ISSN 1330-3651(Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20250425002622
Original scientific paper

An Algorithm for Microphone Array Sparse Arrangement and Acoustic Imaging

Chenyang ZHU, Yuguo ZHOU, Zhongming DONG, Mingming GUO, Qinying ZHANG, Deguang LIU, Hong CAO*

Abstract: Acoustic imaging is a technology for visualizing sound fields, which can be used in fault diagnosis of power equipment and gas leak detection. A spiral microphone array with a large number of microphones is generally used in ultrasonic acoustic imaging, but its application is limited due to the high cost. This paper addresses the issue of the high cost for acoustic imaging microphone arrays. On one hand, an adaptability function suitable for acoustic imaging applications is designed, and the genetic algorithm is used for sparse array arrangement to reduce the cost. On the other hand, the beam forming algorithm is also improved by using a time-varying smoothing coefficient to update the cross-power spectral matrix, further enhancing the performance. Experimental results show that by using the algorithms proposed in this paper, a sparse array can achieve acoustic imaging performance close to that of a large array, the cost of usage is significantly reduced.

Keywords: acoustic imaging; beamforming; genetic algorithm; microphone array; sparse array arrangement

1 INTRODUCTION

Acoustic imaging, also known as an acoustic camera, uses a microphone array to collect multi-channel audio data, and the beamforming algorithm is used to calculate the sound source distribution information on the imaging plane at a specified frequency. This information is then fused with a real-world image to intuitively determine the spatial location and origin of the sound source [1-4].

Since the ultrasonic signals are generated in scenarios such as partial discharges and gas leaks [5-8], and also the directivity of these signals is strong, acoustic imaging can be used for monitoring and detection. Compared with the other array structures, the spiral array [9-12] is suitable for ultrasonic frequency band, not only the main lobe width is narrow, but also better suppresses the side lobes, making it a mainstream array for acoustic imaging applications. However, spiral arrays generally require a large number of microphone elements, resulting in high costs.

The genetic algorithm [13-14], a population-based intelligent optimization algorithm that simulates natural biological evolution, has been widely used in sparse array arrangement due to its advantages in solving complex problems such as nonlinearity, large sample spaces, and combinatorial optimization. The deconvolution beamforming algorithm [15-18] can be used in acoustic imaging, which takes the output of beamforming as an intermediate quantity and iterates multiple times to eliminate side lobes and obtain a clean acoustic image. But the fixed weighting vectors of beamforming limit its adaptability to the environment, which can affect the results of acoustic imaging.

This paper focuses on the issue of high cost in ultrasonic acoustic imaging microphone arrays. First, an adaptability function suitable for acoustic imaging applications is designed, and the genetic algorithm is used to configure a sparse microphone array. Then, the beamforming algorithm is improved by using a time-varying smoothing coefficient to update the cross-power spectral matrix, with the smoothing coefficient adjusted based on the presence probability of ultrasonic signal at each time-frequency bin, further enhancing the algorithm's performance. Finally, the effectiveness of the proposed method is verified through simulation and experimental results.

2 SPARSE ARRAY ARRANGEMENT

2.1 Microphone Array Beam Pattern Function

Assume the microphone array consists of M microphone elements. Taking the center of the microphone array as the origin, a plane Cartesian coordinate system is established as shown in Fig. 1.

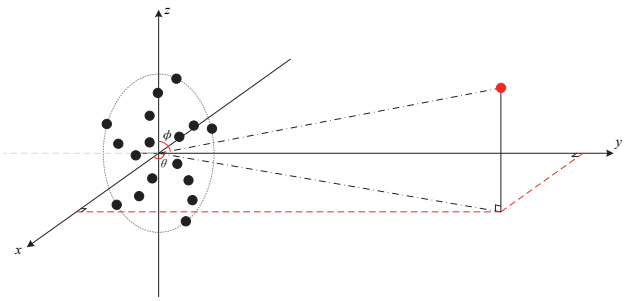


Figure 1 Microphone array coordinate system

where, the azimuth angle $\theta \in [0, 180]°$ is the angle between the positive half-axis of x and the projection of the sound source on the xoy plane, and the elevation angle $\varphi \in [0, 180]°$is the angle between the positive half-axis of z and the sound source.

Assume the main beam of the microphone array is directed at (θ_0, φ_0), and the element pattern of the $m(m = 1, 2, \ldots, M)$-th microphone element at the $f(f = 1, 2, \ldots, F)$-th frequency bin is $p_m(f, \theta, \varphi)$, with an amplitude weighting coefficient of a_m.Thus, the beam pattern function of the microphone array at the f-th frequency bin can be expressed as:

$$P(f, \theta, \varphi) = \sum_{m=1}^{M} a_m p_m(f, \theta, \varphi) e^{j\frac{2\pi f}{c}\left[x_m(\cos\theta\sin\varphi - \cos\theta_0\sin\varphi_0) + y_m(\sin\theta\sin\varphi - \sin\theta_0\sin\varphi_0) + z_m(\cos\varphi - \cos\varphi_0)\right]} \quad (1)$$

where j represents the imaginary part, c represents the speed of sound, and (x_m, y_m, z_m) represents the coordinates of the m-th microphone element.

If the microphone elements are omnidirectional and no amplitude weighting is applied, i.e., $p_m(f, \theta, \varphi) = 1$ and $a_m = 1$, the beam pattern function can be simplified as:

$$P(f, \theta, \varphi) = \sum_{m=1}^{M} e^{j\frac{2\pi f}{c}\left[x_m(\cos\theta\sin\varphi - \cos\theta_0\sin\varphi_0) + y_m(\sin\theta\sin\varphi - \sin\theta_0\sin\varphi_0) + z_m(\cos\varphi - \cos\varphi_0)\right]} \tag{2}$$

2.2 Sparse Array Arrangement Based on Genetic Algorithm

The basic idea of using the genetic algorithm for sparse array arrangement is to encode the working state of the microphone elements in the beam pattern as genes into chromosomes. Then, a fitness function is constructed to evaluate the quality of the chromosomes, and the best chromosomes are selected for reproduction, crossover, and mutation. Finally, the optimal solution is obtained through iterative evolution.

