Acoustic source localization using the open spherical microphone array in the low-frequency range

Boquan Yang³, Shengguo Shi^{1,2,3}, and Desen Yang^{1,2,3}

¹Acoustic Science and Technology Laboratory, Harbin Engineering University, Harbin 150001, China

²Key Laboratory of Marine Information Acquisition and Security (Harbin Engineering University), Ministry of Industry and Information Technology; Harbin 150001, China

³College of Underwater Acoustic Engineering, Harbin Engineering University, Harbin 150001, China

Abstract. Recently, spherical microphone arrays (SMA) have become increasingly significant for source localization and identification in three dimension due to its spherical symmetry. However, conventional Spherical Harmonic Beamforming (SHB) based on SMA has limitations, such as poor resolution and high side-lobe levels in image maps. To overcome these limitations, this paper employs the iterative generalized inverse beamforming algorithm with a virtual extrapolated open spherical microphone array. The sidelobes can be suppressed and the main-lobe can be narrowed by introducing the two iteration processes into the generalized inverse beamforming (GIB) algorithm. The instability caused by uncertainties in actual measurements, such as measurement noise and configuration problems in the process of GIB, can be minimized by iteratively redefining the form of regularization matrix and the corresponding GIB localization results. In addition, the poor performance of microphone arrays in the low-frequency range due to the array aperture can be improved by using a virtual extrapolated open spherical array (EA), which has a larger array aperture. The virtual array is obtained by a kind of data preprocessing method through the regularization matrix algorithm. Both results from simulations and experiments show the feasibility and accuracy of the method.

1 Introduction

Over the last two decades, spherical microphone arrays have been variously applied to localize and identify aeroacoustics source in 3D sound field [1,2,9] due to the directivity of the whole space and the flexibility of 3D beam pattern synthesis [3-5]. Various beamforming algorithms can be combined with the spherical microphone arrays for solving localization and identification. As a basic approach, conventional beamforming is more simple and robust in far-field [4,6] but it is restricted by the poor spatial resolution and the high sidelobe levels at low frequency [2,3-6].

To improve the spatial resolution and to suppress the disturbing sidelobes, Suzuki proposed the generalized inverse beamforming (GIB) [7,12,13], achieving higher computational efficiency and localization accuracy compared with conventional beamforming approaches. However, there are two main limitations of GIB. Firstly, the uncertainties in actual measurements, such as measurement noise and sensors configuration, will lead to instability and errors in the GIB [10,12,13]. Secondly, the GIB algorithm is also not effective in the low-frequency range because of physical and practical limitations, such as the number of microphone sensors and the aperture of the array in the acquisition system.

In order to solve the problem mentioned above, this paper utilizes the double iteration generalized inverse beamforming (DI-GIB) and a novel kind of spherical array extrapolation method. The inverse problem can be stabilized by iterating the beamforming regularization matrix initially and the localization accuracy can be further improved by the next iteration [14]. A virtual extrapolated array (EA) with a larger array aperture can be obtained according to the received signal of the open spherical microphone array without increasing the array element, which will strengthen the performance of the beamformer at low frequency.

This paper is organized as follows. In Sec. 2 a very brief overview of the theory is given, regarding spherical microphone arrays algorithm and array extrapolated method based on the inverse problem regularization theory. Simulations illustrate the performance of the proposed method in Sec. 3. Section 4 is devoted to validate the correctness of the simulation conclusions and effectiveness of these methods in experiments. Conclusions are drawn in Sec. 5.

2 Theory

2.1. Array extrapolated using open spherical arrays

* Corresponding author: shishengguo@hrbeu.edu.cn

In this paper, the inverse problem theory is adapted to sound field extrapolation around a microphone array for adjacent spatial sound. To improve the performance of the algorithm at low frequency, this paper uses an extrapolated spherical array with a larger radius, shown in Fig.1. This method supposed an envelope between the sound source and the actual spherical array. And there are some virtual sources on the envelope surface S . These virtual sound sources are generated by the actual sound source radiation to the envelope surface S [9].

