cramer-Rao Lower Bound on Variance of Model Parameters

Given the high signal to noise ratio of the model components, very accurate estimation of 'single' components is possible, that is components that are orthogonal to all other components. For more complicated portions of the signal, for example a target with several closely spaced highlights, estimation becomes more difficult. To illuminate this, the Cramer-Rao lower bound on variance of estimated model parameters for a pair of linear frequency modulated signals has been calculated. The signals modeled are similar to that transmitted in sea trials here in Narragansett Bay, presented later in this paper. Each signal component is of the form

$$s(A_i, \phi_i, a_i, \tau_i) = w(a_i n - \tau_i) A_i e^{j\phi_i} e^{j2\pi(c+d(a_i n-\tau_i))(a_i n-\tau_i)} \ . \tag{6}$$

The function $w(a_i n - \tau_i)$ is a real valued windowing function, a_i and τ_i time scale and delay the i^{th} signal component. The parameters c and d are selected to sweep the fm signal by a rate of $k = 2d = (\Omega_2 - \Omega_1)/(N-1)$, from digital frequency Ω_1 to Ω_2. The bound can be calculated by evaluating Fisher's Information Matrix for complex processes [4],

$$[\mathbf{I}(\xi)]_{ij} = \mathrm{tr}\left[\mathbf{C}_x^{-1}(\xi)\frac{\partial \mathbf{C}_x(\xi)}{\partial \xi_i}\mathbf{C}_x^{-1}(\xi)\frac{\partial \mathbf{C}_x(\xi)}{\partial \xi_j}\right] + 2\mathrm{Re}\left[\frac{\partial \mu^H(\xi)}{\partial \xi_i}\mathbf{C}_x^{-1}(\xi)\frac{\partial \mu(\xi)}{\partial \xi_j}\right] \tag{7}$$

For our example, the covariance matrix $\mathbf{C}_x(\xi)$ does not depend on the modeling parameters ξ, in fact $\mathbf{C}_x(\xi) = \sigma^2\mathbf{I}$, which simplifies the evaluation of 7. The remaining parameters are

$$\mu(\xi) = s(A_1, \phi_1, a_1, \tau_1) + s(A_2, \phi_2, a_2, \tau_2) \tag{8}$$

$$\xi = [\ A_1 \quad \phi_1 \quad a_1 \quad \tau_1 \quad A_2 \quad \phi_2 \quad a_2 \quad \tau_2\ ]^t \ . \tag{9}$$

The partial derivatives, neglecting minor perturbations due to the windowing, will be

$$\begin{aligned}
\frac{\partial \mu(\xi)}{\partial A_i} &= \frac{1}{A_i} s(A_i, \phi_i, a_i,\ tau_i) \\
\frac{\partial \mu(\xi)}{\partial \phi_i} &= j s(A_i, \phi_i, a_i, \tau_i) \\
\frac{\partial \mu(\xi)}{\partial a_i} &= j 2\pi \left(cn + 2da_i n^2 - 2dn\tau_i\right) s(A_i, \phi_i, a_i, \tau_i) \\
\frac{\partial \mu(\xi)}{\partial \tau_i} &= j 2\pi \left(-c - 2da_i n + 2d\tau_i\right) s(A_i, \phi_i, a_i, \tau_i)\ .
\end{aligned} \tag{10}$$

The variance bounds are found by inverting 7,

$$\sigma_{\xi_i}^2 \geq \left[\mathbf{I}^{-1}(\xi)\right]_{ii} \tag{11}$$

An example has been prepared, using a 50 sample fm signal, sweeping from digital frequencies $\Omega = -0.2$ to $\Omega = +0.2$. Two components are placed close together, and the variance bound of the first component's parameters calculated as the second component is moved closer. The first component has a time scale of $a_1 = 1$, and a delay of $\tau_1 = 0$. Both components have real amplitude $A = 1$, and the signal to noise ratio is 0 dB. These figures

