DOI: 10.61782/fa.2023.0767

APPLICATION OF SOUND SOURCE LOCALIZATION ALGORITHM FOR URBAN NOISE SOURCE MONITORING

In-Jee Jung[1*] **Sung Soo Jung**[1] **Wan-Ho Cho**[1]
Byung Kwon Lee[2] **Taeho Park**[2]
[1] Korea Research Institute of Standards and Science, Daejeon, Korea
[2] Korea Environment Institute, Sejong, Korea

ABSTRACT

Urban noise is an important social issue because it is a major factor that reduces the quality of human life. Recently, the noise level in the urban area has been increasing due to changing the social climate and emerging new mobility systems. Therefore, many countries try to control and reduce urban noise in many ways. In this study, sound source localization algorithms are applied to estimate the position of a moving noise source. To this end, the simulation is conducted with modeling of the moving source. The localization results through TDoA, beamforming, and sound intensimetry are compared for a point source passing through the road. The localization error and latency are analyzed for different sampling rates and array sizes. One can find that beamforming is most affected by the sampling frequency, and it can be observed that the sampling rate is inversely proportional to the accuracy in general. Sound intensimetry has higher accuracy than other algorithms when the array size is small. TDoA has the lowest latency, so it is advantageous for real-time estimation. The test result shows that it is necessary to apply an appropriate localization algorithm according to the system configuration to conduct urban noise source monitoring.

Keywords: *urban noise monitoring, real-time sound source monitoring, source localization algorithm*

[*]***Corresponding author***: *injee@kriss.re.kr*

1. INTRODUCTION

The acoustic environment is one of the important factors for the quality of human life. In particular, noise is a serious problem that causes physical and mental disorders. When the noise is exposed for a long time, one can cause various negative effects on health, such as heart disease, stroke, and sleep disturbance [1]. Nowadays, the soundscape in urban areas has been dramatically changing because of emerging new mobility systems and varying social climates [2]. Therefore, many countries try to control and reduce noise in many ways [3].

Localization or identification of the sound source is the first step of noise control in general. Because a stationary sound source is predictable and takes a low effort to estimate the source position, various countermeasures can be applied to reduce the noise. However, the major noise sources in urban areas, such as vehicles or flying objects, are not easy to control because the dynamic characteristic is unpredictable [4].

One of the methods to reduce the noise from excessively loud vehicles is taking enforcement measures by imposing fines or rectification notices [5]. To this end, a quantitative measure of the noise level and source identification is needed.

In this work, sound source localization algorithms are applied to estimate the direction of arrival(DoA) of a moving noise source and the feasibility of application for localizing vehicles in the real-time. In order to be applied to actual problems, not only accuracy but also fast enough tracking should be possible compared to the speed of the moving sound source. Typical localization algorithms are used to localize a point source passing through the road, and the test results are compared.

2. LOCALIZATION OF MOVING SOURCE

2.1 A moving sound source model for simulation

The sound pressure measured by the stationary microphones for a moving sound source as illustrated in Fig. 1 can be expressed as [6]

$$p_n(t) = \frac{1}{4\pi R_n(t)\left(1 - M\cos\delta_n(t)\right)^2} s\left(t - \frac{R_n(t)}{c}\right). \quad (1)$$

Here, $p_n(t)$ is the sound pressure measured by the n-th microphone, $R(t)$ the distance between the source and the microphone, c the speed of sound, M the Mach number, $\delta_n(t)$ the angle between the moving direction and the position of the microphone, $s(t)$ the source signal defined by the air density and volume velocity of the source. Because the relative magnitude and phase between the measured signals from the microphones are the major parameters in the localization problem, in this study, a recorded vehicle sound whose magnitude was normalized to a maximum value is used as a source signal.

In addition to the Doppler effect, the ground reflection affects the measured sound when the moving source is on the paved road. Therefore, the image source method is adopted to generate the test signal. Figure 2 shows the spectrogram of a test signal for a vehicle driving on the road with an initial velocity of 25 m/s and acceleration of 3 m/s^2, which is generated by using a recorded vehicle sound [7].

2.2 Localization algorithms

The typical localization method, delay-and-sum beamforming, time difference of arrival (TDoA), and sound intensimetry are used to localize the moving source. Although the algorithms are widely used and already well-known methods for sound source localization, a brief explanation is given in this section.

