Wideband Acoustic Array Processing to Detect and Track Ground Vehicles

Tien Pham and Brian M. Sadler
U.S. Army Research Laboratory
Adelphi, MD 20783-1197
phone: 301-394-4282
email: tpham@arl.mil

Abstract

In this study, we apply incoherent and coherent wideband approaches with high-resolution signal subspace algorithms such as MUSIC to estimate the direction of arrival (DOA) of acoustic sources (e.g., ground vehicles). We contrast incoherent and coherent wideband MUSIC techniques and compare them to narrowband MUSIC. We present experimental results for a small baseline circular array illustrating performance and tradeoffs. Performance gain using wideband methods is evident in terms of accuracy and stability, but at higher computational cost than the narrowband approaches.

Introduction

In this paper we contrast and compare incoherent and coherent wideband array processing techniques for acoustic detection and tracking of ground vehicles. We apply these wideband techniques with MUSIC algorithm. Both wideband MUSIC algorithms show significant increase in direction of arrival (DOA) accuracy over the narrowband MUSIC algorithm. We present experimental results for a 12-ft-diameter circular array consisting of 8 sensors, illustrating complexity and performance tradeoffs between incoherent and coherent wideband processing.

In our application, we are motivated to use high-resolution methods because acoustic array baselines are physically constrained by system requirements [1]. The problem is made difficult by several characteristics of the propagation medium and the acoustic sources. Acoustic signatures of ground vehicles are generally nonstationary and undergo severe fading. The usable channel for most instances is largely restricted to [20, 200] Hz, because of wind noise at low frequencies and poor propagation at higher frequencies. There may also be significant time-varying multipath effects. The combined effects of source, terrain, and propagation medium produce large signal variability, even at relatively close ranges.

A typical spectrogram of a moving tracked vehicle with a turbine engine is shown in figure 1. The vehicle exhibits a harmonic structure from the track slaps, but the structure is very nonstationary and displays strong fades during vehicle maneuvering. The turbine engine vehicles exhibit broadband energy from less than 20 Hz to beyond 2 kHz. Because of atmospheric absorption, the energy at the higher frequencies attenuates more rapidly at longer ranges. For this particular test run, note the lack of acoustic energy beyond 250 Hz, except near the closest point approach, [150, 220] s. Beyond 300 s, the signal fades as it moves away from the sensors. However, most ground vehicles of interest have diesel engines that exhibit a more pronounced harmonic structure, because of the engine rate and, to a lesser extent, the track slaps; they do not exhibit the broadband energy that the diesel engine vehicles exhibit. Elsewhere [2] we give detailed analysis of test runs involving tracked vehicles with diesel engines.

Incoherent and Coherent Wideband Processing

A natural extension of narrowband methods is to combine narrowband beampatterns over many temporal frequencies [3]. This is necessary to preserve the narrowband assumption of MUSIC or other high-resolution methods. This approach is useful for acoustic array processing, if there is sufficient SNR in multiple frequency bins, so that narrowband methods (e.g., MUSIC) yield good results independently for each bin. In addition to the relatively high narrowband SNR requirement, disadvantages of this incoherent approach include degradation in the presence of correlated multipath, as well as a general lack of statis-

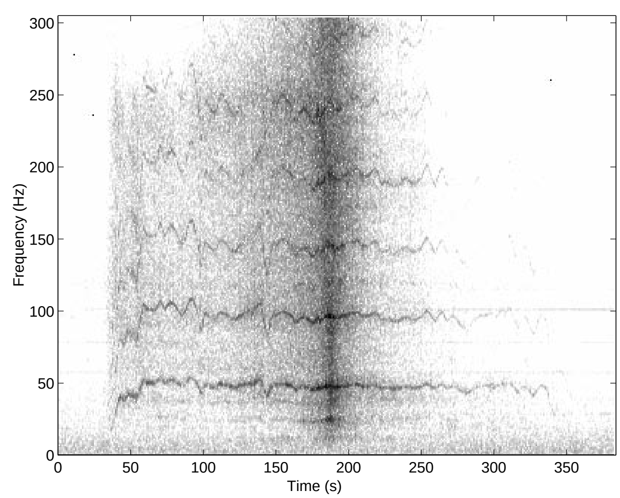

Figure 1. Spectrogram of a turbine engine ground vehicle.

tical stability when compared to wideband coherent methods. Incoherent averaging can lead to false peaks (i.e., false detections) in the resulting averaged beampattern.