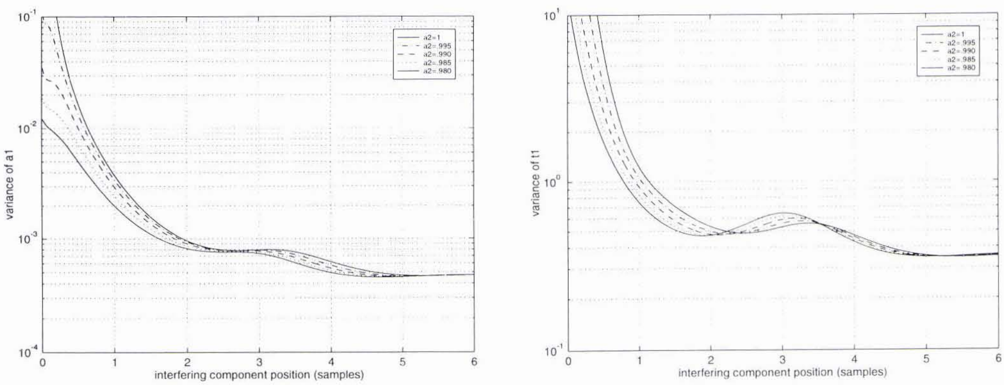

Figure 1: Variance of time scale parameter (left) and time delay parameter (right) of a component placed at $(a = 1, \tau = 0)$,with a second component in the vicinity. Abscissa is the separation of the components, a family of curves are shown for differing values of the second component's time scale, a2. Signal to noise ratio is 0dB

are shown on an expanded scale for the separation of the two components, to show some of the structure on the variance bounds. Basically, as the second component moves closer to the component of interest, the variance of the estimated parameters increases. If the two components have differing time scale parameters, the degradation of the estimation is less severe, since the two signals will not become congruent even as the separation goes to zero. A second case was run, to calculate usable bounds on estimation for the data presented later in this paper. Two components were again placed close together, this time each with a signal to noise ratio of 90dB, each component at a time scale of $a_i = 1$. The results are summarized in table 1, where component separations resulting in the Cramer-Rao bound exceeding a given error tolerance (standard deviation) are shown.

3. Estimation Biasing When the Number of Components is Unknown

In the previous section, estimation variance bounds were calculated for the situation where all the components in the model were accounted for. If an unknown component is present, the problem becomes more difficult, since the error bounds are functionally related to the position and parameters of the unknown component. One approach to calculating the Cramer-Rao bounds on the known component would be to treat the parameters of the unknown component as random variables, and calculate expected values for the bounds. We will not pursue this result

parameter	typical value	tolerance	separation limit
A	1	$\sigma_A = 0.1$	.0011
ϕ	0	$\sigma_\phi = 1$ deg	.0011
a	1	$\sigma_a = .001$	.0021
τ	0	$\sigma_\tau = .0076$	.0076

Table 1: Typical resolution performance for a 40% band 50 sample LFM Signal, 90 dB above measurement noise. Units for separations and τ are in samples.

here, that is we are more interested in detecting situations where an unknown component may be present, and concentrating the effort of the estimation algorithms in these areas.

With this in mind, we have calculated the estimation biasing which occurs when an unknown component nears the location of a component whose parameters we are estimating. This biasing occurs because (in general)

$$\underset{a_1,\,\tau_1}{\arg\max}\ \{x^t * \mathbf{H}_1(\mathbf{H}_1^t\mathbf{H}_1)^{-1}\mathbf{H}_1^t * x\} \neq \underset{a_1,\,\tau_1}{\arg\max}\ \{x^t * \mathbf{H}_{1,2}(\mathbf{H}_{1,2}^t\mathbf{H}_{1,2})^{-1}\mathbf{H}_{1,2}^t * x\}\ , \qquad (12)$$

where $\mathbf{H}_{1,2}$ is the model matrix with a second component added. Continuing the analysis of the 50 sample LFM signal we are interested in, we have run Fast Maximum Likelihood Estimation on a series of simulated returns, when a second unknown component is present. The measurement noise for this example is zero, as we are interested in the biasing of the estimation on the first component. This is equivalent to evaluating the right side of equation 12. Figure 2 shows why it is dangerous to take literally the time scale estimates of a Maximum Likelihood Estimation

