Multipath Cancellation Using a Maximum Likelihood Metric Space

John E. Piper
NSWC PCD, Panama City, FL

***Abstract* – Multipath signals occur in many sonar, radar, and communication applications. It is, of course, generally desirable to eliminate these unwanted signals. Traditional array signal processing techniques often have trouble eliminating these signals when only a small number of array elements are available. The maximum likelihood method explicitly models multiple signals in its mathematical construction. This feature effectively opens up a multidimensional space that allows the desired direct path signal to be completely decoupled from the multipath signals. Examples with both narrowband and broadband signals are presented.**

Introduction

Shallow water sonars can expect to see multipath signals that reflect off the sea surface and the sea bottom. An example of at-sea data collected with a broadband synthetic aperture sonar is shown in fig. 1. The faint direct path is followed by five undesirable multipath signals.

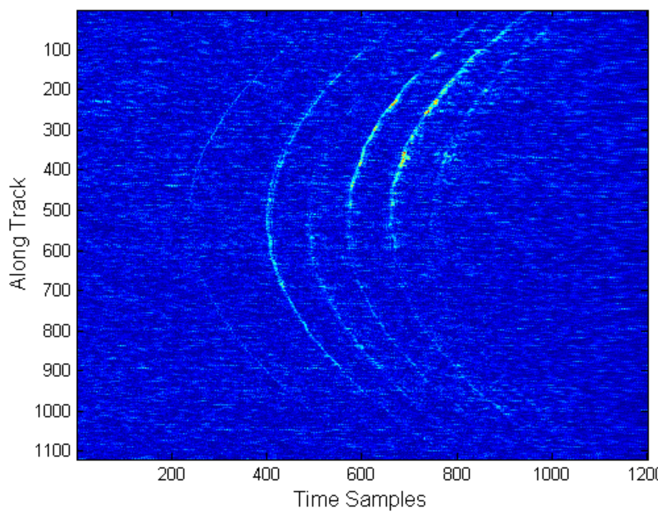

Fig. 1. Sonar data with direct path followed by five multipath signals.

Cancelling multipath signals is a challenging signal processing problem. This is especially true when only a small number of vertical array elements are available, since conventional processing results in beam patterns that are too large to separate the signals. This problem requires a more powerful approach.

The maximum likelihood method is uniquely able to solve this type of problem because it operates in a multidimensional space. In this space it is possible to completely separate or decouple these signals.

Likelihood Function

The maximum likelihood method is an elegant and powerful approach to signal processing problems. It is essentially a statistical approach that attempts to fit the data to a parametric model. This model can be written as

$$y = D\,s + noise \tag{1}$$

where y is the observed array data, s is the signal, and D is the mapping of the data onto the array. This mapping is also known as a steering vector or matrix.

Given y it is possible to solve for s. This is an inverse problem. Assuming Gaussian noise, the general solution can be obtained using a least-squares method [1].

$$\mathrm{L} = \| y - D\,s \|^2 \tag{2}$$

The distance measure between the data, y, and the parametric model, L, is called the likelihood function. Minimizing this function results in the best statistical estimate of the model parameters.

A. One-Signal Model

It is instructive to first consider the one-signal representation of the maximum likelihood method. The steering vector for a uniform linear array with n elements can be written as

$$D = \begin{pmatrix} 1 & e^{i\phi} & \cdots & e^{i(n-1)\phi} \end{pmatrix}^{\mathrm{T}} \tag{3}$$

Where the phase angle, ϕ, is given in terms of the incidence angle, ϑ, element separation, a, and wavelength, λ, as

$$\phi = \frac{2\pi}{\lambda}\, a \sin(\vartheta)\,. \tag{4}$$

This phase angle is the radian measure of the fraction of a wavelength extra distance that the signal travels.

The likelihood function is constructed as the least-squares difference between the measured data and the parametric model.

$$\mathrm{L} = \| y - D\,s \|^2$$
$$\mathrm{L} = (\, y^\dagger - D^\dagger s^* \,)\,(\, y - D\,s \,)$$
$$\mathrm{L} = y^\dagger y - y^\dagger D\,s - D^\dagger y\,s^* + D^\dagger D\,s^*\,s$$

Minimizing the likelihood function with respect to s yields

$$\partial \mathrm{L} / \partial s = 0 = -\,y^\dagger D + D^\dagger D\,s^*$$

or

$$s = D^{\dagger} y / n \tag{5}$$

since

$$D^{\dagger} D = n. \tag{6}$$

Substituting s in (5) into the likelihood function (2) yields

$$\mathrm{L} = y^{\dagger} y - \| D^{\dagger} y \|^2 / n.$$

In the minimization process, constants such as $y^{\dagger} y$ and n can be ignored. It is also convenient to flip the overall sign and make it a maximization process. The best estimate of the direction parameter is found at the maximum of this function, which is simply the periodogram.

$$\mathrm{L} = \| D^{\dagger} y \|^2 \tag{7}$$

Equation (5) indicates that the desired signal, s, can be found by steering the array toward the direction of s. This is the classic signal processing result, which is based on a one-signal model.