Let q_m represent the working state of the m-th microphone element, where $q_m = 1$ indicates the presence of an element at that position, and $q_m = 0$ indicates the absence of an element. Then, the beam pattern function can be expressed as:

$$P(f, \theta, \varphi) = \sum_{m=1}^{M} e^{j\frac{2\pi f}{c}\left[x_m(\cos\theta\sin\varphi - \cos\theta_0\sin\varphi_0) + y_m(\sin\theta\sin\varphi - \sin\theta_0\sin\varphi_0) + z_m(\cos\varphi - \cos\varphi_0)\right]} \cdot q_m \tag{3}$$

Thus, the azimuth and elevation patterns of the microphone array can be expressed as:

$$P_1(f, \theta) = P_{dB}(\theta, \varphi_0; f, \theta_0, \varphi_0) \tag{4}$$

$$P_2(f, \varphi) = P_{dB}(\theta_0, \varphi; f, \theta_0, \varphi_0) \tag{5}$$

where $P_{dB}(f, \theta, \varphi)$ represents the normalized beam pattern function.

To ensure that the optimized microphone array still has a good performance in acoustic imaging, the fitness function of the genetic algorithm consists of three parts:

(1) The average of the maximum side lobe levels at each frequency bin in azimuth and elevation:

$$L_1 = \frac{1}{F}\sum_{f=1}^{F}\frac{1}{2}\left[\max_{\theta\in S_1} P_1(f, \theta) + \max_{\varphi\in S_2} P_2(f, \varphi)\right] \tag{6}$$

where, $\max(\cdot)$ represents the maximum value, S_1 represents the side lobe intervals of azimuth beam pattern function with $\varphi = \varphi_0$, S_2 represents the side lobe intervals of elevation beam pattern function with $\theta = \theta_0$, respectively. $\max_{\theta\in S_1} P_1(f, \theta)$: Maximizes response P_1 over parameter set S_1 (e.g., azimuth angles). $\max_{\varphi\in S_2} P_2(f, \varphi)$: Maximizes response P_2 over parameter set S_2 (e.g., elevation angles).

(2) The variance of the maximum side lobe levels at each frequency bin in azimuth and elevation:

$$L_2 = \frac{1}{F}\sum_{f=1}^{F}\left[\max_{\theta\in S_1} P_1(f, \theta) + \max_{\varphi\in S_2} P_2(f, \varphi) - 2\mu\right]^2 \tag{7}$$

where, the mean value $\mu = L_1$.

(3) The average difference of main lobe width at each frequency bin in azimuth and elevation:

$$L_3 = \frac{1}{F}\sum_{f=1}^{F}\left|\underset{\theta\in\overline{S_1}}{bw} P_1(f, \theta) - \underset{\varphi\in\overline{S_2}}{bw} P_2(f, \varphi)\right| \tag{8}$$

where, $bw(\cdot)$ represents the main lobe width of beam pattern, $\overline{S_1}$ represents the complementary set of S_1, that is the main lobe intervals in azimuth with $\varphi = \varphi_0$, $\overline{S_2}$ is in the same vein.

Thus, the fitness function of the genetic algorithm can be expressed as:

$$L = L_1 - \lambda_1 L_2 - \lambda_2 L_3 \tag{8}$$

where $\lambda_1$and λ_2 are weighting coefficients. The goal is to minimize the side lobe levels while ensuring that the side lobe levels are consistent across frequency bins and that the main lobe widths are similar in both azimuth and elevation. By optimizing the fitness function using the genetic algorithm, the positions of the microphone elements that satisfy a certain sparsity rate can be determined.

3 ACOUSTIC IMAGING WITH MICROPHONE ARRAYS

3.1 Principle of Acoustic Imaging Algorithm

The CLEAN-SC algorithm [19], which is widely used, is based on the principle that the main lobe and side lobes generated by the same sound source are completely coherent. Through multiple iterations, the side lobes coherent with the main lobe can be removed, thereby improving the resolution of acoustic imaging.

The algorithm consists of three steps: Step 1, obtain the initial acoustic imaging result through beamforming, called the dirty map; Step 2, identify the peak in the dirty map and remove the side lobes generated by this peak, update the dirty map; Step 3, iterate multiple times to obtain a clean acoustic image. Since the first step of beamforming usually determines the result of acoustic imaging, the performance of the beamforming algorithm is crucial for acoustic imaging.

Assume the imaging plane is divided into N grids. The beamforming output at the n(n = 1, 2, …, N)-th grid is:

$$b(\theta_n, \varphi_n) = \sum_{f=1}^{F} \boldsymbol{\omega}^H (f, \theta_n, \varphi_n) \boldsymbol{C}(t,f) \boldsymbol{\omega}(f, \theta_n, \varphi_n) \quad (9)$$

where, t represents the frame number after segmentation, f represents the frequency bin after Fourier transform, $\boldsymbol{\omega}(f, \theta_n, \phi_n)$ represents the beamforming weight vector, $(\cdot)^H$ represents the conjugate transpose, $\boldsymbol{C}(t, f)$ represents the cross-power spectral matrix of the received signals from the microphone array, which can be obtained through time-frequency smoothing:

$$\boldsymbol{C}(t,f) = \alpha \boldsymbol{C}(t-1,f) + (1-\alpha) \boldsymbol{X}(t,f) \boldsymbol{X}^H(t,f) \quad (10)$$

where, $\boldsymbol{X}(t, f)$ represents the received signal from the microphone array in time-frequency domain, and $\alpha \in (0, 1)$ is the smoothing coefficient. A larger smoothing coefficient results in slower updates to the cross-power spectral matrix, while a smaller smoothing coefficient results in faster updates.