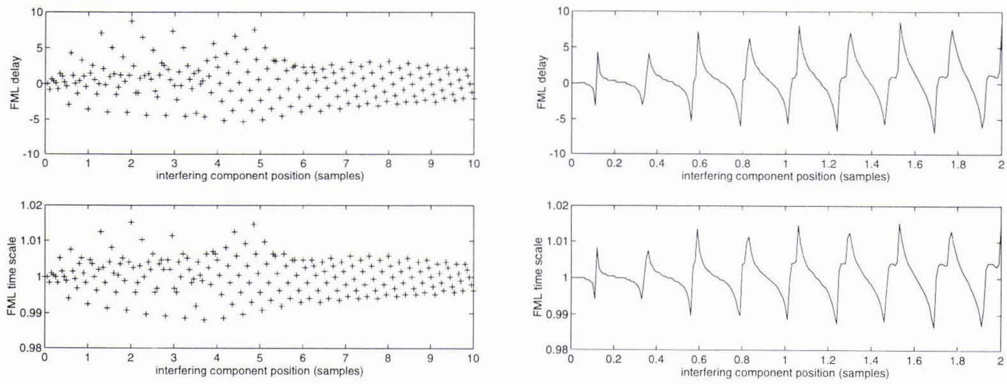

Figure 2: 50 Sample FM Pulse, -0.2 to +0.2 freq. Top left are the delay estimates and bottom left are the time scale estimates for the first FML component. A second unknown is placed near the estimated component, at the indicated delay separation. The correct estimates of the first component are ($a = 1, \tau = 0$) in all cases. On the right are close-ups of the biasing.

procedure. If target or channel features are present with closely spaced time delays, and the estimation is forced to model these features with fewer components than required, the time scale estimates can be significantly biased. In this example, two zero Doppler components can generate an estimate with a time scale approaching $a_i = 1.02$, wh ich, for the sonar system we are modeling, corresponds to a feature moving at approximately 50 feet per second.

4. High Resolution Parameter Estimation

With the fundamental limits on parameter variance and bias in mind, the basic problem to be solved is to decide if any single component of an approximated signal may actually be due to two underlying components. We have shown that two very closely spaced components can generate an estimate of a single component with a radically biased time scale estimate, even though the Cramer-Rao bounds on variance (for the two underlying components) are quite low. We will attempt to exploit this situation, by detecting unusual time scale estimates, and making appropriate substitutions.

One simple approach would be to let the Fast Maximum Likelihood method continue to add components to the estimation, until the signal is resolved. This may result, however, in excessive components being added to the model, before the features we are interested in are resolved. We propose an alternative method, based on maximum posterior probability, a test

$$\max_{\xi} \{\mathsf{L}_{\mathsf{lf}}(x|p(a_1, a_2))\} \overset{?}{>} k \max_{\xi_0} \{\mathsf{L}_{\mathsf{lf}}(x|p(a_0))\} . \tag{13}$$

We will apply this test to components generated from a Fast Maximum Likelihood analysis, focusing on estimates which have unlikely time scale parameters. This situation, of course, drives the right hand side of equation 13 down. This test checks, component by component (s_0), whether it is more likely that two components ($s_1 + s_2$), generated the data sequence x. The notation $\mathsf{L}_{\mathsf{lf}}(x|p(a_0))$ means the likelihood function, modified for a probabilistic prior on the estimated parameter a_0. We are now treating the time scale parameters a_i as random variables, with a multidimensional probability density function with the following properties,

- probability is concentrated near $a_i = 1$
- unlikely that two very closely spaced components have dissimilar time scales, i.e. $a_1 \approx a_2$.

This situation would apply when data were collected with a relatively static platform, in calm seas. Reverberation highlights could have some Doppler (time scale), but since nearby highlights would be moving in unison, the time scale parameters for each highlight would be identical. Let $\mathbf{H}_0 = [s_0]$, $\mathbf{H}_1 = [\ s_1 \quad s_2\ ]$, and the parameter ve ctor ξ be broken up into the linearly entering parameters $\theta_0 = [A_0 e^{j\phi_0}]$, $\theta_1 = [\ A_1 e^{j\phi_1} \quad A_2 e^{j\phi_2}\ ]^t$, and the non-linearly entering parameters $\xi_0 = [\ a_0 \quad \tau_0\ ]^t$, and $\xi_1 = [\ a_1 \quad \tau_1 \quad a_2 \quad \tau_2\ ]^t$. Then,