This one-signal approach has well-known problems when there is more than one signal. Essentially, other signals can cause interference by "leaking" through the sidelobes or mainlobe. With enough array elements, it is sometimes possible to put "nulls" in the direction of the interfering sources. However, this is an inelegant and incorrect use of the mathematics.

B. Multiple Signal Model

To correctly process multiple signals it is absolutely necessary to correctly model the problem. This means that multiple signals need to be explicitly included in the model. Equation (1) can be written with multiple signals as

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = D \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_m \end{pmatrix} + noise$$

where y is the observed array data of length n, s is the signals vector of length m, and D is the multiple-signal steering matrix.

$$D = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ e^{i\phi_1} & e^{i\phi_2} & \cdots & e^{i\phi_m} \\ \vdots & \vdots & \ddots & \vdots \\ e^{i(n-1)\phi_1} & e^{i(n-1)\phi_2} & \cdots & e^{i(n-1)\phi_m} \end{pmatrix} \tag{8}$$

The likelihood function is constructed as in (1)

$$\mathrm{L} = \| y - D\, s \|^2$$

Minimizing this function follows the same procedure as in the one-signal case. The important difference is that $D^{\dagger}D$ is now a matrix and not a simple scalar. So, in (5) instead of simply dividing by n, it is necessary to multiple by the inverse of $D^{\dagger}D$.

$$s = (D^{\dagger}D)^{-1} D^{\dagger} y \tag{9}$$

Inserting (9) into the likelihood function (2) yields the following multiple-source representation.

$$\mathrm{L} = y^{\dagger} y \; - \; y^{\dagger} D\, (D^{\dagger}D)^{-1} D^{\dagger} y$$

Again, it is convenient to drop the constant $y^{\dagger} y$ term and flip the sign to make it a maximization problem.

$$\mathrm{L} = \; y^{\dagger} D\, (D^{\dagger}D)^{-1} D^{\dagger} y \tag{10}$$

C. Direction-of-Arrival Estimation

Equation (10) is often modified to utilize the sample covariance matrix, R. Under the trace operator it is possible to rotate the $y^{\dagger}$ term to the end of (10). The resulting $y\, y^{\dagger}$ term is the sample covariance matrix

$$\mathrm{L} = \mathrm{tr}(\mathrm{L}) = \mathrm{tr}(\, D\, (D^{\dagger}D)^{-1} D^{\dagger} y\, y^{\dagger}\,) = \mathrm{tr}(\, D\, (D^{\dagger}D)^{-1} D^{\dagger}\, \mathrm{R}\,)$$

This representation of the likelihood function has the advantage of allowing many observations to accurately estimate the sample covariance matrix. This is useful in order to accurately estimate the direction-of-arrival angles and thus obtain a good representation for the steering matrix.

The angle parameters are found at the maximum of the likelihood function. Typically, a multidimensional parameter search for the various direction of arrivals is required to find the maximum of this function. Various techniques can be used to solve this problem. For the sonar multipath problem these angles can be approximated from the geometry. This gives a good initial estimate that can lead to quick and accurate convergence of the search technique.

Metric Space Representation

The likelihood function can be viewed as an inner product. Equation (7) can be seen to be the inner product of $D^{\dagger} y$ with its conjugate. Likewise, (10) can also be seen as the inner product of the vector $D^{\dagger} y$ with its conjugate. This equation also contains the term $(D^{\dagger}D)^{-1}$ as part of the inner product. This term is important and is known as an inner product metric. This mathematical form indicates that the space that the $D^{\dagger} y$ vectors exist in is a metric space.