3.2 Improved Smoothing Coefficient Method

As for acoustic imaging, a smaller smoothing coefficient allows a faster response to the changes in the sound source, but with some losses in localization accuracy. A larger smoothing coefficient improves localization accuracy but the slower response. For example, when the sound source stops emitting, the acoustic image will still display the sound source for a period of time before stopping. On the other hand, if the smoothing coefficient is set to a constant, each frame will be smoothed at the same rate, which may lead to the inclusion of more noise components in the cross-power spectral matrix, resulting in degraded acoustic imaging performance.

Since the noise is approximately stationary, the power of each frequency bin in the received signal can be used to determine whether the frequency bin is dominated by ultrasonic signals or noise signals. If it is dominated by ultrasonic signal, the cross-power spectral matrix should be updated more slowly, or if it is dominated by noise, the cross-power spectral matrix should be updated more quickly. Therefore, this paper sets the smoothing coefficient α as a time-frequency variable, adjusting it based on the presence probability of ultrasonic signal. The specific steps are as follows:

First, calculate the average power of each frequency bin using L frames:

$$\boldsymbol{S}_{\text{ave}}(f) = \frac{1}{L} \sum_{l=1}^{L} \left| \boldsymbol{X}(l,f) \right|^2 \quad (11)$$

Then, calculate the power of received signal at the current frame:

$$\boldsymbol{S}(t,f) = \left| \boldsymbol{X}(t,f) \right|^2 \quad (12)$$

Define the following ratio and compare it with a preset threshold σ:

$$\boldsymbol{r}(t,f) = \frac{\boldsymbol{S}(t,f)}{\boldsymbol{S}_{\text{ave}}(f)} \quad (13)$$

$$\hat{\boldsymbol{I}}(t,f) = \begin{cases} 1 & \boldsymbol{r}(t,f) \geq \sigma \\ 0 & \text{else} \end{cases} \quad (14)$$

Next, calculate the presence probability of ultrasonic signal at each frequency bin:

$$\boldsymbol{p}(t,f) = \alpha_p \boldsymbol{p}(t-1,f) + (1-\alpha_p) \hat{\boldsymbol{I}}(t,f) \quad (15)$$

where α_p represents the smoothing coefficient for the presence probability of ultrasonic signal.

Finally, set a range for $\alpha \in (\alpha_{\min}, \alpha_{\max})$, and adjust the smoothing coefficient based on the presence probability of ultrasonic signal. The expression for the smoothing coefficient [20] is as follows:

$$\boldsymbol{\alpha}(t,f) = \frac{\tan^{-1}\left[20\left(\boldsymbol{p}(t,f) - 0.5\right)\right] \cdot (\alpha_{\max} - \alpha_{\min})}{\pi} + \\ + \frac{\alpha_{\max} + \alpha_{\min}}{2} \quad (16)$$

where, $\alpha_{\min}$ is set to 0.50, and $\alpha_{\max}$ is set to 1.00, respectively. The relationship between the smoothing coefficient and the presence probability of ultrasonic signal is shown in Fig. 2.

When the presence probability of ultrasonic signal is 0, the smoothing coefficient is 0.5159, and when the probability is 1, the smoothing coefficient is 0.9841.

Thus, the cross-power spectral matrix can be updated as follows:

$$\boldsymbol{C}(t,f) = \boldsymbol{\alpha}(t,f) \boldsymbol{C}(t-1,f) + \\ + \left[1 - \boldsymbol{\alpha}(t,f)\right] \boldsymbol{X}(t,f) \boldsymbol{X}^H(t,f) \quad (17)$$

where the initial value of the cross-power spectral matrix is:

$$\boldsymbol{C}(L,f) = \frac{1}{L} \sum_{l=1}^{L} \boldsymbol{X}(l,f) \boldsymbol{X}^H(l,f) \quad (18)$$

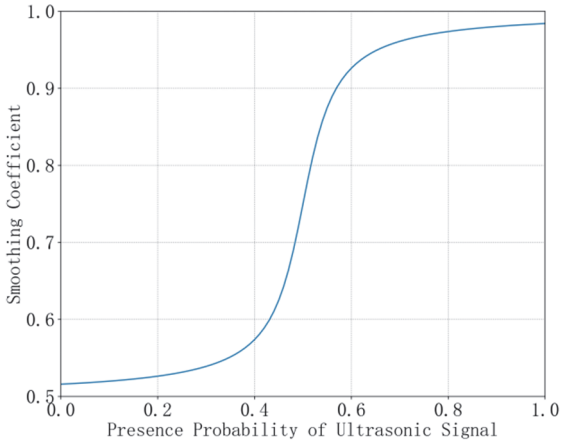


Figure 2 Relationship between smoothing coefficient and presence probability of ultrasonic signal

4 EXPERIMENTS AND RESULTS ANALYSIS

4.1 Sparse Array Arrangement Experiment

This paper considers optimizing a 128-microphone array to a 32-microphone sparse array.

The imaging frequency band of the microphone array is [20, 48] kHz, with a frequency step of 2 kHz. Microphone's frequency range is 6 Hz - 70 kHz (±2 dB), and dynamic range: 14 dB(A) - 146 dB. The main beam is directed at (90°, 90°), and the azimuth and elevation angle ranges are both [55°, 125°], with an angle step of 0.10°. The GA parameters are as follows:

Population size: $N = 50$ (empirically determined via 10 trial runs with $N \in [50, 300]$).

