A FAST DIRECT SOURCE LOCALIZATION APPROACH FOR ACOUSTIC SENSOR ARRAY

Kai Yu\textsuperscript{1,2}, Zhi Wang\textsuperscript{1}\textsuperscript{*}

Zhejiang University
College of Control Science & Engineering\textsuperscript{1}
Hangzhou, 310027 China

Kung Yao\textsuperscript{2}, Ralph E. Hudson\textsuperscript{2}, Charles Taylor\textsuperscript{3}

University of California Los Angeles
Electrical Engineering Department\textsuperscript{2}/
Ecology & Evolutionary Biology Department\textsuperscript{3}
Los Angeles, CA 90095 USA

ABSTRACT

We present a novel Fast Direct Source Localization (FDSL) approach for acoustic sensor array applications. Unlike previous approximate maximum likelihood (AML) approach, the proposed FDSL focuses on the phase shift caused by time delays among sensors and obtains an analytical result without exhaustive search for all the possible locations. Using this phase shift model, conventional phase uncertainty problem can be solved by a light-weight linear search in the limited phase space, which greatly reduces the computational complexity of array processing. Theoretical analysis of the FDSL has been proposed and compared with the Cramer Rao Bound (CRB) of the AML approach. Simulations using real bird call data validate the advantage of proposed method.

Index Terms—Source Localization, Sensor Array, Direction of Arrival (DOA).

1. INTRODUCTION

Source localization using sensor array has been one of the key problems in many applications such as radar, sonar, acoustic tracking etc.\textsuperscript{[1]} \textsuperscript{[2]} \textsuperscript{[3]}. For the past few decades, a wide variety of source localization algorithms have been proposed\textsuperscript{[4]}.

For low-cost embedded systems, reducing the complexity of array processing is the key to near real-time localization updating. For the AML\textsuperscript{[5]} \textsuperscript{[6]} approach, normal 2D exhaustive search needs to cover all the possible locations via many iterations. Similarly, although the result of sparsity signal reconstruction (SSR)\textsuperscript{[7]} can be obtained in one step, the large number of variables (number of grid) in sparse reconstruction require much computational resource. In summary, the reason for exhaustive search and sparse signal reconstruction is that the location parameters are hidden in the complex exponential part of the array data spectrum. Thus the $2\pi$ periodicity makes it difficult to find the unique connection from complex array data spectrum to location parameters (which can be called the ambiguity problem in array processing).

In this paper, we focus on the phase shift among sensor nodes rather than the complex exponential values of phase shifts and find an analytical solution for source location estimation problem. We call this approach a Fast Direct Source Localization (FDSL) estimation. Using the FDSL approach, the source localization is formulated as a matrix calculation. The array geometry information helps to transform the exhaustive search and optimization search to light-weight linear search in limited phase space.

The main contributions are:
1) A novel Fast Direct Source Localization (FDSL) estimation scheme is proposed, which provides analytical solutions with greatly reduced complexity;
2) The error covariance of the FDSL is provided and compared with the Cramer Rao Bound of the traditional AML approach;
3) Simulation using real data validates the feasibility of our approach and shows its superior performance.

This paper is organized as follows. In Section 2, the array signal model and conventional array processing are introduced. Then the problem formulation of the FDSL is explored in Section 3. In section 4, we provide some theoretic analysis of the FDSL approach. In Section 5, we provide some simulation results that show the advantages of the proposed algorithm. Finally, we conclude our work in Section 6.

2. ARRAY SIGNAL MODEL AND PROCESSING

2.1. Array Signal Model

Consider a source localization system with $H$ sensor arrays and each sensor array is equipped with $M$ omni-directional microphones collecting signals from $Q$ sources with distinct directions in the far field. The data received by the $m$th microphone of the $h$th array is given by

$$x_{hm}(t) = \sum_{q=1}^{Q} s_{qh}(t - r_{qhm}(r_q)) + n_{hm}(t), t = 0, \ldots, N-1,$$

(1)

in which $s_{qh}(t)$ is the signal of the $q$th source in the centroid of the $h$th array, $r_{qhm}(r_q)$ is the time delay of the $q$th page.
between the \( m \)th sensor and the centroid of the \( h \)th array, \( r_q \) is the position of the \( q \)th source, \( n_{hm}(t) \) is the zero-mean white Gaussian noise with variance \( \sigma^2 \), and \( N \) is the number of signal samples. For simplification, we denote \( (h,m) \)th as the \( m \)th sensor of the \( h \)th array.