$$\mathsf{L}_{\mathsf{lf}}(x|p(a_1, a_2)) = \frac{1}{(2\pi)^N |\mathbf{C}|} e^{-(x - \mathbf{H}_1\theta_1)^h \mathbf{C}(x - \mathbf{H}_1\theta_1)} p(a_1, a_2) , \tag{14}$$

$$\mathsf{L}_{\mathsf{lf}}(x|p(a_0)) = \frac{1}{(2\pi)^N |\mathbf{C}|} e^{-(x - \mathbf{H}_0\theta_0)^h \mathbf{C}(x - \mathbf{H}_0\theta_0)} p(a_0) , \tag{15}$$

and

$$\begin{aligned} p(a_1, a_2) &= \delta(a_1 - a_2) p_a(a_1) \\ p(a_0) &= p_a(a_0) . \end{aligned} \tag{16}$$

Then, the test 13 would be equivalent to, letting $\mathbf{C} = \sigma^2 \mathbf{I}$,

$$\max_{\xi_1'} \left\{ \ln p_a(a_1) + \frac{1}{\sigma^2} ||\mathbf{P}_{\mathbf{H}_1} x||^2 \right\} \overset{?}{>} k' + \max_{\xi_0} \left\{ \ln p_a(a_0) + \frac{1}{\sigma^2} ||\mathbf{P}_{\mathbf{H}_0} x||^2 \right\} , \tag{17}$$

where $\xi_1{}' = \left[\ a_1 \quad \tau_1 \quad a_1 \quad \tau_2\ \right]^t$. This test has been re-written by noting that the maximization of $\ln p_a(a_i)$ does not depend on θ_i, and substituting the maximum likelihood estimate (MLE) for θ_i into each side of the equation. It is well known that substitution of the MLE will maximize this form, see for example [5]. This likelihood test, equation 17, will be the test we are primarily concerned with. The simplest case would be a uniform prior on the time scale parameter, $p_a(a) = (a_U - a_L)^{-1} \mathsf{U}(a_L, a_U)$. For this case, time scale estimates outside the limits of $p_a(a)$ drive either side of equation 17 to $-\infty$, and the likelihood test is reduced to, dropping the record keeping on the different likelihood thresholds 'k',

$$\max_{a_1, \tau_1, \tau_2} ||\mathbf{P}_{\mathbf{H}_1} x||^2 \overset{?}{>} k + \max_{a_0, \tau_0} ||\mathbf{P}_{\mathbf{H}_0} x||^2 . \tag{18}$$

5. Data Analysis

We have applied our algorithm to a data set collected in Narragansett Bay, in November of 1994. In this test, a series of linear frequency modulated signals were transmitted against a field of targets situated in about 120 feet of water. This was a static test, that is neither the transmitter, receiver, or the targets were moving. We have chosen a transmit which, in the decimated signal domain, lasts about fifty samples, and sweeps from digital frequencies of approximately $\Omega = -0.2$ to $\Omega = +0.2$. Records of reverberation data were extracted from the raw data by beamforming the data away from the targets, and selecting data segments of manageable record size. We have implemented the equation 17 test as follows.

- Perform Fast Maximum Likelihood estimation on a data record.
- Identify 'singleton' components with unreasonable time scale.
- Perform the likelihood test, equation 17, searching over 2-d space of $(\Delta\tau, a)$, where $\tau_1 = \tau_0(a) - 1/2\Delta\tau$, and $\tau_2 = \tau_0(a) + 1/2\Delta\tau$. Replace the single component with two components, if the test 17 is satisfied. τ_0 is a function of a, and follows the likelihood ridge passing through the single component's location.

Figures 3 and 4 show a typical result of the analysis. In figure 3, the reverberation signal has been approximated using Fast Maximum Likelihood. The analysis has been purposely stopped short, that is only the most significant components of the signal were approximated to avoid over-resolving the signal. A prior probability on the time scale parameter was not invoked, the algorithm was free to fit using any time scale between $a_i = .98$ and $a_i = 1.02$. Figure 4 shows the result of implementing the likelihood test of equation 17. We chose the time scale threshold to be $|a_i - 1| = .005$, and the noise parameter σ to be equivalent to the residual of the Fast Maximum Likelihood analysis. The left hand side of figure 4 shows the wide band cross ambiguity function of the reconstructed signal,