Fig. 1. The geometry of array extrapolated.

The Helmholtz equation is [8]

$$\nabla^2 p_r(a) + k^2 p_r(a) = \delta(a - r') \quad (1)$$

For sound sources confined in a source volume S , the resulting continuous sound field is given by the simple source formulation.

$$p(a) = \int_S G(a, r') q(r') dS(r') \quad (2)$$

where $G(a, r')$ denotes the free field green function.

In the actual numerical calculation, suppose there are K virtual sources distributed evenly on the sphere S , so the formula needs to be discretized as [7]

$$\mathbf{p}(\mathbf{a}_1) = \mathbf{G}(\mathbf{a}_1, \mathbf{r}'_k) \mathbf{q}(\mathbf{r}'_k) \quad (3)$$

where $\mathbf{G}(\mathbf{a}_1, \mathbf{r}'_k)$ denotes the transfer matrix between secondary sound sources and array elements. Through the formula (3), the inverse question can be presented in different forms

$$\mathbf{q} = \mathbf{G}^{-1} \mathbf{p} \quad (4)$$

Under the actual circumstance, the number of virtual sources used in the inverse problem is always larger than or equal to the number of measurement microphone $K \leq L$. It is an ill-posed question for $K > L$. Therefore, the solution of the inverse problem needs to use a regularization method. It is also possible to regularize the inverse using Tikhonov regularization

$$\mathbf{q}_{reg} = \arg \min \{ \|\mathbf{p} - \mathbf{G}\mathbf{q}\|_2^2 + \lambda^2 \Omega(\mathbf{q})^2 \} \quad (5)$$

where λ stands for the regularization parameter, $\|\cdot\|_2$ is 2-norm, $\Omega(\mathbf{q}) = \|\mathbf{L}\mathbf{q}\|_2$ which is a smoothing norm of $\mathbf{q}$, where $\mathbf{L}$ is the penalty matrix. When matrix $\mathbf{L}$ is the identity matrix, the problem corresponds to the basic Tikhonov regularization method. But when $\mathbf{L} \neq \mathbf{I}_n$, the solution of the functional problem is [12]

$$\mathbf{q}_{reg} = \frac{\mathbf{G}^H \mathbf{p}}{\mathbf{G}^H \mathbf{G} + \lambda^2 \mathbf{L}^H \mathbf{L}} \quad (6)$$

where $[\cdot]^H$ denotes the complex conjugate transpose.

In order to avoid the lack of information caused by the penalty matrix, the beamforming regularization matrix is adopted

$$\mathbf{L} = [\text{diag}(\|\mathbf{G}^H \mathbf{p}\| / \|\mathbf{G}^H \mathbf{p}\|_\infty)^{-1}] \quad (7)$$

The formula for solving the strength of the virtual sources is

$$\mathbf{q}_{BF} = \frac{\mathbf{G}^H \mathbf{p}}{\mathbf{G}^H \mathbf{G} + \lambda^2 [\text{diag}(\|\mathbf{G}^H \mathbf{p}\| / \|\mathbf{G}^H \mathbf{p}\|_\infty)^{-1}]^2} \quad (8)$$

Substituting the virtual source strength into Eq. (3), it can get the sound pressure approximation in the spherical surface S_0 . Using this method, we are able to get an extrapolated spherical array. Assuming that the virtual array has M elements with radius r shown in Fig.1, the received signal of m th element is

$$p_v(r_m) = \mathbf{G}(r_m, r') \mathbf{q}_{BF}(r') \quad (9)$$

The received sound pressure vector of the virtual array is

$$\mathbf{P}_v = [p_v(r_1), p_v(r_2), \dots, p_v(r_M)]^T \quad (10)$$

2.2 Iterative generalized inverse beamforming method

On the basis of the inverse problem theory mentioned above, the Tikhonov regularization solution of source strength is

$$\mathbf{q}_s = \arg \min \{ \|\mathbf{P}_v - \mathbf{G}\mathbf{q}_s\|_2^2 + \lambda^2 \Omega(\mathbf{q}_s)^2 \} \quad (11)$$

