Applied Acoustics 136 (2018) 41–47

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Fast direction-of-arrival estimation algorithm for multiple wideband acoustic sources using multiple open spherical arrays

Xi Pan, Huayang Wang*, Zhiyi Lou, Yongchao Su

School of Mechatronical Engineering, Beijing Institute of Technology, Beijing 100081, China

ARTICLE INFO

Keywords:
Direction of arrival(DOA)
Multiple wideband acoustic sources
Orthonormal propagator method (OPM)
Spherical harmonic (SH)
Spherical microphone arrays (SMAs)

ABSTRACT

The orthonormal propagator method (OPM) is extended to the spherical harmonic (SH) domain using multiple open spherical microphone arrays (SMAs). Compared with the multiple signal classification method (MUSIC), the computational loads of the OPM can be lower (by one or two orders), because the OPM does not require eigendecomposition of the cross spectral matrix (CSM) of the received sensor signals. Moreover, multiple open SMAs have wider frequency ranges of operation because their large radii reduce their lower frequency bounds and their small radii increase their higher frequency bounds. In this work, the performances of the OPM using both the spherical Fourier transform (SFT) components in the SH domain (SH-SFT-OPM) and the cross spectral matrix of the SFT components in the SH domain (SH-CSM-OPM) are analyzed theoretically and then evaluated in terms of their root-mean-square errors under various signal-to-noise ratio (SNR) conditions and direction-of-arrival estimation snapshots. The performance results are then compared with those of the SH-MUSIC. The simulation results confirm that the SH-CSM-OPM performs similarly to the SH-MUSIC, even for low SNRs, and that the computational complexity of the SH-SFT-OPM is significantly lower than that of the SH-MUSIC.

1. Introduction

Spherical microphone arrays (SMAs) have been used in various applications, including sound-field reproduction [1–4], beamforming [5–7], and direction-of-arrival (DOA) estimation of sound sources [8–12]. Subspace-based methods in the spherical harmonic (SH) domain, a DOA estimation method using SMAs, and the division of observation space into a noise subspace and a source subspace have all demonstrated high resolution capabilities in a three-dimensional space [8,10,11]. Subspace-based methods in the SH domain all use the eigendecomposition of the cross-spectral matrix (CSM) of the SH coefficients, which is obtained using the spherical Fourier transform (SFT). One of the most famous of these high-resolution methods—SH multiple signal classification (SH-MUSIC)—has shown significant improvements over MUSIC alone [13]. However, obtaining the eigenvectors is computationally expensive when the number of microphones is large. The orthonormal propagator method (OPM) is a subspace-based method that does not require the eigendecomposition of the CSM of the received sensor signals [14], and the propagator is a linear operator that depends only on the steering vectors. Thus, the OPM can significantly reduce the computational complexity.

There are two main methods for the DOA estimation of wideband signals. One is the incoherent signal subspace method (ISSM), and the other is the coherent signal subspace method (CSSM). The key idea of the ISSM is to decompose wideband signals into narrowband sectors and apply narrowband techniques to each sector independently. After then, the results from all of the frequency sectors are averaged [15]. This method, though simple, suffers from performance degradation when the signal-to-noise ratio (SNR) varies at each frequency, and it requires a DOA estimation for each frequency sector, which results in a high computational cost. The CSSM transforms the CSMs of many frequency sectors into one general CSM at a single focusing frequency using a focusing matrix that depends on the frequency sector [16]. The drawbacks of CSSM are that these focusing matrices require initial DOA values and that the estimation performance is very sensitive to these initial values. However, the SFT using SMAs has the advantage that the frequency-dependent components are decoupled from the angular-dependent components [10]. As a result, the sound source DOA is no longer dominated by the frequency of the source. By removing the frequency-dependent components, the DOA estimation methods can be applied to narrowband signals without requiring either focusing matrices or knowledge of the initial DOA values. Thus, by using the SFT together with the frequency removal techniques, both a better DOA estimation performance and a lower computational complexity can be achieved.

For open spheres, the modal strength contains poor numerical

* Corresponding author at: Beijing Institute of Technology, 3rd Building 5, South Zhongguancun Street Haidian District, Beijing 100081, China.
E-mail address: wanghuayang@bit.edu.cn (H. Wang).

https://doi.org/10.1016/j.apacoust.2018.01.014
Received 16 July 2017; Received in revised form 8 November 2017; Accepted 14 January 2018

conditioning at the fixed frequencies [5], which causes excessive noise at these frequencies. To overcome this problem, a method using multiple open SMAs (MOSMAs) is proposed. By choosing an appropriate radius ratio for a MOSMA, not only can numerically robust open SMAs be obtained, but the frequency range of operation can be increased as well, which means that a smaller radius increases the upper frequency range bound and a larger radius increases the lower frequency range bound.