We assume that the position of the \( (h,m) \)th sensor in the Cartesian coordinate system is \( r_{hm} = [x_{hm}, y_{hm}]^T \) and all the \( H \) array have the same array manifold. Then the relative time delay of the \( q \)th source between the \( (h,m) \)th sensor and the array centroid (the origin of the coordinate system) is given by

\[
\tau_{qhm} = \frac{1}{c} r_{hm}^T a_{qh},
\]

where \( c \) is the propagation speed of acoustic signal, \( a_{qh} = (r_q - r_h) / \|r_q - r_h\| \) is defined as the unit vector of the \( q \)th source observed in the \( h \)th array. We further assume that \( s_{qh}(t) \) is a wideband signal. After dividing the time domain signal into multiple frames (each frame has the same size and is separated by the same number of samples between snapshots) and employing the Discrete Fourier Transform (DFT) on each frame, the array data spectrum \( X_h(k) = [X_{h1}(k), X_{h2}(k), \cdots, X_{hM}(k)]^T \) in the \( h \)th array at frequency \( f_k \) can be given by

\[
X_h(k) = D_h(r, k)S_h(k) + N_h(k), k = 0, \ldots, N/2 - 1. \tag{3}
\]

The AML approach calculates the likelihood function \( J(\mathbf{r}) \) over all the possible locations, then returns the most likely location. After ignoring some irrelevant constant terms, the maximum likelihood estimation of the source location is given by

\[
\max_r J(\mathbf{r}) = \min_r \sum_{k=1}^K \sum_{h=1}^H ||X_h(k) - D_h(\mathbf{r}, k)S_h(k)||^2,
\]

where \( K \) is the number of active frequencies that are chosen for array processing.

3. Fast Source Localization

3.1. Problem Formulation

Recall Eq. (2), the phase shift between the \( (h,m) \)th sensor and the array centroid at frequency \( f_k \) is

\[
P_{hm,c}^{(k)} = 2\pi f_k r_{hm} T c^{(k)}.
\]

Then the phase shifts vector \( \mathbf{P}_{hm,c}^{(k)} = [P_{h1,c}^{(k)}, \cdots, P_{hM,c}^{(k)}] \) of the \( h \)th array is given by

\[
\mathbf{P}_{hc}^{(k)} = \frac{2\pi f_k}{c} \mathbf{R}_h a_{h}^{(k)}.
\]

in which \( \mathbf{R}_h = [r_{h1}, r_{h2}, \cdots, r_{hM}]^T \) is the array position matrix. We assume all the \( H \) arrays have the same position (\( \mathbf{R}_1 = \cdots = \mathbf{R}_H = \mathbf{R} \). \( \mathbf{R}_h \) is full-rank and is an isotropic array \[8\] (\( \mathbf{R}^T \mathbf{R} = \kappa I \), \( \sum_{m=1}^M r_m = 0 \)). The unit vector can be obtained by

\[
a_{h}^{(k)} = \frac{c}{2\pi f_k} (\mathbf{R}_h^T \mathbf{R}_h)^{-1} \mathbf{R}_h^T \mathbf{P}_{hc}.
\]

Therefore, the triangulation based source localization can be formulated as:

\[
\mathbf{A}(k) r_s = \mathbf{F}(k),
\]

where \( \mathbf{A}(k) = [a_{11}, a_{21}, \cdots, a_{12}, a_{13}, a_{14}]^T \mathbf{T} \), \( \mathbf{F}(k) = [a_{11} \mathbf{T}_1, a_{12} \mathbf{T}_2, \cdots, a_{14} \mathbf{T}_H] \), \( a_{h} = \sum_{k=1}^K \omega_k a_{h}(k) / \sum_{k=1}^K \omega_k \), \( \omega_k \) is the summation weight of different frequency and is a steering vector that rearranges \( a_{h} \). Then the least square solution is

\[
r_s = (\mathbf{U}^T \mathbf{U})^{-1} \mathbf{U}^T \mathbf{V},
\]

where \( \mathbf{U} = \mathbf{T} \sum_{h=1}^H a_{h} a_{h}^T \mathbf{T}^T \), \( \mathbf{V} = \sum_{h=1}^H (\mathbf{T} a_{h} a_{h}^T \mathbf{T}^T) r_h \).