Crossover rate: $p_c = 0.80$ (two-point crossover).

Mutation rate: $p_m = 1/\ell$ (where ℓ = chromosome length), ensuring ~1 mutation per chromosome.

Selection: Tournament selection with $k = 3$.

Termination: Generation limit = 1000 OR fitness stagnation (< 0.1% improvement for 50 generations).

After optimization using the genetic algorithm, the average maximum side lobe level in the azimuth and elevation at each frequency bin is 16.11 dB, the variance is 0.11, and the average difference in main lobe widths is 0.28°. The array structure is shown in Fig. 3, where blue elements represent the 128-microphone array, and red elements represent the 32-microphone sparse array:

The beam patterns of the 32-microphone sparse array at 20 kHz, 34 kHz, and 48 kHz are shown in Fig. 4, where the blue curve represents the azimuth beam pattern, and the red curve represents the elevation beam pattern.

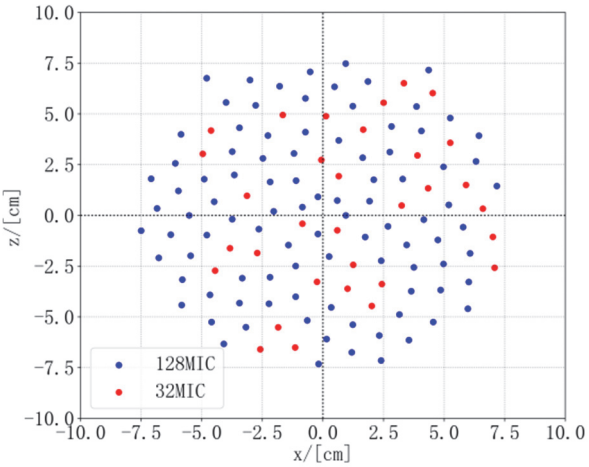
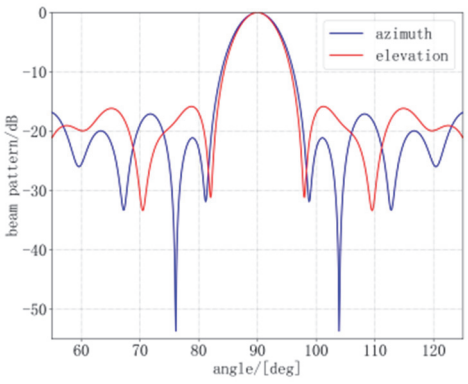
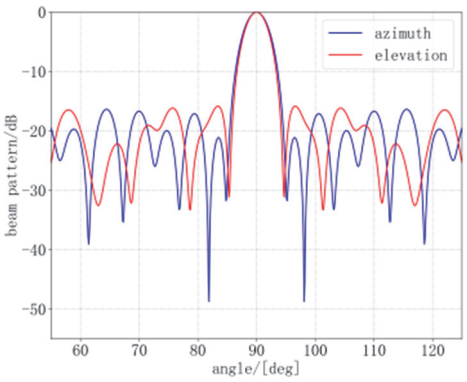


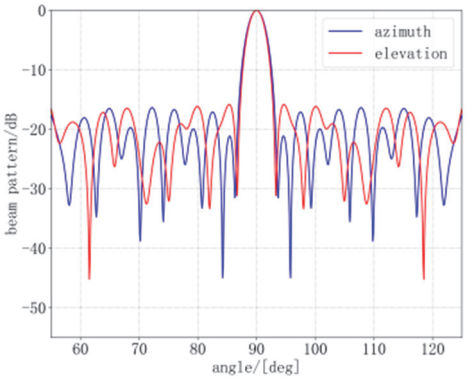
Figure 3 Comparison of array structures



20KHz



34KHz



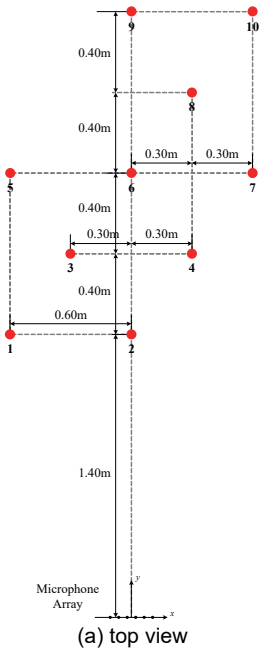
48KHz

Figure 4 Beam patterns of 32-microphone sparse array

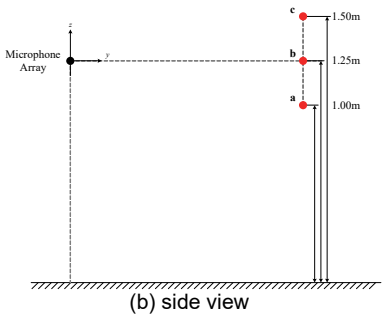
The result shows that the 32-microphone sparse array optimized by the genetic algorithm has a high dynamic range in the beam pattern, with consistent maximum side lobe levels and main lobe widths across frequency bins in both azimuth and elevation. While reducing the number of microphone elements by 96, the performance of the microphone array is well maintained.

4.2 Acoustic Imaging Experiment with Microphone Arrays

The following experiment uses a 128-microphone array in an office environment for multi-point acoustic imaging. The experimental setup is shown in Fig. 5.