A. Inner Product Metric

It is worthwhile to investigate the properties of this metric for the two-signal and n-element case. The metric can be constructed from the steering matrix.

$$D^{\dagger} D = \begin{pmatrix} 1 & e^{-i\phi_1} \cdots & e^{-i(n-1)\phi_1} \\ 1 & e^{-i\phi_2} \cdots & e^{-i(n-1)\phi_2} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ e^{i\phi_1} & e^{i\phi_2} \\ \vdots & \vdots \\ e^{i(n-1)\phi_1} & e^{i(n-1)\phi_2} \end{pmatrix}$$

$$= \begin{pmatrix} n & \dfrac{1 - e^{in\Delta}}{1 - e^{i\Delta}} \\ \dfrac{1 - e^{-in\Delta}}{1 - e^{-i\Delta}} & n \end{pmatrix}$$

where $\Delta = \phi_2 - \phi_1$. The inverse is then

$$(D^\dagger D)^{-1} = \frac{1}{n^2 - \frac{(1-\cos(n\Delta))}{(1-\cos(\Delta))}} \begin{pmatrix} n & -\frac{1-e^{in\Delta}}{1-e^{i\Delta}} \\ -\frac{1-e^{-in\Delta}}{1-e^{-i\Delta}} & n \end{pmatrix}. \quad (11)$$

This metric can be seen to be Hermitian. This is always the case since the likelihood function is real valued.

B. *Off-Diagonal Terms*

The off-diagonal terms in (11) are particularly interesting. These terms are a measure of the coupling between the signals. As the angle between the two signals becomes small, the off-diagonal terms approach the magnitude of the diagonal terms.

$$\lim_{\Delta \to 0} \frac{1-e^{in\Delta}}{1-e^{i\Delta}} = n$$

Since the off-diagonal terms can grow to be nearly as large as the diagonal terms, they clearly cannot be simply ignored.

It should be noted that there are special cases when the off-diagonal terms can be safely ignored. These special cases occur when the off-diagonal terms are zero, which occur when

$$n\Delta = \pm 2\pi, \pm 4\pi, \cdots .$$

Because the off-diagonal terms are zero at these values, the problem naturally decouples and the signals can be completely separated. This is the goal of null steering, which is generally accomplished by weighting the steering matrix to produce nulls in the beam pattern in the direction of the unwanted signal. There are limits and compromises to null steering, especially when signals are spatially close and there are only a small number of array elements.

It should also be noted that when n is large, the inner product metric tends to become diagonally dominant. In this case the off-diagonal terms tend to be relatively less important, and the conventional signal processing approach starts to assume some validity.

C. *Number of Resolvable Signals*

The inner product metric, $(D^\dagger D)^{-1}$, gives a bound on the maximum number of signals, m, that can be spatially resolved by an array of n sensors. Essentially, if $D^\dagger D$ can be inverted, then the metric exists and a solution can be found. If $D^\dagger D$ cannot be inverted, then no solution can be found. Therefore, the requirement is for $D^\dagger D$ to be full rank. This may be determined from the following equation.

$$rank(D^\dagger D) = \min(rank(D^\dagger), rank(D))$$

Since $D^\dagger D$ has dimension m x m, it needs to be rank m in order to be inverted. By extension both D and $D^\dagger$ need to be rank m. This can be seen to be generally true if the number of array elements, n, is equal to or greater than m, the number of signals. Otherwise, if n is less than m, then the rank of D and $D^\dagger$ will be n and $D^\dagger D$ will not be full rank.

This result is not in agreement with eigenvector methods such as MUSIC. These techniques use the sample covariance matrix eigenvectors to determine direction-of-arrival estimates. Since at least one eigenvector is assigned to span the noise subspace, the maximum number of signals that can be resolved with MUSIC is only n-1 [2], [3].

NARROWBAND BEAMFORMER

Beamforming generally involves steering the array towards the signal of interest. For the narrowband case this steering is often done by introducing the phase shifts in (3) to the measured signals. These phase shifts effectively shift the signals in time so that all the array elements receive the signal from the direction of interest at the same time.

For multiple sources the steering matrix, D, is the fundamentally important object. The maximum likelihood beamformer (9) can be expressed as

$$\begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_m \end{pmatrix} = (D^\dagger D)^{-1} D^\dagger \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

This representation clearly shows that all the signals are decoupled.

An example of this approach for two narrowband signals is shown in Fig. 2. The first narrowband signal of length 100 time samples and 10 cycles arrives at a two-element array with half-wavelength separation at an angle of +3° at time sample 50. A similar second signal arrives at the array at an angle of −3° at time sample 100. Most readers will recognize that simple beam steering cannot separate these signals since the main beam is ±30° for a two-element array. However, using the maximum likelihood beamformer, which explicitly calculates the off-diagonal interference terms, the two signals can be completely decoupled or separated.