3.2. Phase Measurement and Uncertainty

The \( 2\pi \) periodicity of \( \exp(-j \omega_k \tau_{qhm}) \) makes it difficult to find the right \( \tau_{qhm} \). To solve this issue, a fast phase search approach is proposed that finds the right phase shift from the array data spectrum. In this paper, only the single source case is considered. However, this FDLS approach also works for multiple sources case when the orthogonal projection technology [5] is introduced.

3.2.1. Phase Measurement

Recall Eq. (5), \( P_{hm,c}^{(k)} \) is defined as the phase difference between the \( (h,m) \)th sensor and the array centroid at frequency \( f_k \). In this section, we only focus on single sensor array, thus the index \( h \) is ignored in all notations. In single source case, the array data spectrum is consistent with the steering vector, and hence the phase shift vector can be easily calculated by

\[
\mathbf{P}_{hc}^{(k)} = [P_{1c}^{(k)}, P_{2c}^{(k)}, \cdots, P_{M_c}^{(k)}]^T,
\]

where \( c \) is the propagation speed of acoustic signal, \( a_{qh} = (r_q - r_h) / \|r_q - r_h\| \) is defined as the unit vector of the \( q \)th source observed in the \( h \)th array. We further assume that \( s_{qh}(t) \) is a wideband signal. After dividing the time domain signal into multiple frames (each frame has the same size and is separated by the same number of samples between snapshots) and employing the Discrete Fourier Transform (DFT) on each frame, the array data spectrum \( X_h(k) = [X_{h1}(k), X_{h2}(k), \cdots, X_{hM}(k)]^T \) in the \( h \)th array at frequency \( f_k \) can be given by

\[
X_h(k) = D_h(r, k)S_h(k) + N_h(k), k = 0, \ldots, N/2 - 1. \tag{3}
\]

The AML approach calculates the likelihood function \( J(\mathbf{r}) \) over all the possible locations, then returns the most likely location. After ignoring some irrelevant constant terms, the maximum likelihood estimation of the source location is given by

\[
\max_r J(\mathbf{r}) = \min_r \sum_{k=1}^K \sum_{h=1}^H ||X_h(k) - D_h(\mathbf{r}, k)S_h(k)||^2,
\]

where \( K \) is the number of active frequencies that are chosen for array processing.
where \( F_{nc}^{(k)} = \angle (X_m(k)X^*_c(k)) \), \( X_c(k) \) is the spectrum at the array centroid, \( \angle(x) = \tan^{-1}(\text{Im}(x)/\text{Re}(x)) \in [-\pi, \pi] \) is the operation that calculates the phase of a complex value.

On condition that the array data spectrum at the array centroid is unknown, \( P_c^{(k)} \) can be obtained by

\[
P_c^{(k)} = \left( I - \frac{1}{2M \times M} \right) P_m^{(k)},
\]

where \( P_m^{(k)} = [P_{1m}(k), P_{2m}(k), \cdots, P_{Mm}(k)] \) is the phase shift vector with respect to the \( m \)th sensor. This is because \( \sum_{m=1}^{M} F_{mc}^{(k)} = 2\pi f_k/c \sum_{m=1}^{M} r_m^2 \mathbf{a}^{(k)} = 0 \).

### 3.2.2. Phase Search Formulation

Fig. 1. (a) Phase search of low frequency, (b) Phase search of high frequency

Fig. 1 shows the uncertainty of phase shift, it is related to the largest distance \( D \) between two sensor. Suppose \( D \) is smaller than the wavelength (as shown in Fig.1.a), then the corresponding phase shift between two sensors should be smaller than \( 2\pi \). This condition is equivalent to \( f_k \leq \frac{c}{D} = f_{\text{Low}} \). Frequencies less than \( f_{\text{Low}} \) have smaller phase shifts than \( 2\pi \).