(a) top view



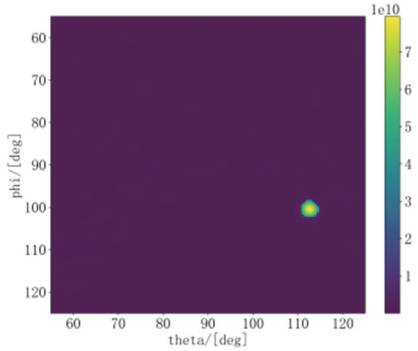
(b) side view

Figure 5 Experimental setup

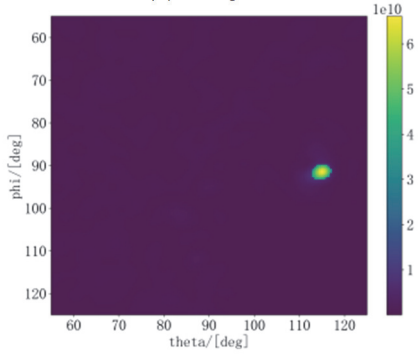
There are 10 test points, each of which includes 3 heights, totaling 30 acoustic imaging experiments. In each experiment, acoustic imaging is performed for approximately 10 seconds, with the ultrasonic sound source active for about 7 seconds.

4.2.1 Acoustic Imaging Comparison

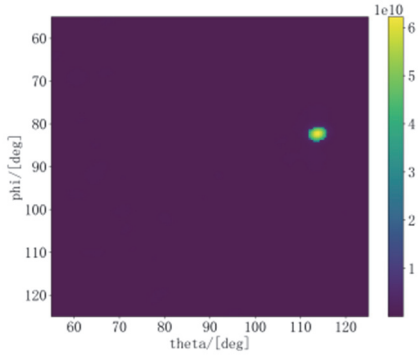
First, the imaging results of the 128-microphone array and the 32-microphone sparse array at test point 1, on different heights and the same time, with a smoothing coefficient of 0.5159, are compared in Fig. 6:



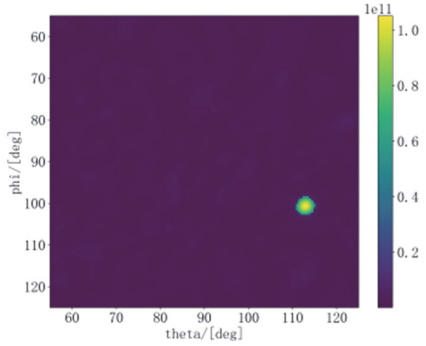
(a)-1 height a



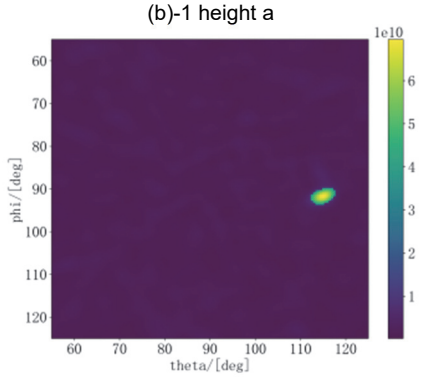
(a)-2 height b



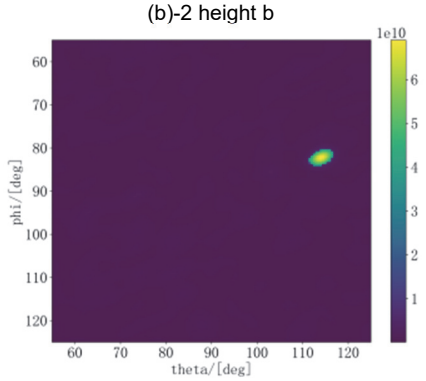
(a)-3 height c
128-microphone array



(b)-1 height a



(b)-2 height b



(b)-3 height c
32-microphone sparse array

Figure 6 Test point 1 imaging result comparison

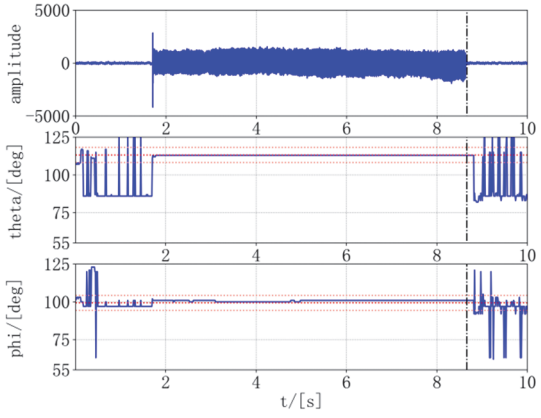
The results show that the imaging results of the optimized 32-microphone sparse array are consistent with those of the 128-microphone array, although the main lobe width of the sparse array is slightly wider.

4.2.2 Sound Source Localization Performance Comparison

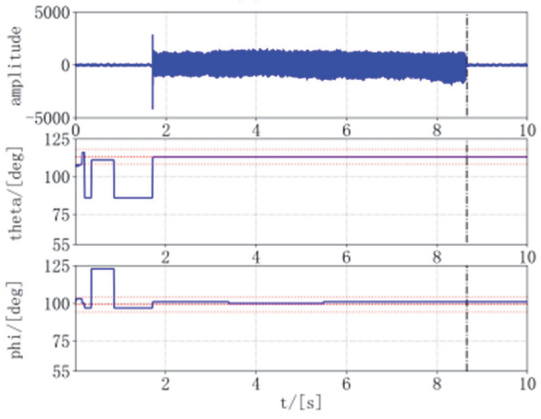
Next, the impact of different smoothing coefficients on sound source localization performance is compared, with the noise segment length L set to 1 second when the smoothing coefficient is time-varying.