Wideband coherent processing gain is possible with the steered covariance matrix (STCM) method originated by Wang and Kaveh [4,5]. STCM is based on forming the composite covariance matrix given by

$$\hat{R}(\theta) = \sum_{m=1}^{M} T(\omega_m, \theta)\hat{R}_Y(\omega_m)T(\omega_m, \theta)^H \qquad (1)$$

where M is the number of narrowband frequency bins, and $\hat{R}_Y(\omega_m)$ is the estimated spatial correlation matrix at frequency ω_m. The steering or focusing matrix, $T(\omega_m, \theta)$, is a function of both frequency and look direction or angle, θ. Here it is defined as

$$T(\omega_m, \theta) = \mathrm{diag}\{[e^{2\pi f_m \Delta t_1}, e^{2\pi f_{mk} \Delta t_2}, ..., e^{2\pi f_m \Delta t_N}]\}, \qquad (2)$$

where $\Delta t_i = (d/c)\sin\phi_i$, $\phi_i = \theta - \alpha_i$; α_i is the angle relative to the normal for sensor i, $i = 1, 2, ..., N$; d is the radius of the circular array; and c is the speed of sound in air. The resulting STCM, $\hat{R}(\theta)$, focuses a signal in the respective narrowband correlation matrices into the same subspace, yielding coherent processing gain over multiple frequencies. Conventional subspace methods (such as MUSIC) can then be applied to $\hat{R}(\theta)$.

The complexity of the coherent STCM approach is increased by the need for computing $\hat{R}(\theta)$ and performing eigenanalysis for every θ. However, the computational load can be lowered by using preliminary estimates of the source locations, obtained, for example, by conventional beamforming. The relative computational complexity between the coherent and incoherent techniques depends on the relative size of the number of look directions versus the number of narrowband frequency bins over which wideband processing occurs. A lower complexity alternative to STCM for linear arrays employs spatial resampling [5].

Implementation

In this section we describe our processing schemes using MUSIC as the means of computing the beampattern. To overcome the nonstationary nature of the source, the data are segmented before processing into fixed data blocks of 1 s, and stationarity is assumed over each data block. The primary steps are (i) compute the fast Fourier transforms (FFTs) for all channels over the data block; (ii) use block-adaptive preprocessing to adaptively select the narrowband frequency bins; (iii) apply incoherent or coherent techniques with MUSIC; and (iv) estimate the directions of the sources from the resulting beampatterns [2]. In practice, the DOA estimates are fed into a tracker that is reasonably robust and, therefore, able to fill in missing or to remove outlying data.

Let $y_i(n)$ denote the output of the ith sensor from an array of N sensors, and let $Y_i(k)$ denote the discrete Fourier transform of $\{y_i(n)\}$. The sum $(1/N)\Sigma_i|Y_i(k)|^2$ is formed in order to adaptively select frequency bins of interest. We simply select the M highest-power bins within the range ω_{low} to ω_{high} for the subsequent analysis.

The incoherent approach proceeds as follows. The conventional narrowband MUSIC beampattern is computed M times. For each θ, we compute

$$\hat{P}_{Incoh}(\theta) = \sum_{m=1}^{M} [E(\omega_m, \theta)^H \Pi^{\perp}(\omega_m) E(\omega_m, \theta)]^{-1}, \quad (3)$$

where $E(\omega_m, \theta) = \mathrm{diag}\{T_s(\omega_m, \theta)\}$ is the steering vector, and the noise orthonormal projector is defined as

$$\Pi^{\perp}(\omega_m) = \hat{U}_n(\omega_m)\hat{U}_n(\omega_m)^H, \qquad (4)$$

where $\hat{U}_n(\omega_m)$ is the noise subspace estimate at ω_m. We estimate the number of signals using a formulation of the Akaike information criterion (AIC) developed by Wax and Kailath [6]. The computational complexity of (3) is approximately $M[O(N^2) + O(N^3) + S \cdot O(N^2)]$, where M is the number of frequency bins and S is the number of look angles. The dominant term in the bracket is the cubic term, which reflects the singular value decomposition (SVD) calculation to form $\hat{U}_n(\omega_m)$ at each ω_m [2].