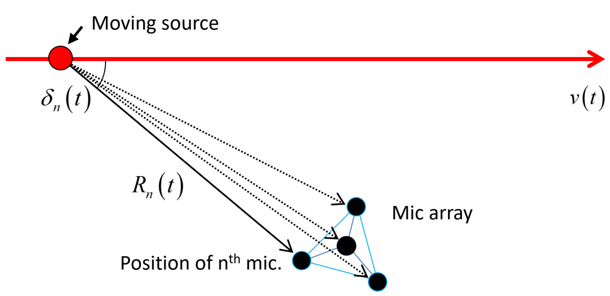

Figure 1. Measured sound pressure using a microphone array for a moving sound source.

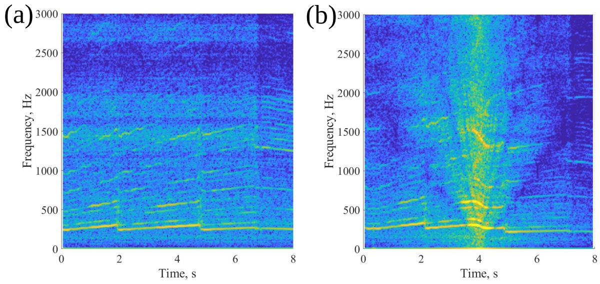

Figure 2. The spectrogram of the test signal. (a) Original source, (b) moving source.

Beamforming is a method of radiating or receiving wave energy by concentrating a beam in a specific direction through a transducer array. The beam is a function of bearing angles, which is defined as a beamformer, and various types of beamforming algorithms are applied for the sound field control, and a delay-and-sum (DAS) beamformer [8] is popularly used in many acoustic problems.

A beam of a specific direction can be calculated by adjusting the time delay of the signal measured from N microphones, and the output of the beam in the time domain can be expressed as

$$b_t(t,m) = \sum_{n=1}^{N} w_n p_n\left(t - D_n(m)\right), \quad (2)$$

where w_n is the aperture shading coefficient of the beam [9], p_n the signal measured from the n-th microphone, and D_n the time delay applied to the n-th microphone to form a beam in the m-direction. The output of the beam in the frequency domain can be obtained by Fourier transform of Eq. (2), which can be written as

$$b_f(f,m) = \sum_{n=1}^{N} w_n P_n(f) e^{-j2\pi f D_n(m)}. \quad (3)$$

When the beamforming is used to localize the sound source, the direction that has the maximum beam output value is determined as the DoA of the sound source. The estimated azimuth angle and elevation angle by using the beamforming method can be calculated by

$$\left(\phi_{DAS}, \theta_{DAS}\right) = \arg\max_{\phi,\theta} b(\phi,\theta), \quad (4)$$

where $\left(\phi_{DAS}, \theta_{DAS}\right)$ denote the azimuth angle and elevation angle estimated by beamforming, $b(\phi,\theta)$ is the spatial response of the delay-and-sum beamformer for azimuth angle, ϕ and elevation angle, θ.

TDoA is a method for the localization of a sound source by calculating the time difference of sound waves arriving at the microphones. Since the time difference for the sound wave can be simply estimated through the cross-correlation function between the microphones, the method has a low computational cost and can be easily implemented in any system. Also, generalized cross-correlation algorithm is widely used, which has the advantage of high robustness in a reverberant sound field with high spatial resolution [10]. Figure 3 shows the microphone array randomly arranged in three-dimensional space to localize the plane wave source using the TDoA method. The time difference can be calculated from a set of two microphones, which can be expressed as

$$\mathbf{d}_{\mathbf{ij}} \cdot \mathbf{s}_{\mathbf{ij}} = c \cdot \tau_{ij}, \tag{5}$$

where $\mathbf{d}_{\mathbf{ij}} = \mathbf{r}_{\mathbf{j}} - \mathbf{r}_{\mathbf{i}}$, $\mathbf{r}_{\mathbf{i}}$, $\mathbf{r}_{\mathbf{j}}$ indicate the position vector of the i-th and j-th microphone in Cartesian coordinate, $\mathbf{s}_{\mathbf{ij}}$ the DoA vector calculated from the microphone set, and τ_{ij} the time difference between the i-th and j-th microphones.

Because the $\mathbf{s}_{\mathbf{ij}}$ have the same direction vector for the plane wave source, the DoA vector estimated from the microphone array, which is consisted of N microphones, can be written as [11]

$$\mathbf{D} \cdot \mathbf{s}^{\mathbf{T}} = c \cdot \mathbf{T}. \tag{6}$$

Here, $\mathbf{s} = \left[s_x,\ s_y,\ s_z \right]$ is the DoA vector estimated from the acoustic center of the microphone array, $\mathbf{D} = \left[\mathbf{d}_{12}, \mathbf{d}_{13}, \cdots, \mathbf{d}_{\mathrm{ij}} \right]^T$, $\mathbf{T} = \left[\tau_{12},\ \tau_{13}, \cdots,\ \tau_{\mathrm{ij}} \right]$ denotes the position vector and time difference matrix of the microphone array, respectively.