where $\mathbf{G}$ denotes the transfer matrix between sound sources and the virtual array elements. In this paper, it is assumed that the number of microphones is fewer than that of grid points of source scanning plane. An optimal inverse solution for Eq. (11) found by solving this least-mean-square problem is [11,13,14]

$$\mathbf{q}_{\lambda\beta} = \mathbf{L}^{-1} \mathbf{G}^H \mathbf{J}_\lambda \beta \mathbf{P}_v \quad (12)$$

where $\mathbf{G} = \mathbf{G}\mathbf{L}^{-1}$, $\mathbf{J}_\lambda = (\mathbf{G}\mathbf{G}^H + \lambda \mathbf{I})^{-1}$, β is the scaling parameter.

$$\beta = \|\mathbf{G}\mathbf{L}^{-1}(\mathbf{G}\mathbf{L}^{-1})^H + \lambda \mathbf{I}\|_2 \quad (13)$$

The proposed algorithm (EA-DI-GIB) can be realized through substituting Eq (10) into the Generalized inverse beamforming and iterate the regularization parameters and the beamforming output. The iterative process is as follows [13].

The first iteration:

The regularization matrix $\mathbf{L}$ can be redefined by the solution of generalized inverse beamforming [13].

$$\mathbf{L}^{(1)} = \text{diag}(\|\mathbf{W}^H \mathbf{P}_v\| / \|\mathbf{W}^H \mathbf{P}_v\|_\infty) \quad (14)$$

$$\mathbf{q}^{(1)} = \mathbf{L}^{-1(1)} \mathbf{G}^H \mathbf{J}_\lambda \beta \mathbf{P}_v \quad (15)$$

$$\mathbf{L}^{(i+1)} = \text{diag}(\|\mathbf{q}^{(i)}\| / \|\mathbf{q}^{(i)}\|_\infty) \quad (16)$$

$$\mathbf{q}^{(i+1)} = \mathbf{L}^{-1(i+1)} \mathbf{G}^H \mathbf{J}_\lambda \beta \mathbf{P}_v \quad (17)$$

The second iteration:

The regularization matrix $\mathbf{L}$ can be redefined by the cross-spectral form of the final solution of the first step. The cross-spectrum of the output of the first iteration process is taken as the initial value of the second iteration process [13].

$$\mathbf{Q}^{(1)} = \mathbf{q}^{(n+1)} (\mathbf{q}^{(n+1)})^H \quad (18)$$

$$\mathbf{L}^{-1(j+1)} = \text{diag} \left(\left\| \mathbf{Q}^{(j)} \right\| / \left\| \mathbf{Q}^{(j)} \right\|_{\infty} \right) \quad (19)$$

$$\mathbf{q}^{(j+1)} = \mathbf{L}^{-1(j+1)} \mathbf{G}^H \mathbf{J}_{\lambda} \beta \mathbf{P}_v \quad (20)$$

The proposed method is substituting the virtual array received sound pressure Eq.(10) into the iterative process above mentioned.

3 Simulation examples

In this section, simulations are conducted using a 64-element open spherical array that the elements are randomly distributed evenly on the spherical surface with a radius of $a=0.3\text{m}$, whose geometrical setup is in Fig.2. The distance between the spherical array center and the source is 1 m, the sound speed in air is 340m/s. The focus surface is assumed to be spherical, where its radius equal to 1m and the focus surface are divided to 37×73 points with an interval $[\Delta\theta=5^\circ, \Delta\phi=5^\circ]$. GIB, EA-GIB and EA-DI-GIB methods are used to focus the sound pressure signals sampled by microphones to the surface of the interest, and then map the acoustic sources.

Fig. 2. Randomly distributed spherical array.