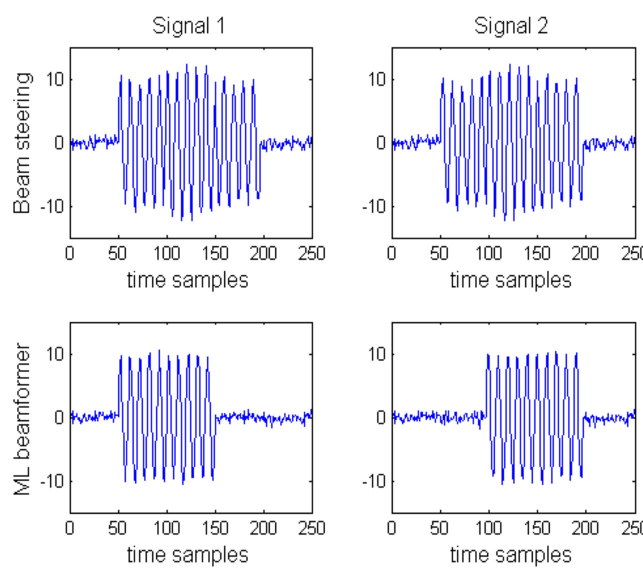

Fig. 2. Output from simple beam steering and maximum likelihood beamformer with 2-element array. Two overlapping narrowband signals arrive at ±3° on a two-element array.

Multipath data was collected from a synthetic aperture rail system at the NSWC test pool. A short 20-kHz narrow beam signal was projected and two wide beam elements received the signal. Fig. 3 shows the data using a traditional beamformer steered toward a spherical target on the bottom. The vertical axis corresponds to the ping number along the rail length, and the horizontal axis corresponds to the relative time sample.

The target is clearly seen between time samples 1100 and 1500. The multipath signal can also be clearly seen between time samples 2500 and 2900. This multipath signal reflects off the water-air boundary and arrives at an angle that is within the array's main beam.

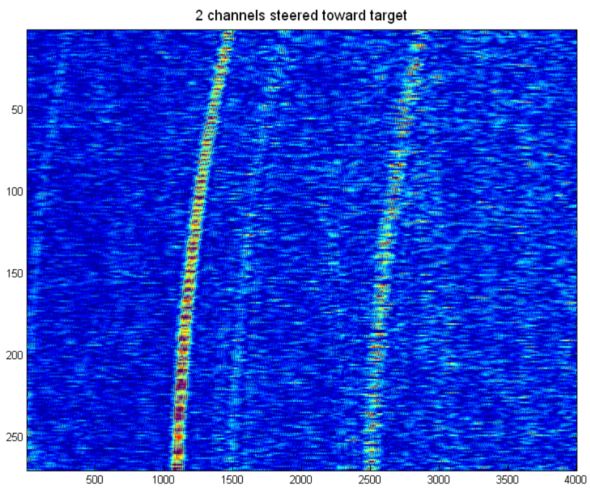

Fig. 3. Direct and multipath signals seen using simple beam steering.

The maximum likelihood beamformer output is shown in Fig. 4. This is the beamformer output that is steered toward the target, while the second beamformer output is steered toward the multipath signal.

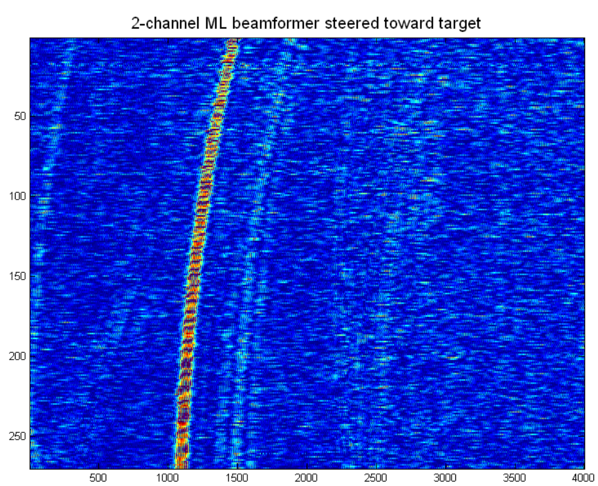

Fig. 4. Only direct path seen using maximum likelihood beamformer.

Ideally, the beamformer should be adaptive in the sense that as the range increases the two angles of interest, (1) angle to the bottom and (2) angle to the multipath signal, change. These angles can be calculated based on the range, depth of the sonar, and depth of the bottom.