The STCM approach requires focusing as a function of look direction. Experimental results shown in the next section are based on computing over 360° in 1° steps. After computation of $\hat{R}(\theta)$ for some angle θ, the SVD of $\hat{R}(\theta)$ yields the unitary noise subspace estimate $\hat{U}_n(\theta)$. Here, we estimate the number of wideband signals, based on the formulation of AIC developed by Hong and Kaveh. This formulation is similar to the narrowband case, except for an extra M factor in the first term of the AIC equation [7]. The coherent wideband MUSIC spatial spectrum is then calculated via

$$\hat{P}_{Coh}(\theta) = \left[L^H \hat{U}_n(\theta)\hat{U}_n(\theta)^H L\right]^{-1}, \qquad (5)$$

where L is an N-element vector of ones. The computational complexity of (5) is approximately $S[M \cdot O(N^2) + O(N^3) + O(N^2)]$. Again, the dominant term in the bracket is the cubic term from the SVD calculation of $\hat{U}_n(\theta)$ for each θ [2].

For both methods, the SVD calculation, which is $O(N^3)$, tends to dominate the complexity comparison. Note that, for the incoherent method, it is $M \cdot O(N^3)$, while it is $S \cdot O(N^3)$ for the coherent method, so that the relative complexity is controlled by the relative size of M and S. To reduce the number of frequency bins, M, harmonic line association techniques can be used to group a set of frequency bins for each source and only applying MUSIC to the largest narrowband frequency for each set. To reduce the number of look angles S, we can use coarse angle estimates obtained (e.g., by conventional beamforming) to narrow the field of view [2,4].

Experimental Results

In this section we present experimental results for DOA estimation of one turbine-engine tracked vehicle traveling on a straight path from approximately 2 km west to 2 km east of the sensor array in an open desert field [8]. The vehicle was equipped with a Global Positioning System (GPS) sensor to provide accurate positioning ground truth. However, GPS ground truth was not available for every update, so we perform analysis only for updates where GPS data were available. Figure 2 shows raw experimental DOA estimates for a single source, for narrowband, incoherent wideband, and coherent wideband MUSIC, versus the GPS angles on a test run of 250 s in length. Mean square error (MSE) and mean absolute error (MAE) results for various sets of M frequency bins are shown in table 1. The M frequency components were chosen based on the highest bin signal to noise ratios (SNRs) in the frequency range of [20, 200] Hz, without a priori information. We assume only one signal for each frequency bin (for incoherent processing) and for each look angle (for coherent processing), so that the signal subspace consists of one eigenvector, with the other $N-1$ eigenvectors forming the noise subspace. This assumption reduces the processing load.

The MSEs and MAEs are calculated with the outliers removed [2]. Outliers can be caused by several factors including fading, wind noise, and acoustic source variations. For the error analysis in table 1, the number of outliers ranges from 20 to 30 out of a total of 250 processing intervals of 1 s each in length, sampling rate of 1 kHz, and 1024-pt FFTs. For $M = 1$, incoherent and coherent wideband MUSIC reduce to the narrowband case. Processing gain is evident for both methods, in that the estimates generally improve with increasing M. For this single source experiment, the coherent approach produced smaller errors in terms of both MSE and MAE, reflecting the generally low SNR in each frequency bin from the broadband source and the statistical stability of the coherent method over the incoherent method. Note that, for M = 50, the coherent method begins to degrade in performance accuracy. This may reflects a bias due to the large bandwidth [5]. The bias introduced in the coherent processing has been ignored, and results in table 1 may partly reflect this fact.

Figure 3 shows the pseudospectrograms (a 3-dimensional plot of the beampattern as a function of time) of incoherent and coherent wideband MUSIC for M = 20. Both methods produce excellent results. However, inspection reveals that the coherent method consistently yields sharper beampatterns.

Table 1. MSE and MAE for wideband processing over *M* frequency.