The azimuth and elevation angle estimates corresponding to the maximum value in each frame's acoustic image are recorded as the sound source localization results. The localization results of the 128-microphone array and the 32-microphone sparse array at test point 1, height a, with different smoothing coefficients, are shown in Fig. 7. The first row shows the received signal from the first microphone element of the 128-microphone array, the second row shows the azimuth angle estimates, and the third row shows the elevation angle estimates. The red dashed line indicates the true position of the sound source, the orange dashed line

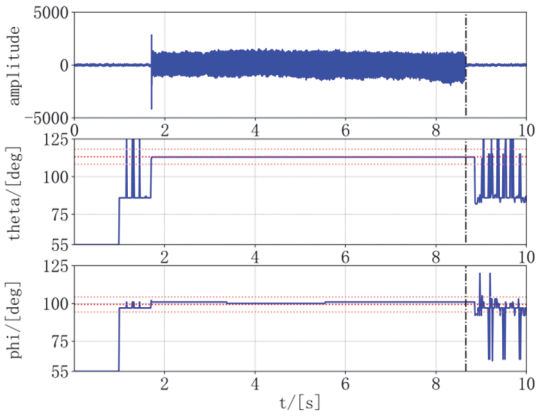
indicates the ±5° error range around the true position, and the black dotted line indicates the moment when the sound source stops emitting:



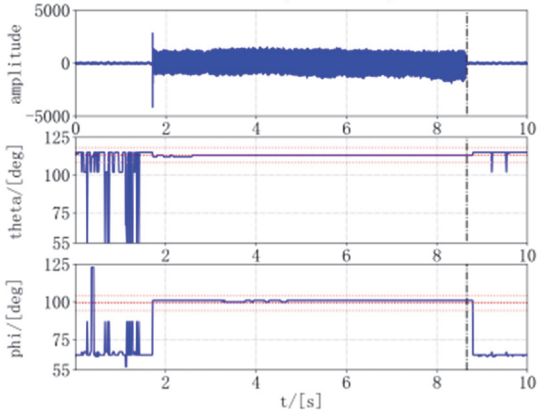
(a)-1 0.5159



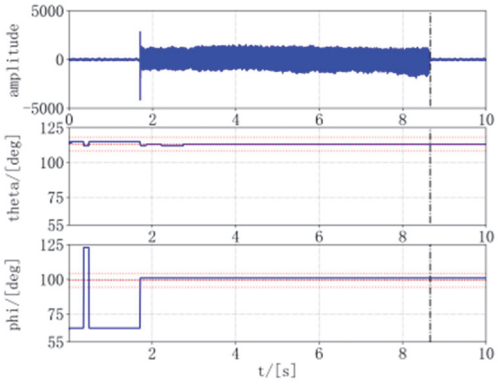
(a)-2 0.9841



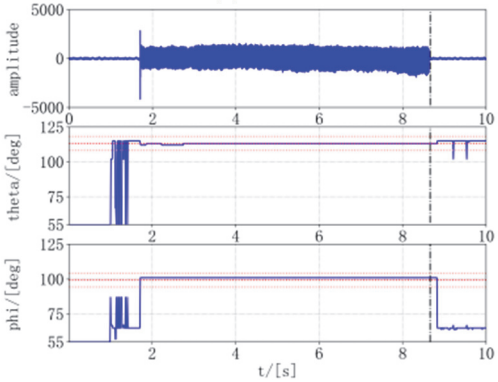
(a)-3 time-varying
128-microphone array



(b)-1 0.5159



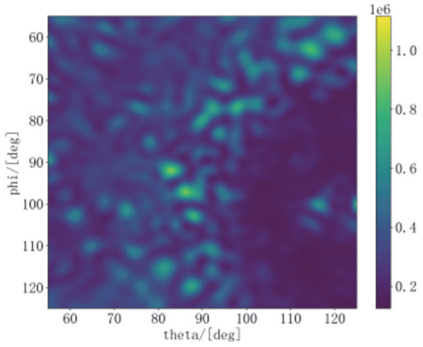
(b)-2 0.9841



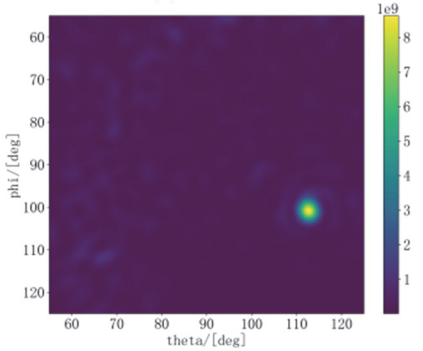
(b)-3 time-varying
32-microphone sparse array

Figure 7 Comparison of localization results with different smoothing coefficients

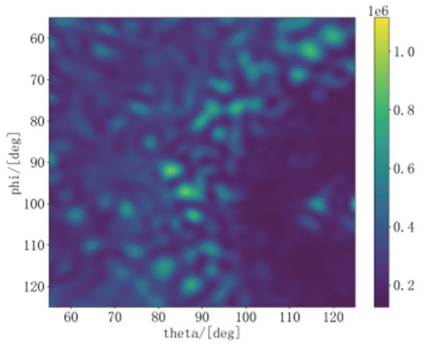
The results show that for both the 128-microphone array and the 32-microphone sparse array, the acoustic image quickly stops localizing the sound source after it stops emitting when the smoothing coefficient is small or time-varying. However, when the smoothing coefficient is large, the acoustic image continues to display the sound source for some time after it stops emitting.