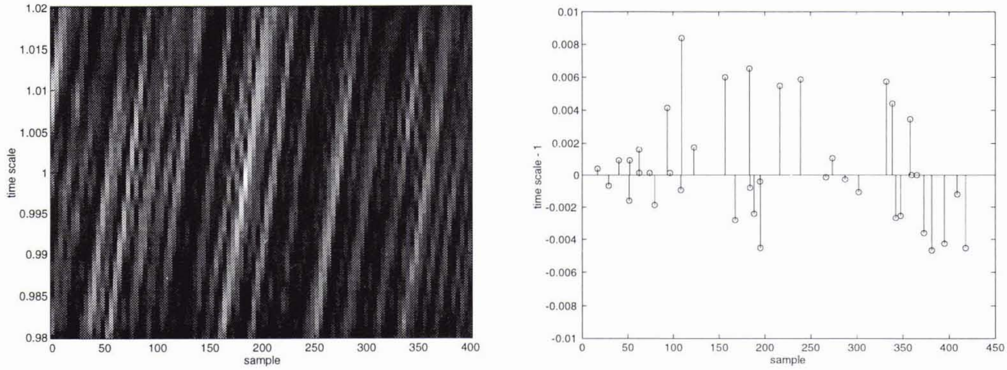

Figure 3: Fast Maximum Likelihood Analysis Result of High Frequency Reverberation Data. Left is wide-band cross ambiguity function of a section of return with the transmit. Right are the delays and time scalings which represent this signal.

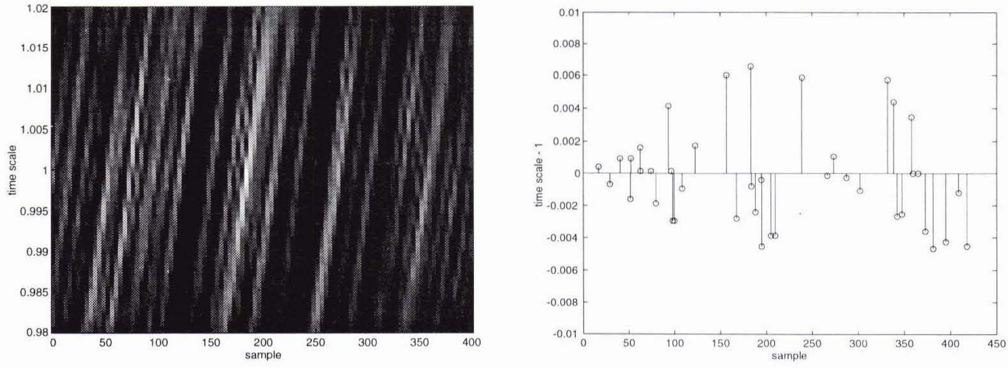

Figure 4: High Resolution Analysis Result of High Frequency Reverberation Data. Left is wide-band cross ambiguity function of reconstructed signal with the transmit. Right are the delays and time scalings which represent this signal.

to be contrasted with the left hand side of figure 3. The high resolution test has replaced two of the Fast Maximum Likelihood components with closely spaced components at more likely time scales. The reconstructed signal retains the important features of the original signal, especially the highly structured nature of the wide band ambiguity function.

6. Conclusion

We have shown a method to extend the performance of Fast Maximum Likelihood estimation, by performing localized likelihood searches on individual signal components. This method can be used to introduce prior probability information to the estimation procedure, particularly information related to propagation and the environment. We hope to extend this work to address probabilistic scattering models for structured targets present in the reverberation.

References

[1] D. Middleton, "A statistical theory of reverberation and similar first-order scattered fields, part 1: Waveforms and the general process," *IEEE Transactions on Information Theory*, vol. IT-13, no. 3, pp. 372–392, July 1967.

[2] D. Middleton, "A statistical theory of reverberation and similar first-order scattered fields, part 2: Moments, spectra, and special distributions," *IEEE Transactions on Information Theory*, vol. IT-13, no. 3, pp. 393–414, July 1967.

[3] D. W. Tufts, H. Ge, and S. Umesh, "Fast maximum likelihood estimation of signal parameters using the shape of the compressed likelihood function," *IEEE Journal of Oceanic Engineering*, vol. 18, no. 4, pp. 388–399, October 1993.

[4] S. M. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1993.

[5] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Transactions on Signal Processing*, vol. 42, pp. 2146–2157, August 1994.