M	Incoherent MUSIC		Coherent MUSIC	
	MSE	MAE	MSE	MAE
1	6.00	1.86	6.00	1.86
10	2.83	1.33	2.64	1.26
20	2.32	1.21	2.13	1.12
30	1.94	1.11	1.75	1.05
40	1.76	1.05	1.71	1.05
50	1.38	0.96	1.87	1.09

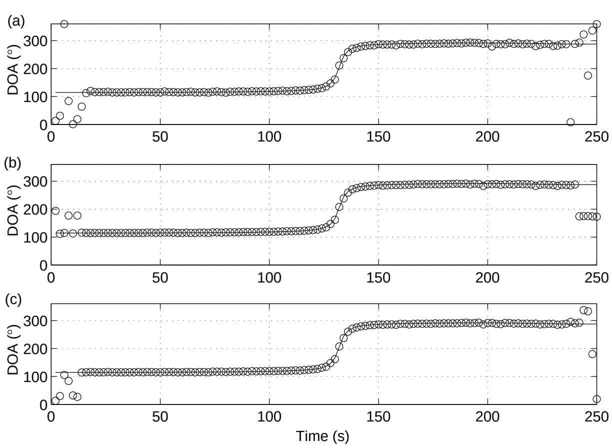

Figure 2. Raw DOA estimates (circles) for (a) narrowband, (b) incoherent wideband (*M* = 20), and (c) coherent wideband (*M* = 20) MUSIC and GPS ground truth (solid line).

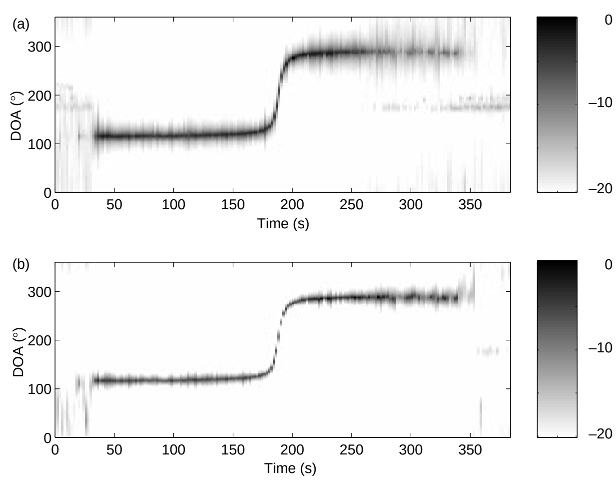

Figure 3. Pseudospectrogram for (a) incoherent and (b) coherent wideband MUSIC with *M* = 20.

Conclusions

Both the incoherent and coherent wideband MUSIC methods provide processing gain over narrowband MUSIC, as shown by our experiments. Given adequate SNR, incoherent wideband methods performed well and yielded sharp and distinct peaks in the beampattern. However, frequency selection is an issue, because the inclusion of low SNR bins tends to degrade the resulting beampattern, reducing source peaks and introducing spurious peaks. In contrast, the coherent MUSIC approach is much more statistically stable, with a beampattern that generally improves with the addition of lower SNR bins. However, as mentioned before, inclusion of more frequency bins can introduce larger bias errors. Here, the coherent method outperforms the incoherent method with sources that have relatively flat spectra. Conversely, for sources with highly peaked spectra the incoherent approach yields better results [2].

The computational complexity comparison between the two methods is largely governed by the SVD calculation, which is $O(N^3)$, with a multiplier given by the number of spectral components M (for incoherent processing) or number of look angles S (for coherent processing).

References

1. T. Pham and B. Sadler, "Aeroacoustic wideband array processing for detection and tracking of ground vehicles," *J. Acoust. Soc. Am*. **98**, No. 5, pt. 2, 2969 (1995).

2. T. Pham and B. Sadler, "Adaptive wideband aeroacoustic wideband array processing," *8th IEEE SP Workshop on Statistical Signal and Array Processing*, 295–298 (June 1996).

3. M. Wax, T. Shan, and T. Kailath, "Spatio-temporal spectral analysis by eigenstructure methods," *IEEE Trans. ASSP*. **32**, 817–827 (August 1994).

4. H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wideband sources," *IEEE Trans. ASSP* **33**, 823–831 (August 1985).

5. J. Krolik, "Focused wideband array processing for spatial spectral estimation," *Advances in Spectrum Analysis and Array Processing*, Vol. II, Ch. 6, S. Haykin, ed., Prentice-Hall (1991).

6. M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. ASSP* **33**, 387–392 (April 1995).

7. H. Wang and M. Kaveh, "On the performance of signal-subspace processing — part II: Coherent wideband systems," *IEEE Trans. ASSP* **35**, 1583–1591 (November 1987).

8. B. Kennedy, R. Olsen, et al., "Joint Acoustic Propagation Experiment: Project Summary," JAPE Catalog number 007 (Illinois Institute of Technology Research internal report) (September 1991).