In this case, the possible alternative phase shifts between the \( n \)th and the \( m \)th sensors are

\[
P_{nm}^{(k)} = \begin{cases} P_{nm}^{(k)} & P_{nm}^{(k)} - 2\pi \text{sign}(P_{nm}^{(k)}) \end{cases}.
\]

Consider each \( P_{nm}^{(k)} \) has 2 alternatives, the total alternative number of \( P_{nm}^{(k)} \) is \( 2^{M-1} \). Actually, the number of \( 2^{M-1} \) can be further reduced. An easy approach is to divide the \( 2^{M-1} \) alternatives into \( M \) cases, in which each sensor is assumed to be nearest to the source and has the smallest phase. So if the \( m \)th sensor has the smallest phase, \( P_{nm}^{(k)}(n \neq m) \) should be positive and hence negative phase shift \( P_{nm}^{(k)}(n \neq m) \) can be removed from the possible phase shift set. After deleting the negative phase shifts, the size of possible phase shift set becomes \( M \). Those are

\[
P_{m}^{(k)} = P_{m}^{(k)} + \pi \text{sign}(P_{m}^{(k)}) \left( \text{sign}(P_{m}^{(k)}) - 1 \right),
\]

\( m = 1, 2, \cdots, M \).

Till now, the possible phase shift has been limited to \( M \) alternatives, and the remaining problem is to find the unique one among them.

Fig.1.b is the high frequency cases of \( f_k \geq f_{\text{Low}} \), the maximum phase shift is \( 2\pi [d_{f_k}/c] \). This means the alternative number of \( P_{nm}^{(k)} \) becomes \( L_k = [d_{f_k}/c] \). In this case, we need to extend the phase search space from \( 2\pi \) to \( 2\pi L_k \), and the total number of possible \( P_{m}^{(k)} \) is \( ML_k^{M-1} \).

### 3.3. Solution for Phase Uncertainty

Given the possible phase shift vectors, the unique \( P_c^{(k)} \) can be selected from all the alternatives. Based on the the unit norm definition of \( a^{(k)} \), the right phase shift should satisfy

\[
(P_c^{(k)})^T P_c^{(k)} = \frac{4\pi^2 f_k^2 k}{c^2}.
\]

Thus the norm constraints of \( P_c^{(k)} \) can be used as a criterion that tests the validity of all the possible phase shifts and finds the right one.

### 4. PERFORMANCE ANALYSIS

To evaluate a source localization method, criteria such as error covariance and CRB are used. In this section, the error covariance of our FDSL is provided and compared with the CRB of the AML approach.

Recall Eq. (9), the error of \( r_s \) can be approximated by

\[
\Delta r_s = \frac{\partial r_s}{\partial a^{(k)}} \Delta a,
\]

where \( \frac{\partial r_s}{\partial a^{(k)}} = \left[ \frac{\partial r_s}{\partial a_1^{(k)}}, \frac{\partial r_s}{\partial a_2^{(k)}}, \cdots, \frac{\partial r_s}{\partial a_M^{(k)}} \right] \), \( \frac{\partial r_s}{\partial a^{(k)}} = \left[ \frac{\partial r_s}{\partial a_1^{(k)}}, \frac{\partial r_s}{\partial a_2^{(k)}}, \cdots, \frac{\partial r_s}{\partial a_M^{(k)}} \right] \),

\[
\Delta a = [\Delta a_1^{(k)}, \Delta a_2^{(k)}, \cdots, \Delta a_M^{(k)}]^T.
\]

Assume the 1st sensor of the \( h \)th array has the smallest phase, the error of \( P_{ch}^{(k)} \) can be given by

\[
\Delta P_{hc}^{(k)} = \left( I - \frac{1}{M} \right) \begin{bmatrix} 0 & 0^T & 0 & 0^T \\ E_1 & E_2 & E_3 & E_4 \end{bmatrix} \begin{bmatrix} \Delta \text{Re}(N_h^{(k)}) \\ \Delta \text{Im}(N_h^{(k)}) \end{bmatrix},
\]

\[
E_1 = [\epsilon_{1,1}^1, \epsilon_{1,1}^2, \cdots, \epsilon_{1,1}^M]^T, \quad E_2 = \text{diag}(\epsilon_{2,1}^1, \cdots, \epsilon_{2,1}^M), \quad E_3 = [\epsilon_{1,2}^1, \epsilon_{2,2}^1, \cdots, \epsilon_{1,2}^M]^T, \quad E_4 = \text{diag}(\epsilon_{4,1}^1, \cdots, \epsilon_{4,1}^M),
\]

\[
\epsilon_{1,1}^m, \epsilon_{1,2}^m, \epsilon_{2,1}^m, \epsilon_{4,1}^m = \frac{\text{Im}(X_m(k))}{|X_m(k)|^2}, \quad \epsilon_{1,1}^m, \epsilon_{1,2}^m, \epsilon_{2,1}^m, \epsilon_{4,1}^m = \frac{\text{Re}(X_m(k))}{|X_m(k)|^2}.
\]

Here \( \epsilon_{1,1}^m, \epsilon_{1,2}^m, \epsilon_{2,1}^m, \epsilon_{4,1}^m \) are the partial derivatives of \( P_{ch}^{(k)} \) to \( \text{Re}(X_1(k)), \text{Im}(X_1(k)), \text{Re}(X_m(k)) \) and \( \text{Im}(X_m(k)) \) respectively.