To perform a fast DOA estimation for multiple wideband acoustic sources, the proposed method incorporates a wider frequency band of operation with the faster abilities of the OPM method in the SH domain by using MOSMAs. The remainder of this paper is organized as follows. In Section 2, the MOSMA model is presented. Section 3, the proposed DOA algorithm, using both the SH-SFT-OPM and the SH-CSM-OPM for multiple wideband acoustic sources, is derived, and its computational complexity is analyzed. In Section 4.1, the performance of the MOSMA is discussed, and the frequency ranges of operation for various dual open sphere radii are presented for the given number of microphones. In Section Section 4.2, the performances of the SH-SFT-OPM, the SH-CSM-OPM, and the SH-MUSIC are analyzed theoretically and simulated in terms of their RMSEs for both various SNRs and various DOA estimation snapshots. Finally, Section Section 5 concludes the paper.

Throughout this paper, vectors are denoted by lower-case bold letters, and matrices are represented by upper-case bold letters. The superscripts $_{\mathrm{T}}$, $_{*}$, and $_{\mathrm{H}}$ denote transpose, complex conjugate, and conjugate-transpose, respectively.

2. Multiple open SMA model

To simplify both the description and the physical construction of the algorithm, we first restrict the problem to that of a dual open SMA, as shown in Fig. 1, composed of two open spheres: one with a radius of r_1 and another with a radius of $r_2 = \rho r_1$, where ρ is the ratio of the radii of the two spheres. Distributed evenly over each of the surfaces of these two spheres are L omnidirectional pressure microphones with positions of $\Omega_l = (\theta_l,\phi_l)$ $(l = 1, 2 \ldots \mathrm{L})$. The position vector of the lth microphone of the dth sphere, in Cartesian coordinates, is $\boldsymbol{r}_{dl} = (r_{dl}\sin\theta_l\cos\phi_l, r_{dl}\sin\theta_l\sin\phi_l, r_{dl}\cos\theta_l)^T$ $(d = 1, 2)$, where r_{dl} is the distance from the origin to that microphone. The azimuth angle ϕ is measured counterclockwise from the x-axis, and the elevation angle θ is measured down from the z-axis.

Consider S plane waves with the wave number k impinging on the array and with $k = 2\pi f/c$, where f is the frequency, c is the speed of sound, and the DOA of the xth plane wave is $\Omega_s = (\theta_s,\phi_s)$ $(s = 1, 2 \ldots \mathrm{S})$. Then, the wavenumber vector of the sth propagating plane wave is denoted by $\boldsymbol{k}_s = -(k_s\sin\theta_s\cos\phi_s, k_s\sin\theta_s\sin\phi_s, k_s\cos\theta_s)^T$. In the spatial domain, the sound pressure at each of the L microphones on the two SMAs can be written as

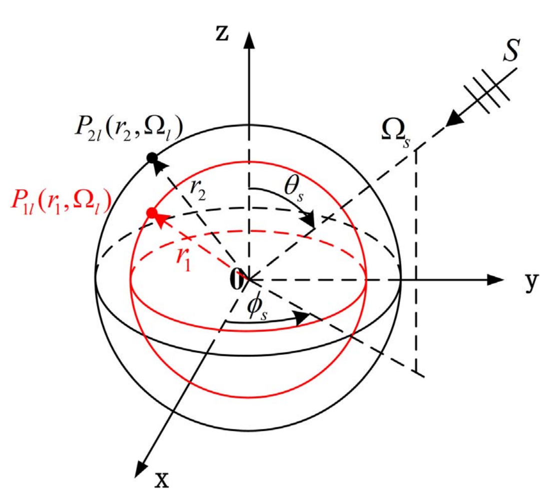

Fig. 1. Illustration of the dual spherical array geometry.

$$\boldsymbol{x}_1(k) = \boldsymbol{A}_1(k)\boldsymbol{s}(k) + \boldsymbol{n}_1(k), \tag{1}$$

$$\boldsymbol{x}_2(k) = \boldsymbol{A}_2(k)\boldsymbol{s}(k) + \boldsymbol{n}_2(k), \tag{2}$$

where $\boldsymbol{x}_d(k) = [x_{d1}(k), x_{d2}(k), \cdots x_{dL}(k)]^T \in C^{L\times 1}$ represents the observation vectors of the two SMAs, respectively; $\boldsymbol{s}(k) = [s_1(k), s_2(k), \cdots s_S(k)]^T \in C^{S\times 1}$ represents the signal waveform vectors, and $\boldsymbol{n}_d(k) = [n_{d1}(k), n_{d2}(k), \cdots, n_{dL}(k)]^T \in C^{L\times 1}$ represents the respective additive noise vectors that are both uncorrelated with the signal vector $\boldsymbol{s}(k)$, $\boldsymbol{A}_d(k) \in C^{L\times S}$ is the steering matrix, and $\boldsymbol{A}_d(k)$ can be written as