3.1 Single source

This simulation compares the performance of the GIB, EA-GIB, and EA-DI-GIB methods for locating the single

single source. The point source is located at $(1\text{m}, 89^\circ, 160^\circ)$. Fig.3 and 4 respectively show the acoustic source localization images with the signal frequency $f=180\text{Hz}$ ($ka=1$) and 540Hz ($ka=3$). For EA-GIB and EA-DI-GIB methods, the radius of the virtual spherical array is set to 0.45m in Fig.3, and similarly, the virtual spherical array radius is equal to 0.4m in Fig.4. In Fig.3 and 4, the GIB method can locate the single acoustic source, and it has lower sidelobes and narrower main lobe as the signal frequency increases. The EA-GIB method has a narrower main lobe compared to the GIB method, but its sidelobes raise when the signal frequency is increased. As a contrast, The EA-DI-GIB method outperforms the other method in terms of the sidelobes and spatial resolution by combining the array extrapolated technique and EA-DI-GIB. Consequently, the EA-DI-GIB method is more suitable for locating acoustic sources.

3.2 Double source

This simulation examines the performance of the above methods for identifying two point sources, where two sources respectively are located at $(1\text{m}, 94^\circ, 78^\circ)$ and $(1\text{m}, 94^\circ, 259^\circ)$. The other simulation conditions are the same as in section 3.1. Fig.5 and Fig.6 respectively show the localization results with signal frequency $f=250\text{Hz}$ and 500Hz . Three methods are used for source localization: GIB, EA-GIB and EA-DI-GIB. In Fig.5, the EA-GIB method has lower sidelobes and narrower main lobes by exploiting the extrapolated array technique. But the EA-GIB method cannot effectively improve the localization results, since the advantage of the array extrapolated is not obvious for the higher signal frequency. As a contrast, the EA-DI-GIB method not only sharpens the mainlobe but also removes the sidelobe contaminations effectively. Moreover, the EA-DI-GIB method has better localization performance for the higher signal frequency. Thus, the EA-DI-GIB is a superior method for locating acoustic source.

Fig. 3. Acoustic source simulation results of a single loudspeaker source at 180Hz after different post-processing techniques: (a)GIB, (b)EA-GIB, (c)EA-DI-GIB.

Fig. 4. Acoustic source simulation results of a single loudspeaker source at 540Hz after different post-processing techniques: (a)GIB, (b)EA-GIB, (c)EA-DI-GIB.

Fig. 5. Acoustic source simulation results of double point sources at 250Hz after different post-processing techniques: (a) GIB, (b) EA-GIB, (c) EA-DI-GIB.

Fig. 6. Acoustic source simulation results of double point sources at 500Hz after different post-processing techniques: (a) GIB, (b) EA-GIB, (c) EA-DI-GIB.

4 Validation examples

With the purpose of validating the correctness of simulation conclusions and the effectiveness of the proposed method in practical applications, experimental measurements are performed using an open spherical array and small loudspeakers in a big room. The configuration is depicted in Fig.7. The open spherical array consists of 64 microphones and its radius is equal to 0.3m. Taking the center of the spherical as the origin of the coordinate system, two loudspeakers are located at $(1\text{m}, 94^\circ, 78^\circ)$ and $(1\text{m}, 94^\circ, 259^\circ)$, respectively. Sound pressure signals received by microphones are acquired

by Brüel&Kjær PULSE Data Acquisition System. The sampling frequency is 65536 Hz.

Fig. 7. Picture of the experiment setup.

Fig. 8. Acoustic source localization experiment results of a single loudspeaker source at 180Hz ($ka=1$) after different post-processing techniques: (a) GIB, (b) EA-GIB, (c) EA-DI-GIB.

Fig. 9. Acoustic source localization experiment results of a single loudspeaker source at 540Hz ($ka=3$) after different post-processing techniques: (a) GIB, (b) EA-GIB, (c) EA-DI-GIB.

Fig. 10. Acoustic source localization experiment results of double loudspeakers source at 250Hz after different post-processing techniques: (a) GIB, (b) EA-GIB, (c) EA-DI-GIB.

Fig. 11. Acoustic source localization experiment results of double loudspeakers source at 500Hz after different post-processing techniques: (a) GIB, (b) EA-GIB, (c) EA-DI-GIB.

Fig.8 and Fig.9 show the localization results of the single loudspeaker excited by the mono-frequency at 180Hz and 540Hz, respectively. It can be seen from Fig.8 and 9 that these methods can correctly locate the position of the single source. Comparing Fig.8(b) and 8(a), the EA-GIB method improves the localization performance of the GIB method by extrapolating the array. However, the EA-GIB method cannot obviously improve the localization performance of the GIB method at the high signal frequency by comparing Fig.9(a) and 9(b). It is illustrated that the array extrapolated technique is effective for the low signal frequency. As a contrast, the EA-DI-GIB method outperforms the other methods in terms of spatial resolution and sidelobe suppression. From the above, the experimental conclusions are in line with the simulation ones, and they have proved that the conclusions are correct and the proposed method is effective for the practical single source.

Fig.10 and 11 show the localization results of the double loudspeaker synchronously excited by the mono-frequency at 250Hz and 500Hz, respectively. In Fig. 10, the GIB method is unable to identify two sources, and the EA-GIB method can locate the positions of two sources by exploiting the extrapolated array technique. But the sidelobes of the EA-GIB method are still high. As a contrast, the EA-DI-GIB method can suppress the sidelobes and sharpen the main lobe by combining the extrapolated array technique and DI-GIB. In Fig.11, the above methods can correctly locate the positions of two sources. Moreover, the location performance of EA-DI-GIB method is better when the signal frequency increases.

5 Conclusions

A double iteration GIB algorithm has been combined with the spherical microphone array to improve the localization accuracy. The inverse problem has been

stabilized by iterating the beamforming regularization matrix initially. A higher localization accuracy can be achieved through the second iteration. The virtually spherical array with a big array aperture is utilized to strengthen the performance of the beamformer in the low-frequency range in this paper. Both simulation and experimental results showed that the proposed method can obtain accurate localization results with high spatial resolution and narrower main-lobe especially in the low frequency range, which demonstrates the potential of the method for localizing acoustic sources in actual 3D environment.

References

1. Sloan I H, Womersley R S, *Advances in Computational Mathematics*, **21**(1-2):107-125, (2004)
2. Petersen S O, Odense: University of Southern Denmark, 2004
3. Agmon M, Rafaely B, *Electrical and Electronics Engineers in Israel, 25th Convention of IEEE*, IEEE, 2008, 227-230
4. Yan S, Sun H, Svensson U P, et al. Audio, Speech, and Language Processing, *IEEE Transactions on*, **19**(2): 361-371(2011)
5. B. Rafaely, *Journal of the Acoustical Society of America* **116** (4) 2149-2157 (2004)
6. B. Rafaely, Y. Peled, M. Agmon, D. Khaykin, E. Fisher, SpringerVerlag, Germany, 281-305 (2010)
7. P.A. Gauthier, C. Camier, Y. Pasco, A. Berry, E. Chambatte, R. Lapointe, M.A. Delalay, *Journal of Sound and Vibration*, **330** (24) 5852-5877 (2011)
8. Williams E G, *Fourier acoustics: sound radiation and nearfield acoustical holography*, NewYork: Academic press, 1999

9. D. B. Ward, T. D. Abhayapala, *ICASSP*, **2**, 1781-1784 (2002)
10. S. Li, Z. Xu, Z. Zhang, Z. Zhang, J. Mao, et al., *J. Vib. Control* 1-12 (2015)
11. Y. Kim, P. A. Nelson, *J. Sound Vib.* **275**, 463-487 (2004)
12. T. Padois, P. A. Gauthier, A. Berry, *J. Sound Vib.* **333**, 6858-6868(2014)
13. T. Suzuki, *Proceedings of the 14th AIAA/CEAS Aeroacoustics Conference*, Vancouver, May, 2008
14. Z. Zhang, S. Chen, Z. Xu, Y. He, S. Li, *J. Sound Vib.* **396**, 108-121 (2017)