BROADBAND BEAMFORMER

Broadband signals require a somewhat different approach. Simple phase shifts work well for narrowband signals since they effectively shift the signals in time. However, the range in frequencies in a broadband signal prevents a simple phase shift approach from working.

The solution is to directly use a time-shift operator. Assuming the signal can be expressed as $e^{if(t)}$, then multiplying the signal by $e^{i\tau}$ results in $e^{i(f(t)+\tau)}$. This yields a general time shift, τ, to the waveform.

The steering vector for a broadband signal with *n* uniformly spaced elements can then be written as

$$D = (\,1 \;\; e^{i\tau} \;\; e^{i2\tau} \;\; \ldots \;\; e^{i(n-1)\tau}\,)^{T}$$

The dimensionless time-shift parameter, τ, is given as

$$\tau = 2\pi \text{ time shift / sample period.}$$

This time shift is analogous to the narrowband case, which is a radian measure of the fractional extra time. The broadband maximum likelihood beamformer can then be developed using the time-shift parameter, τ, instead of the phase angle, ϕ .

An example of this approach for two chirp signals is shown in Fig. 5. The first chirp signal of length 100 time samples and 10 cycles arrives at a two-element array with half-wavelength separation at an angle of +3° at time sample 50. A similar second signal arrives at the array at an angle of −3° at time sample 100. Again, simple beam steering cannot separate these signals. However, the maximum likelihood beamformer easily separates or decouples the two signals.

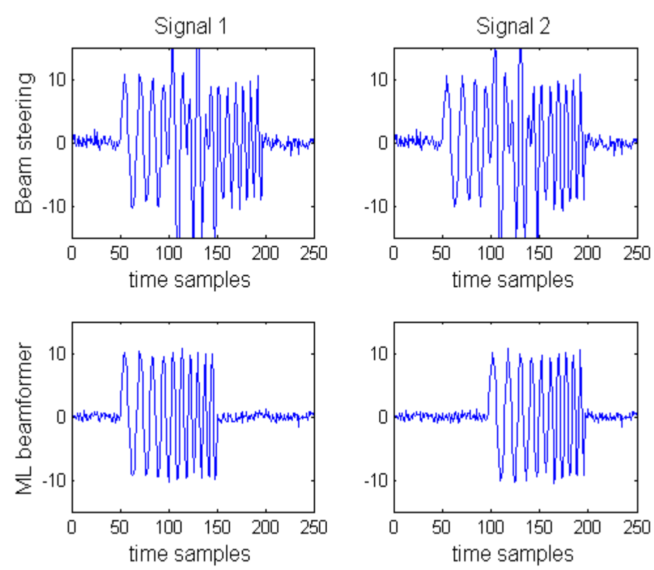

Fig. 5. Output from simple beam steering and maximum likelihood beamformer with 2-element array. Two overlapping chirp signals arrive at ±3° on a two-element array.

COMMENTS

Signal separation with only a small number of array elements is a challenging problem. The maximum likelihood method offers an elegant solution that is able to completely decouple the signals in the spatial domain. Some insights and findings of this approach are given below.

- Although the maximum likelihood method is well known for its high resolution capabilities in determining the angles-of-arrival in a multiple signal environment, this paper has shown that the maximum likelihood method is also capable of completely separating multiple signals. This capability is widely unrecognized but can be easily exploited.
- Signal separation or decoupling occurs naturally in this maximum likelihood approach. This is a direct result of explicitly formulating the problem with multiple signals. This multiple dimensional space leads to the calculation of the off-diagonal terms. These terms represent the real power of this approach.
- It is convenient to consider the likelihood function to be a distance measure that can be represented as an inner product metric space.
- This maximum likelihood approach indicates that a maximum of n signals can be resolved by an n-element array. These signals are resolved in the sense that they can in theory and practice be fully decoupled.
- The steering matrix is not constrained to a uniform linear array geometry for this approach to be valid.
- This maximum likelihood technique to separate signals can be used in other signal processing applications, such as radar and communications.

ACKNOWLEDGMENT

This work was supported in part by the In-Laboratory Independent Research Program funded by the Office of Naval Research.

REFERENCES

[1] S.J. Orfanidis, *Optimal Signal Processing*, New York, NY: Macmillan, p 369, 1988.

[2] Y. Bresler and A. Macovski, “On the number of signals resolvable by a uniform linear array,” *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 34, pp. 1361-1375, December 1986.

[3] B. Friedlander and A.J. Weiss, “On the number of signals whose directions can be estimated by an array,” *IEEE Trans. on Signal Processing*, vol. 39, pp.1686-1689, July 1991.