$$\boldsymbol{A}_d(k) = \begin{bmatrix} e^{-i\boldsymbol{k}_1^T\boldsymbol{r}_{d1}} & e^{-i\boldsymbol{k}_2^T\boldsymbol{r}_{d1}} & \cdots & e^{-i\boldsymbol{k}_S^T\boldsymbol{r}_{d1}} \\ e^{-i\boldsymbol{k}_1^T\boldsymbol{r}_{d2}} & e^{-i\boldsymbol{k}_2^T\boldsymbol{r}_{d2}} & \cdots & e^{-i\boldsymbol{k}_S^T\boldsymbol{r}_{d2}} \\ \vdots & \vdots & & \vdots \\ e^{-i\boldsymbol{k}_1^T\boldsymbol{r}_{dL}} & e^{-i\boldsymbol{k}_2^T\boldsymbol{r}_{dL}} & \cdots & e^{-i\boldsymbol{k}_S^T\boldsymbol{r}_{dL}} \end{bmatrix} = [\boldsymbol{a}_{d1}(k) \;\; \boldsymbol{a}_{d2}(k) \;\; \cdots \;\; \boldsymbol{a}_{dS}(k)] \tag{3}$$

where $\boldsymbol{a}_{ds}(k) \in C^{L\times 1}$ is the steering vector, with

$$\boldsymbol{a}_{ds}(k) = \left[e^{-i\boldsymbol{k}_s^T\boldsymbol{r}_{d1}}, e^{-i\boldsymbol{k}_s^T\boldsymbol{r}_{d2}}, \ldots, e^{-i\boldsymbol{k}_s^T\boldsymbol{r}_{dL}}\right]^T, \tag{4}$$

Here, represents the pressure of the sth unit plane wave that is incident from the direction Ω_s on the lth microphone of the dth SMA. In the SH domain, this can be expressed as [17]

$$e^{-i\boldsymbol{k}_s^T\boldsymbol{r}_{dl}} \approx \sum_{n=0}^{N}\sum_{m=-n}^{n} b_n(kr_d)Y_{nm}^*(\Omega_s)Y_{nm}(\Omega_l), \tag{5}$$

where $i = \sqrt{-1}$; n is the order, ranging from 0 to N; N is the highest SH order, which is determined by the sampling set Ω_l of the microphones on the sphere [1]; m is the degree, ranging from $-n$ to n; and $b_n(kr_d) = 4\pi i^n \times j_n(kr_d)$ is the far-field modal strength for open SMAs, where $j_n(kr_d)$ is the spherical Bessel function. Fig. 2(a) shows the magnitude of b_n as a function of f with $r_d = 0.1$ m for several orders n. Y_{nm} is the SH of both the order n and degree m, defined as

$$Y_{nm}(\Omega) = \sqrt{\frac{2n+1}{4\pi}\frac{(n-|m|)!}{(n+|m|)!}}P_{nm}(\cos\theta)e^{im\phi}, \tag{6}$$

where $P_{nm}(\cos\theta)$ is the associated Legendre function. According to (5), the steering matrix $\boldsymbol{A}_d(k)$ can be expressed as

$$\boldsymbol{A}_d(k) = \boldsymbol{Y}(\Omega_L)\boldsymbol{B}(kr_d)\boldsymbol{Y}^H(\Omega_S), \tag{7}$$

where $\boldsymbol{Y}(\Omega_L) \in C^{L\times(N+1)^2}$ is an SH matrix given by

$$\boldsymbol{Y}(\Omega_L) = \begin{bmatrix} Y_{00}(\Omega_1) & Y_{1-1}(\Omega_1) & Y_{10}(\Omega_1) & Y_{11}(\Omega_1) & \cdots & Y_{NN}(\Omega_1) \\ Y_{00}(\Omega_2) & Y_{1-1}(\Omega_2) & Y_{10}(\Omega_2) & Y_{11}(\Omega_2) & \cdots & Y_{NN}(\Omega_2) \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ Y_{00}(\Omega_L) & Y_{1-1}(\Omega_L) & Y_{10}(\Omega_L) & Y_{11}(\Omega_L) & \cdots & Y_{NN}(\Omega_L) \end{bmatrix}, \tag{8}$$

$\boldsymbol{B}(kr_d) \in C^{(N+1)^2\times(N+1)^2}$ is a far-field modal strength diagonal matrix written as

$$\boldsymbol{B}(kr_d) = diag(b_0(kr_d), b_1(kr_d), b_1(kr_d), \cdots, b_N(kr_d)), \tag{9}$$

and $\boldsymbol{Y}(\Omega_S) \in C^{S\times(N+1)^2}$ is an SH matrix containing the source signal directions, which can be expressed as

$$\boldsymbol{Y}(\Omega_S) = \begin{bmatrix} Y_{00}(\Omega_1) & Y_{1-1}(\Omega_1) & Y_{10}(\Omega_1) & Y_{11}(\Omega_1) & \cdots & Y_{NN}(\Omega_1) \\ Y_{00}(\Omega_2) & Y_{1-1}(\Omega_2) & Y_{10}(\Omega_2) & Y_{11}(\Omega_2) & \cdots & Y_{NN}(\Omega_2) \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ Y_{00}(\Omega_S) & Y_{1-1}(\Omega_S) & Y_{10}(\Omega_S) & Y_{11}(\Omega_S) & \cdots & Y_{NN}(\Omega_S) \end{bmatrix}, \tag{10}$$

Substituting (7) into both (1) and (2) and multiplying each of the results from the left by $\boldsymbol{Y}^H(\Omega_L) \times \boldsymbol{V}$ provides the SFT coefficient model of the observation signal, which is given as follows: