2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI 2017)

Improving the Accuracy of Beamforming Method for Moving Acoustic Source Localization in Far-field

Qiqi Bao, Feng Luan*, Jingxuan Yang
School of Computer Science and Engineering
Northeastern University
Shenyang, China

Abstract—**Microphone arrays are widely used in the applications for acoustic source localization. The interest in signal processing techniques for deriving directional information from small-sized arrays of microphones in far field model is steadily growing. And beam forming is a robust method for source localization, which aims at estimating the source position by maximizing the response output by the microphone arrays in the source direction. In this paper, we focus on the sound source orientation in far field modal. In order to estimate the probability of the sound source point in the plane that we have assumed, the plane is divided into grids with the same size and then we use autocorrelation method to evaluate the possibility degree of each grid and after being normalized we use Matlab to show the distribution of the possibility to orient the actual position. And in the experiment, the moving noise sound, which gives out continuous sound and keeps changing the position, is used to detect the accuracy of the algorithm. Furthermore, the use of interpolation method and autocorrelation matrix in beamforming estimation can overcome the orientation error caused by the data limit and compute at a corresponding fast speed.**

Keywords-acoustic source; beamforming; far-field; moving source

I. Introduction

Nowadays the application of far-field microphone arrays is becoming increasingly popular in the context of acoustic testing for sound source localization [1-3]. Both time-domain and frequency-domain methods used for analyzing the system are widely applied in the digital signal processing area. And time delay estimation is the primary measuring procedure in source localization or beamforming applications, as a result, microphone array is widely used in many fields including radar, sonar, geophysics, communication and medical imaging and so on [14].

In this paper, with the aim to track the moving sound, particularly the noise, beamforming algorithm based on far-field time domain model is established [11]. The delay-and-sum beamforming algorithm computes a beamforming map of sound pressure contribution from the array microphone output signals [4-7][13]. In order to estimate the probability distribution of the possible sound in the chosen plane, we first divide the scope of the target region into small rectangular grids with the same size and use their center coordinate to represent the approximate position of each grid in the algorithm. Since the position of each grid is known, we can use the geometry method to calculate the difference of the distance between each microphone and the central position, and calculate the time delay between those two microphones. Then the moving time or the moving unit grid number which is calculated by rounding down the time delay multiplied by the sampling frequency is used to shift the sound data gathered by the microphones to evaluate the correlation of them by using autocorrelation algorithm, for if the grid point is near the actual sound source point, it can be easily seen that the possible degree will be higher than those far from the actual sound position. Through this operation, we aim at analyzing the degree of the possibility of each grid to be the actual sound source point and show up the range of the possible occurrence. However, since the value, which we calculated by using the time delay multiplied by the sampling frequency is not an integer, the rough calculation can be somewhat inaccuracy, as a result, we then use the interpolation algorithm to improve the accuracy of the moving grids with more smaller unit value which can get a more accurate result [9].

By choosing the center coordinate as the grid's location, the microphone arrays can first judge the approximate areas and then do the accurate inspection, using the method of geometry, time shift and autocorrelation matrix which is capable to gain an accurate probability value through the algorithm [8][10].

In Section 2, we explain the geometry method to calculate the time delay of each grids and the theory of beam forming method. In Section 3, we present the simulated examples with the method of autocorrelation and interpolation. And then we discuss the results and do a comparison between those two methods. In Section 4, we apply the technique in the real experiment with the moving noise.

The main purpose of this paper is to develop an algorithm based on time domain method and locates the coherent sound with beamforming.

II. Methodology

A. Estimate the Time Delay

The brief sketch of the microphone array is placed in a circle and the eight microphones are allocated equally in the plane and the measurement range also can be seen in Fig.1. First, setup with a planar microphone array with 8 microphones and a 2-D region of interest, which is divided into rectangular homogeneously. And the origin of the coordinate system is placed in the center of the planar microphone arrays. The acoustic sources are located in an x-y plane at a distance of z_0 from the center of the microphone array. The camera angle for

*Corresponding author: luanfeng@mail.neu.edu.cn

978-1-5386-1937-7/17/$31.00 ©2017 IEEE

landscape orientation angle is θ_1, while the portrait orientation angle is θ_2. And we assume that the sound source is only positioned in this plane.

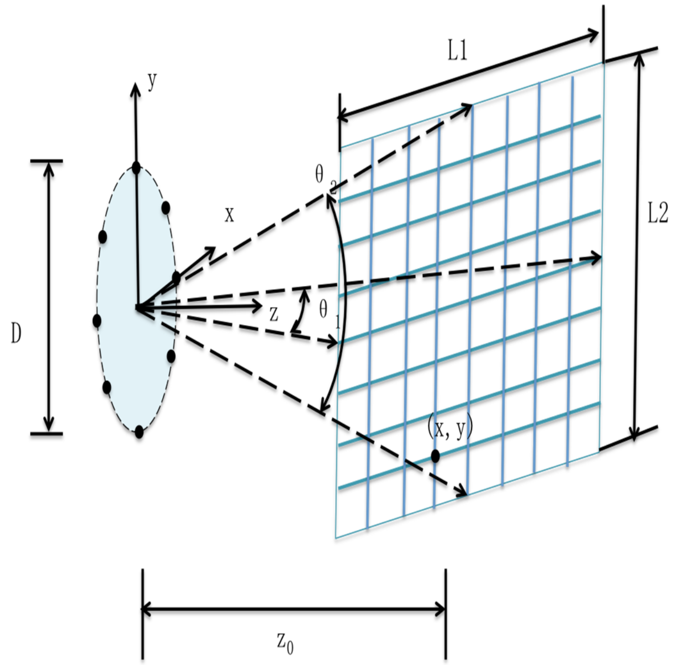

Figure1. A brief sketch of a planar microphone array placed in circle and a 2-D region of interest.

Fig.2 shows the far-field model, where the radius of the microphones arrays are placed relatively compact, so that the source data gathered by each microphone can be seen from the same angle and direction. The red point in the picture is thought as the original sound source and its spreading direction to each microphone and the central position is thought to be parallel. In the far field model, the direction of the sound source can be described by two angles, which are azimuth angle and altitude angle. With such data information, the time shift can be solved [12][14].

Through geometric methods, by doing the a perpendicular line from the sound source to the microphone plane, the angle of the sound source can be written as

$$K=-[\sin(\theta)\cos(\varphi); \sin(\theta)\sin(\varphi); \cos(\theta)] \tag{1}$$

In equation (1), K is a vector of the original sound source angle, θ is the elevation angle of the acoustic source and φ equals to $\tan^{-1}(\frac{y}{x})$, where (x, y) represents the coordinate of the acoustic source.

Then, we use far-field modal, based on its plane incident wave, the equation of the time-delay deviation and the incident angle is

$$\tau\_m=r\cdot K/v \tag{2}$$

In equation (2), τ_m denotes the time delay, r is the direction vector of the microphone, which has three vectors, and v is the speed of the sound. Note that τ_m should be round to the integer for further calculation.

Then shift the time delay of τ_m with the data that the microphones captured in matrix.

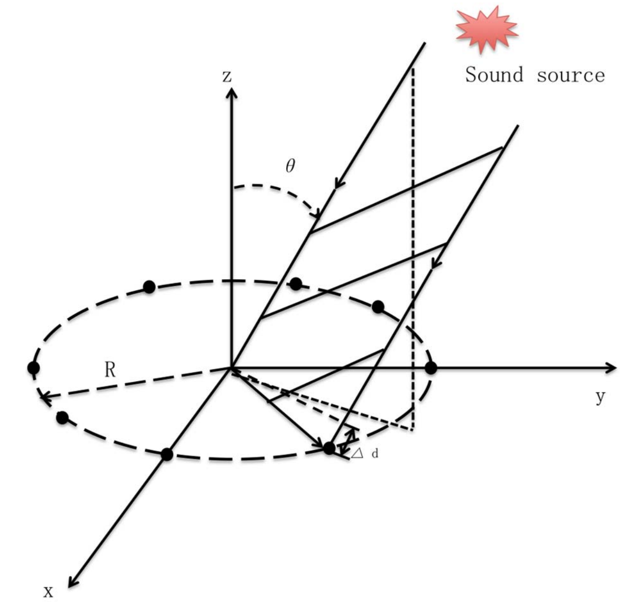

Figure2. Far-field modal with plane incident wave from the sound source.

B. Use beamforming method to compute grid

The beamforming method can estimate the orientation of the sound source. Since the sound of each source travels to every microphone along different paths, the signals captured by the microphones are similar in waveform, but show different delays, which are proportional to the travel distances.

So after choosing a focus point in one of the grids, shift the signal captured by each microphone based on the corresponding runtime difference, the possible orientation of the sound source can be compute by comparing the relative sound pressure strength shown in Fig.3. Note that testing point1 and testing point2 are seen as the test points. Then the signal components of testing point1 in all channels are in phase, whereas the signal components originating from testing point2 are out of phase. If the signals of all channels are summed together, the amplitude of the signal component of testing point1 in the sum signal is strong and the signal components originating from testing point2 are negligible. And after shifting and summing, the normalization should be done to get the final results.

In order to simplified the calculation and get the probability of acoustic sources on each point, divide the source plane into grid G with N equidistant points and each point is seen as a test point. If the real acoustic source is located in the grid point, then after shifting the waveform with the length of the time-delay calculated before and then does the normalization, the acoustic sources can be estimated and get corresponding probability value.

Since the limitation of the sample frequency, the use of autocorrelation matrix to solve the main propagation source direction can reduce the estimation error.

$$M=P^{H}P \tag{3}$$

The alphabet shown in (3), both P and M are matrixes. P denotes the data of the sound source after shifting the sample data by the estimated time-delay while M denotes the data after finishing doing the autocorrelation.

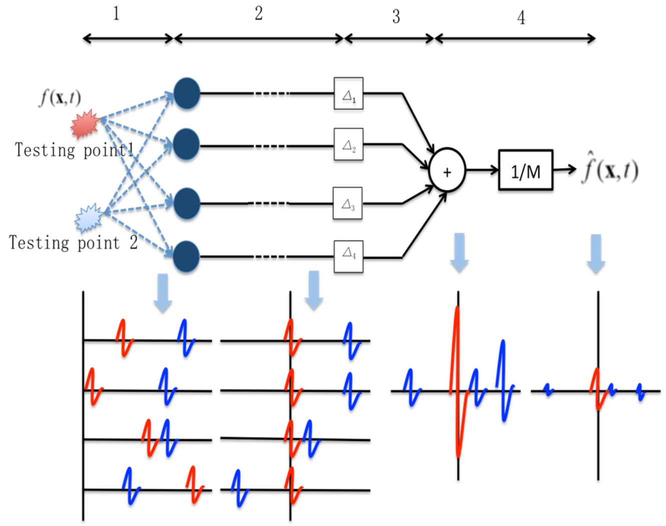

Figure3. Block diagram beam forming.

III. SIMULATIONS

The simulations are carried out with Matlab and the plane array contains 8 microphones are positioned in a circle, whose diameter is R and the intersection angle between the microphones is 45°.

$$R=\frac{1}{7\times\sin\left(\frac{2\times\pi}{M1}\right)}=0.2\text{m} \tag{4}$$

In equation (4), M1 represents the number of the microphones, which equals to 8. In the geometrical setup, the observation plane is paralleled to the array plane. The distance between array plane and observation plane z_0 equals to 1.0 m. The opening angle is assumed as 100° and 80°. And the sampling frequency is 20 kHz. In this simulation, we also add noise into the system. Then divide the scanning area into 51×41 grids in the equal size. And two sources are placed simultaneously at the grid point (－15°, 10°) and (30°, 0°) in the plane respectively. And by using geometry and vector analysis in calculation, the delayed time can be stored. After shifting and storing the sampling sound source value of each microphone in an n (n represents the total number of the line)×m (m represents the number of the channel, in this simulation m equals to 8) matrix A, use autocorrelation matrix to calculate the relative possibility of each testing point.

$$B=A^{H}A \tag{5}$$

Equation (5) is the autocorrelation matrix, where $A=[\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8]$, and α_i contains n elements of sound voltage. In theory, if the test point is the actual sound source position, the sound voltage value gathered by each channel should be equal in the shape of the envelope with a bit difference in amplitude, ignoring the interruption of the system's noise first. So the summation of the matrix shows the relative possibility of each testing point. As is shown in the Fig.4, the relative possibility of the sound can be seen, where the red color represents the maximum probability and the following color of yellow, green, blue shows that the possibility decrease progressively.

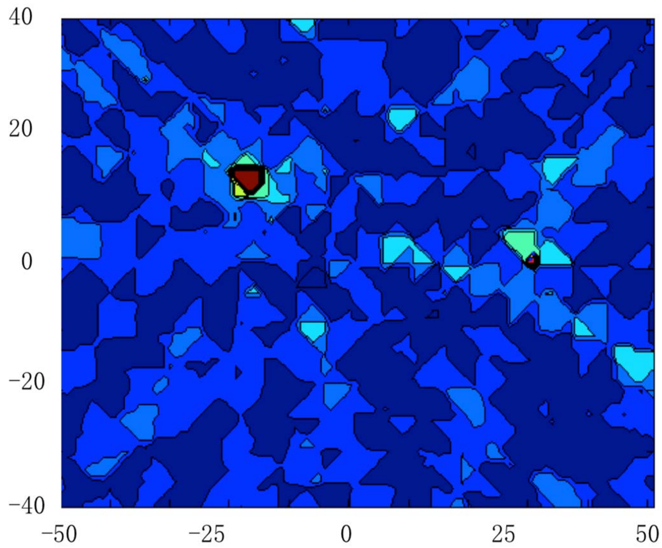

Figure4. Contour plots with autocorrelation matrix algorithm showing simulations of sound sources.

However, most of time delay have decimal part, so rather than simply round them, in our second simulation, interpolation method is used to accurate the result by scaling the sampling value, which is equivalence to expanding the sampling frequency. After such operation, round the time delay and the result can be more accurate. As can be seen in Fig.5, the interpolation method is used, and the size of the red color becomes smaller, thus the orientation is more accurate, for it can orient a more exact position rather than a relative big range of the position.

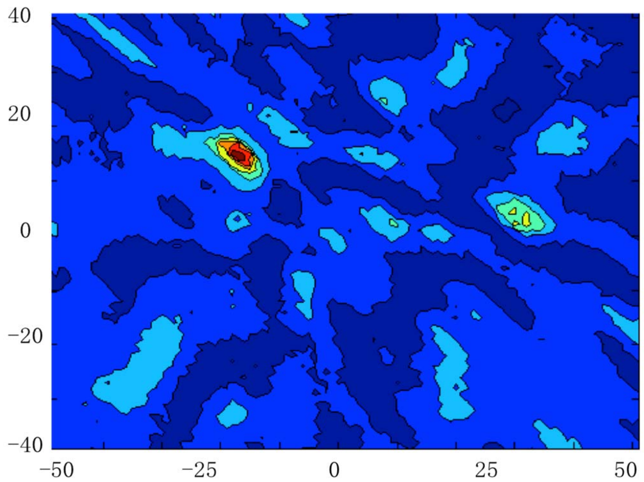

Figure5. Contour plots with interpolation method showing simulations of sound sources.

In order to develop the testing speed and reduce the number of times of multiplication in the algorithm, matrix B in (5) is simplified by the following matrix.

$$B=A^H\cdot\alpha_8 \qquad (6)$$

Here, $\langle\alpha_i, \alpha_8\rangle=\alpha_i^T\cdot\alpha_8$, so each data gathered by the one of the microphones is multiplied only once with the other microphones. Such operation can also manipulate the sound's position in a much quicker speed but with the cost of a bit inaccuracy. Fig.6 shows that though the maximum possibility of the sound can be computed, the distribution of the possibility is not satisfied.

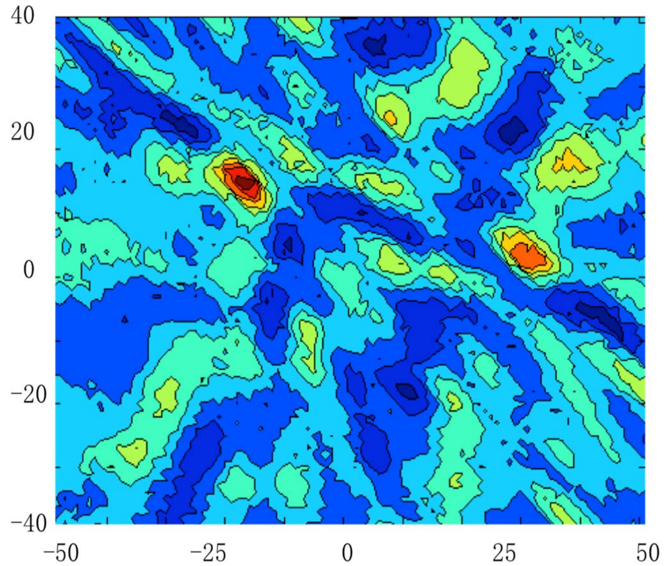

Figure6. Contour plots with the method of using the matrix multiply a vector in the matrix showing simulations of sound sources.

So in order to improve the accuracy of the algorithm mentioned above, the method followed is manipulated. Since eight microphones are used in the simulation, matrix B in (6) is a 8×1 matrix, and the eighth element, which is $\alpha_8^T\cdot\alpha_8$, in B is the multiple of itself. And $\alpha_8^T\cdot\alpha_8$ can be seen as the reference value. Then, abstract this real-number from each of the other element in matrix B, and through this way can we normalize all the value in matrix B with the same reference value and then put the new matrix in B. Additionally, sum the value of each element in the lately formed matrix B and use matrix B as the divisor to get the relative ratio. Such operation can develop the accuracy of the result. As can be seen in the Fig.7, the relative probability distribution is more focused and outstanding than that in Fig.6, where the maximum possibility is more centralized.

After all points are scanned, the possibility value of all the grids can be stored in the matrix and the visualization result can be acquired. And the function of the contour plots in Matlab shows the degree of possibilities simulations of sources at different locations.

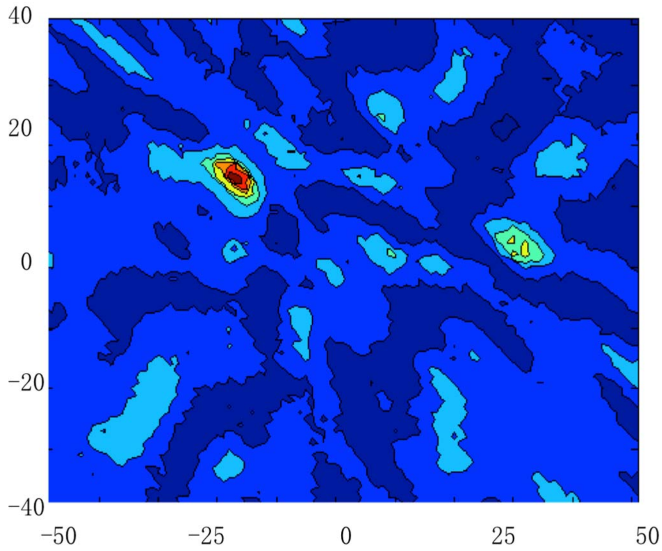

Figure7. Contour plots with the normalization method showing simulations of sound sources.

IV. EXPERIMENT

For the experiment, an array with 32 microphones, which is shown in Fig.8 is used. The red points represent the first layer of microphones. The green points represent the second layer of microphones. The blue points represent the outside layer of microphones. There are 32 microphones in the array, while only the 7 microphones in the first layer (pained in green*) are used or the 6 microphones in the second layer (painted in red*) are used for the far-field modal. The experiment uses a speaker positioned at the distance of 1.6m from the arrays, and the analysis size with the camera angle is 43° in horizon and 31.7° in vertical and the speaker is held and shook by someone as a moving sound source. The sampling frequency of the mechanism is 20000 Hz and sampling time is 15 s.

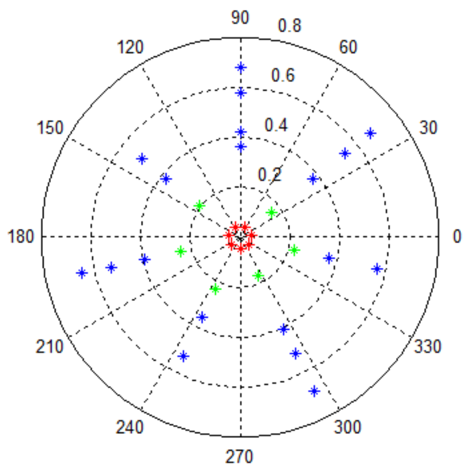

Figure8. The position of the microphone arrays in the experiment.

Divide the plane into 20*20 grids and regard each grid as a test point. Then calculate the time delay based on the far-field modal, using a matrix to save the data from the microphone array and shift the column of data based on the time delay. Then use the autocorrelation matrix to solve the main propagation source and put the relative probability value into each grid and through the imaging method shows the acoustic source on the image. By using the loop statements and importing the video and audio data, the overlaid acoustic imaging is able to track the moving noise.

Focusing on the location of the microphones, there is a phenomenon that the accuracy of the orientation result depends on the distance among each microphone. In this experiment, choose (5) to get the value of each testing points. When measure the far-field acoustic sources, if place the microphones into a circle, smaller radius leads to the inaccurate results, since the time-delay can only be calculated in integer and the round algorithm cause some error, especially when the sound locates almost in the center of the arrays. The interpolation method will improve the inaccurate results from the compact arrays. Make the length of the original data collected by the microphone expand some times to improve the sampling frequency. In this way, higher sampling frequency can make time delay more accurate and have the better acoustic sources.

However when choosing the first layer of the microphone array, the result is not accurate, and the deviation is unsatisfied. For the second layer of the microphones have the smaller radius than that in the first layer. The fractional part of the real time shift is not negligible, so the method of interpolation should be added.

In Fig.9, these two figures show different locations of the moving acoustic source and both choose the second layer of microphones. By using the loop statements and importing the video and audio data, the overlaid acoustic imaging is able to track the moving noise. When choosing suitable microphone arrays and algorithm, similarly get accurate results.

Figure9. Tracking the moving acoustic source.

In Fig.10, both two figures choose the first layer of the microphone array and compare with the accuracy of the acoustic source in the same situation and the only difference is that whether it uses interpolation method or not. From the Fig.10, the algorithm, which uses interpolation method (the right figure), gets more accurate results than that don't use the method (the left figure).

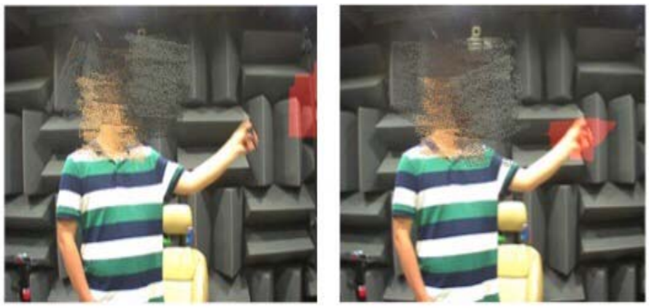

Figure10. The improvement of accuracy with interpolation method.

CONCLUSION

The beam forming with the combination of geometry and autocorrelation matrix is used for acoustic source localization. To verify the validity of the proposed method, the simulation and experiment are performed to compare and analyze the compute results of the moving sources. The results of simulation show that the far-field modal in time domain, and the use of interpolation method and autocorrelation matrix in beamforming estimation can acquire the accuracy orientation results. The results of experiment suggest that the proposed method has a good computing ability to tracking the moving sound sources.

ACKNOWLEDGMENT

This work was supported by National Training Program of Innovation and Entrepreneurship for Undergraduates (Project 201710145084), the National Natural Science Foundation of China (No. 61201054), the Fundamental Research Funds for the Central Universities (N130404006), and the Scientific Research Starting Foundation for Returned Overseas Chinese Scholars, Ministry of Education, China (No.50-4).

REFERENCES

[1] D. H. Johnson, and D. E. Dudgeon, Array Signal Processing, Concepts and Techniques, First ed. Prentice-Hall, New Jersey, 1993.

[2] U. Michel, History of acoustic beam forming, in Proceedings of the 1st Berlin Beam forming Conference, pp.1–17, 2006.

[3] Caleb Rascon, and Ivan Meza, "Localization of sound sources in robotics: A review, " Robotics and Autonomous Systems, vol.96, pp. 184-210, 2017.

[4] Ning, Fangli, Jingang Wei, Lianfang Qiu, Hongbing Shi, and Xiaofan Li, "Three-dimensional acoustic imaging with planar microphone arrays and compressive sensing," Journal of Sound and Vibration, vol. 380, pp. 112-128, 2016.

[5] Wei Ma, and Xun Liu, "Improving the efficiency of DAMAS for sound source localization via wavelet compression computational grid," Journal of Sound and Vibration, vol. 395, pp. 341-353, 2017.

[6] Zhigang Chu, and Yang Yang, "Comparison of deconvolution methods for the visualization of acoustic sources based on cross-spectral imaging function beamforming," Mechanical Systems and Signal Processing, vol. 48(1-2), pp. 404-422, 2014.

[7] Zhigang Chu, Yang Yang, Linbang Shen, and Guoli Ping, "Enhancement of two-dimensional acoustic source identification with Fourier-based deconvolution beamforming," Journal of Vibroengineering, vol. 18(5), pp. 3337-3361, 2016.

[8] Feng Miao, Diange Yang, Junjie Wen, and Xiaomin Lian, "Moving sound source localization based on triangulation method," Journal of Sound and Vibration, vol. 385, pp. 93-103, 2016.

[9] Shuangyan Quan, and Yang Cao, “The Application and Research of Interpolation Method,” Science and Technology Information (Science and Technology), vol(36), pp. 413-414, 2007

[10] Yilu Zhao, Xiong Chen, and Bin Wang, "Real-time sound source localization using hybrid framework, " Applied Acoustics, vol.74, pp.1367-1373, 2013.

[11] J. J. Christensen, and J. Hald, “Improvements of cross spectral beamforming,” 32nd International Congress and Exposition on Noise Control Engineering(Inter-Noise), Seogwipo, Korea, pp. 2652–2659, August 25–28, 2003.

[12] F. Ripamonti, A. Cigada, and M. Vanali, “The delay and sum algorithm applied to microphone array measurements, numerical analysis and experiment a validation,” Mech. Syst. Signal Process, vol. 21(6), pp. 2645–2664, 2007.

[13] Z. G. Chu, Y. Yang, W. D. Wang, X. B. Xiao, Z. H. Jiang, and Y. S. He, “Identification of truck noise sources under passby condition based on wave beamforming method,” J. Vib. Shock, vol.31(7), pp.66–70, 2012.

[14] T. Yardibi, C. Bahr, N. Zawodny, F. Liu, L. N., Cattafesta III, and J. Li, “Uncertainty analysis of the standard delay-and-sum beamformer and array calibration,” J. Sound Vib, vol. 329(13), pp. 2654–2682, 2010.