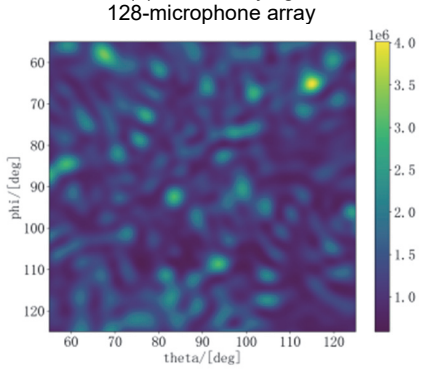
(a)-1 0.5159



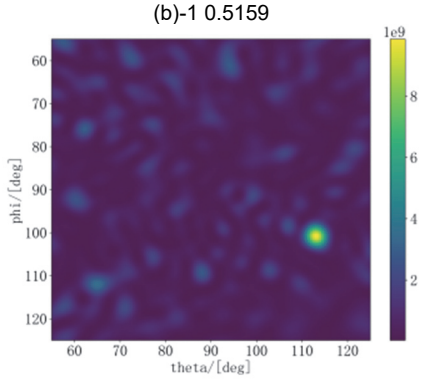
(a)-2 0.9841



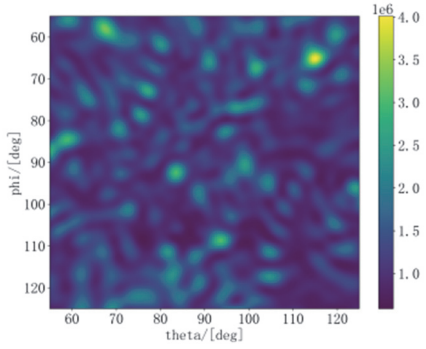
(a)-3 time-varying
128-microphone array



(b)-1 0.5159



(b)-2 0.9841



(b)-3 time-varying
32-microphone sparse array

Figure 8 Comparison of beamforming imaging results with different coefficients at the same time

Taking the moment approximately 1 second after the sound source stops emitting as an example, i.e., the time point at 9.65 seconds, the imaging results of the CLEAN-SC algorithm's first step (beamforming) for the 128-microphone array and the 32-microphone sparse array are shown in Fig. 8.

The results show that after the sound source stops emitting, the acoustic image still displays a clear and high-power sound source when the smoothing coefficient is large. However, when the smoothing coefficient is small or time-varying, the power in the acoustic image is significantly reduced, and no sound source is displayed. Since acoustic imaging should quickly reflect the actual sound source distribution, a large smoothing coefficient is not suitable for acoustic imaging applications, while a small or time-varying smoothing coefficient provides better imaging performance.

Finally, if the azimuth and elevation angle estimates of the current frame are within ±5° of the true sound source position, the localization is considered accurate. The sound source localization accuracy rates for the 128-microphone array and the 32-microphone sparse array, with different smoothing coefficients, are compared in Fig. 9. The blue bars represent the results for the 128-microphone array, and the red bars represent the results for the 32-microphone sparse array:

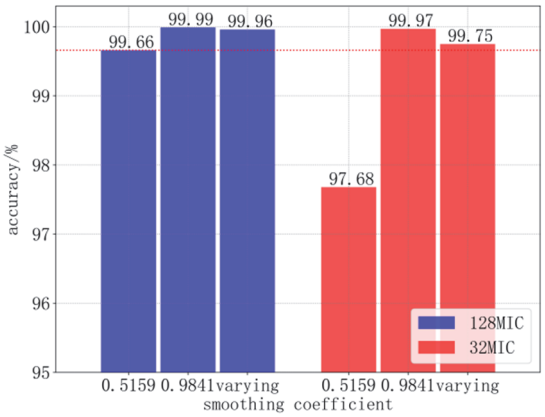


Figure 9 Comparison of sound source localization accuracy with different smoothing coefficients

The result shows that for both the 128-microphone array and the 32-microphone sparse array, the localization accuracy is higher when the smoothing coefficient is large, but this is not suitable. When the smoothing coefficient is small, the localization performance of the 32-microphone sparse array is significantly worse than that of the 128-microphone array, with a 1.98% loss in accuracy. However, when the smoothing coefficient is time-varying, the localization accuracy of the 32-microphone sparse array improves by 2.07%, and it further surpasses the accuracy of the 128-microphone array with a small smoothing coefficient by 0.09%. Additionally, when the smoothing coefficient is time-varying, the localization accuracy of both arrays is close to that achieved with a large smoothing coefficient.

In conclusion, when the smoothing coefficient is time-varying, it combines the advantages of both small and large smoothing coefficients, allowing for a fast response to actual sound source distribution while maintaining high localization accuracy. Furthermore, the localization performance of the sparse array is comparable to that of the large array, significantly improving the performance of the algorithm.

5 CONCLUSION

This paper addresses the high-cost challenge in acoustic imaging microphone arrays through two key innovations: (1) a genetic algorithm-based sparse array

optimization method that maintains imaging performance while reducing hardware costs, and (2) an improved beamforming algorithm using time-varying smoothing coefficients to enhance cross-power spectral matrix updates. Experimental results demonstrate that the proposed approach achieves acoustic imaging performance comparable to dense arrays while significantly lowering deployment costs. Future research directions include Dynamic Array Reconfiguration and Hybrid Beamforming Architectures. These directions aim to bridge the gap between computational acoustics and practical implementation, ultimately enabling low-cost, high-performance acoustic imaging across scientific and industrial domains.

Acknowledgements

This work was supported by China Yangtze Power Co., Ltd. (Contract No. Z532302001).