After some derivations, the covariance of \( \Delta P_{ch}^{(k)} \) and \( \Delta a_h^{(k)} \) are

\[
E(\Delta P_{ch}^{(k)}) = \frac{\sigma_h^2}{2} (I + 1_{M \times M}) , \quad E(\Delta a_h^{(k)}) = \frac{N^2 \sigma^2}{8\pi^2} \frac{1}{f_k^2 c^2 h^2}.
\]
Assume the noise \(N_h(k)\) at different frequency is independent, the error covariance of \(\Delta a_h\) is

\[
\Delta a_h = \sum_{k=1}^{K} w_k \Delta a_h^{(k)} / \sum_{k=1}^{K} w_k.
\]  

where \(\mu_h = \sum_{k=1}^{K} f_k^2 S_{h,k}^2\). Recall Eq. (9), the partial derivation of \(r_s\) is given by

\[
\frac{\partial r_s}{\partial a_h} = -U^{-1} \frac{\partial U}{\partial a_h} V + U^{-1} \frac{\partial V}{\partial a_h} U^{-1} = -U^{-1} z_h (\Delta r_h),
\]

in which \(z_{h1} = T (a_h e_1^T + e_1 a_h^T) T^T\), \(e_1 = [1,0]^T\). Combine Eqs. (15-20), the covariance of \(\Delta r_s\) is

\[
\Delta r_s \Delta r_s^T = -U^{-1} \sum_{h=1}^{H} \frac{\partial r_s}{\partial a_h} \Delta a_h \Delta a_h^T \frac{\partial r_s}{\partial a_h} U^{-1} = \frac{N c^2 \sigma^2}{8 \pi^2} U^{-1} \sum_{h=1}^{H} \alpha_h U^{-1},
\]

in which \(\alpha_h = z_h \Delta r_h \Delta r_h^T z_h^T\). After some simple calculations, the error covariance of \(r_s\) is

\[
\text{COV}(r_s) = \frac{N c^2 \sigma^2}{8 \pi^2} U^{-1} \sum_{h=1}^{H} \Delta r_h \Delta r_h^T T U^{-1}.
\]

For comparison, the CRB [9] of source localization is

\[
\text{CRB}(r) = \frac{N \sigma^2 \sigma^2}{8 \pi^2} \left( \sum_{k=1}^{K} \frac{1}{f_k} \sum_{h=1}^{H} S_{h,k}^2 \Delta r_h \Delta r_h^T T \right)^{-1}.
\]

5. SIMULATION RESULTS

In this section, simulations are carried out to validate our FDSL approach. Four circular arrays with diameter 7.5cm are implemented at \([-120, -120]m, [-120, 120]m, [120, -120]m, [120, 120]m\), and a real bird call of Bewick’s Wren (BEWR) is used as source signal in simulations.

Fig. 2 shows the comparison between the CRB of AML approach and the COV of the proposed FDSL approach. The \(SNR=10dB\) and 20 snapshots are used for each source localization estimation. The COV of the FDSL approach is 37% higher than the CRB of the AML approach on average, but it is still within an acceptable range.

Further simulations are conducted to compare the CRB and COV values with the RMSEs of the AML and the FDSL approaches. Both methods approach their theoretical performance limits with increasing SNR. Although the FDSL has 37% performance degradation with respect to the AML, the processing time of the FDSL is only 3% of the AML approach. Consider both the computational complexity and accuracy, the FDSL approach shows its superiority compared to the AML approach.

6. CONCLUSION

In this paper, a novel FDSL estimation approach is proposed. The source localization problem is considered from an alternative perspective of phase calculation. Thus, traditional exhaustive search based AML approach can be replaced by a light weight linear search problem. Also the array geometry is used to solve the phase uncertainty problem. Simulation using experiment data validates the feasibility of the FDSL approach and shows its superiority over the AML approach.
7. REFERENCES