6 REFERENCES

[1] Sandro, G. & Roland, S. (2001). Advanced source localization techniques using microphone arrays. *SAE International Journal of Passenger Cars - Mechanical Systems*, *4*(2), 1241-1249. https://doi.org/10.4271/2011-01-1657

[2] Huang, X. (2009). Real-time algorithm for acoustic imaging with a microphone array. *Journal of the Acoustical Society of America*, *125*(5). https://doi.org/10.1121/1.3100641

[3] Orman, M., Rzeszucinski, P., & Pinto, C. T. (2014). Low cost hand held acoustic camera. *IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*. https://doi.org/10.1109/ICSPCC.2014.6986223

[4] Ilkhechi, H. D. & Samimi, M. H. (2021). Applications of the acoustic method in partial discharge measurement: a review. *IEEE Transactions on Dielectrics and Electrical Insulation*, *28*(1), 42-51. https://doi.org/10.1109/TDEI.2020.008985

[5] Choi, Y. C., Park, J. H., Yoon, D. B., & Kwon, H. S. (2008). Noise source identification in a reverberant field using spherical beamforming. *Modern Physics Letters B*, *22*(11), 1147-1151. https://doi.org/10.1142/S021798490801598X

[6] Orman, M., Orkisz, M., & Pinto, C. T. (2011). Parameter identification and slip estimation of induction machine. *Mechanical Systems & Signal Processing*, *25*(4), 1408-1416. https://doi.org/10.1016/j.ymssp.2010.11.004

[7] Liang, W., Zhang, L., Xu, Q., & Yan, C. (2013). Gas pipeline leakage detection based on acoustic technology. *Engineering Failure Analysis*, *31*, 1-7. https://doi.org/10.1016/j.engfailanal.2012.10.020

[8] Pihera, J., Hornak, J., Trnka, P., Turecek, O., & Albrecht, R. (2020). Partial discharge detection using acoustic camera. *IEEE 3rd International Conference on Dielectrics (ICD)*. https://doi.org/10.1109/ICD46958.2020.9341902

[9] Messer, H. (1992). Source localization performance and the array beampattern. *Signal Processing*, *28*(2), 163-181, https://doi.org/10.1016/0165-1684(92)90033-S

[10] Rafaely, B. (2005). Analysis and design of spherical microphone arrays. *IEEE Trans. Speech Audio Process*, *13*(1), 135-143. https://doi.org/10.1109/TSA.2004.839244

[11] Rafaely, B., Weiss, B., & Bachmat, E. (2007). Spatial aliasing in spherical microphone arrays. *IEEE Transactions on Signal Processing*, *55*, 1003-1010. https://doi.org/10.1109/TSP.2006.888896

[12] Li, Z. & Duraiswami, R. (2007). Flexible and optimal design of spherical microphone arrays for beamforming. *IEEE Transactions on Audio, Speech, and Language Processing*, *15*(2), 702-714. https://doi.org/10.1109/TASL.2006.876764

[13] Holland, J. H. (2000). Building Blocks, Cohort Genetic Algorithms, and Hyperplane-Defined Functions. *Evolutionary Computation*. *8*(4), 373-391. https://doi.org/10.1162/106365600568220

[14] Chen, K., He, Z., & Han, C. (2006). A modified real GA for the sparse linear array synthesis with multiple constraints. *IEEE Transactions on Antennas & Propagation*, *54*(7), 2169-2173. https://doi.org/10.1109/TAP.2006.877211

[15] Castellini, P. & Martarelli, M. (2008). Acoustic beamforming: analysis of uncertainty and metrological performances. *Mechanical Systems and Signal Processing*, *22*(3), 672-692. https://doi.org/10.1016/j.ymssp.2007.09.017

[16] Chu, Z. & Yang, Y. (2014). Comparison of deconvolution methods for the visualization of acoustic sources based on cross-spectral imaging function beamforming. *Mechanical Systems & Signal Processing*, *48*, 1-2. https://doi.org/10.1016/j.ymssp.2014.03.012

[17] Chiariotti, P., Martarelli, M., & Castellini, P. (2019). Acoustic beamforming for noise source localization - reviews, methodology and applications. *Mechanical Systems and Signal Processing*, *120*, 422-448. https://doi.org/10.1016/j.ymssp.2018.09.019

[18] Merino-Martinez, R., Sijtsma, P., Snellen, M., Ahlefeldt, T., & Spehr, C. (2019). A review of acoustic imaging methods using phased microphone arrays. *CEAS Aeronautical Journal*, *10*, 197-230. https://doi.org/10.1007/s13272-019-00383-4

[19] Sijtsma, P. (2009). Clean based on spatial source coherence. *International Journal of Aeroacoustics*, *6*(4), 357-374. https://doi.org/10.1260/147547207783359459

[20] Li S F, Li Y, & Yan Z Y. (2007). Self-adaptive noise elimination with single channel RLS algorithm and the determination of the forgotten factor. *Journal of Changchun University of Science and Technology (Natural Science Edition)*, *28*, 224.

Contact information:

Chenyang ZHU
Baihetan Hydropower Plant, China Yangtze Power Co., Ltd.,
Liangshan Sichuan, 615421, China

Yuguo ZHOU
Baihetan Hydropower Plant, China Yangtze Power Co., Ltd.,
Liangshan Sichuan, 615421, China

Zhongming DONG
Wudongde Hydropower Plant, China Yangtze Power Co., Ltd.,
Kunming Yunnan, 650000, China

Mingming GUO
Baihetan Hydropower Plant, China Yangtze Power Co., Ltd.,
Liangshan Sichuan, 615421,China

Qinying ZHANG
Baihetan Hydropower Plant, China Yangtze Power Co., Ltd.,
Liangshan Sichuan, 615421,China

Deguang LIU
Tsinghua AI Plus Co., Ltd.,
Beijing, Beijing, 100000, China

Hong CAO
(Corresponding author)
Tsinghua AI Plus Co., Ltd.,
Beijing, Beijing, 100000, China
E-mail: aithu6798@163.com