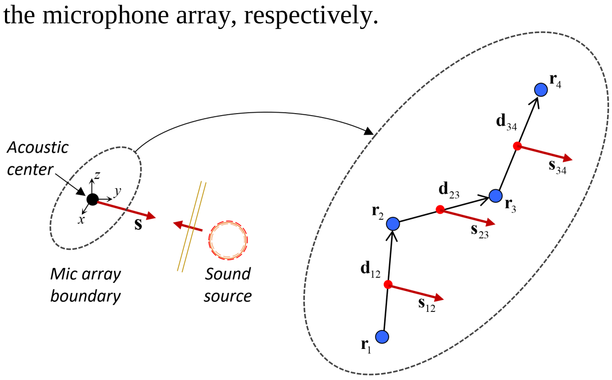

Figure 3. An illustration of a randomly arranged microphones used for source localization by using TDoA method.

A number of microphones are used to reduce the localization uncertainty and avoid localization error due to cone of confusion. Accordingly, the DoA vector estimated from the microphone array can be calculated by

$$\mathbf{s}^{\mathbf{T}} = c\left(\mathbf{D}^T \cdot \mathbf{D}\right)^{-1} \cdot \mathbf{D}^T \cdot \mathbf{T}. \tag{7}$$

The estimated azimuth angle and elevation angle by using the TDoA method can be calculated by

$$\phi_{TDOA} = \tan^{-1}\left(\frac{s_y}{s_x}\right),\ \theta_{TDOA} = \tan^{-1}\left(\frac{s_z}{\sqrt{s_x^2 + s_y^2}}\right) \tag{8a,b}$$

where ϕ_{TDOA}, θ_{TDOA} denotes the azimuth angle and elevation angle estimated by TDoA method, respectively.

Sound intensimetry estimates the active intensity at the acoustic center of the array module, which is consisted of microphones or particle velocity sensors. Therefore, one can be used to localize the sound source in the far field.

In the estimation method using a microphone array, i.e. p-p method, the microphone spacing must be far shorter than a wavelength to approximate the particle velocity, so this method is suitable for implementation to a compact hardware system or localization of a sound

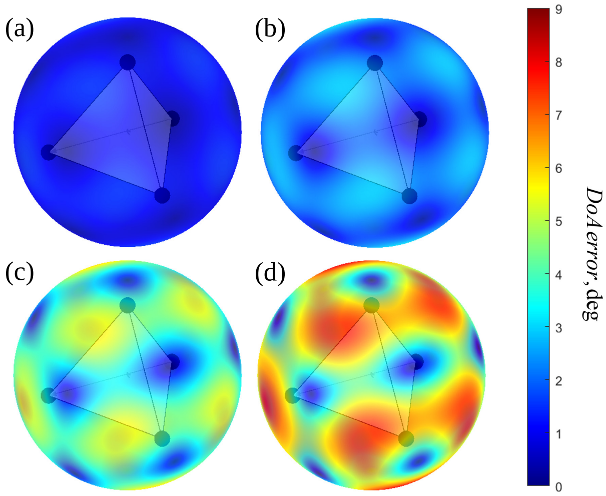

Figure 4. Spatial bias error of the sound intensimetry using tetrahedron microphone array for (a) kd=0.5, (b) kd=1.0, (c) kd=2.0, (d) kd=3.0.

source for the low Helmholtz numbers [12, 13]. However, spatially irregular array directional response appears at the high Helmholtz numbers due to the finite difference error for particle velocity approximation, which makes a large spatial bias error in the localization result [14].
Figure 4 shows how the spatial bias error affects to the localization result when using the tetrahedron microphone array. One can see that a large error occurs at a high *kd* range, where *k* is the wave number and *d* is the microphone spacing. The frequency range of interest is limited to the low *kd* range for the typical sound intensimetry due to such an error, however, one can expand the frequency range of interest by applying the error compensation methods or phase linearization algorithms [14, 15, 16]. The one-dimensional (1D) active intensity vector in Cartesian coordinates, which is calculated from the sound pressure measured from two microphones, can be expressed as [14]

$$\mathbf{I}_{ij}(\omega)=\begin{bmatrix}\mathrm{I}_{x,ij}\\ \mathrm{I}_{y,ij}\\ \mathrm{I}_{z,ij}\end{bmatrix}=\begin{bmatrix}\chi_{x,ij}\\ \chi_{y,ij}\\ \chi_{z,ij}\end{bmatrix}\frac{\mathrm{Im}\left(\hat{\mathbf{G}}_{ij}\right)}{2\rho_0\omega d_{ij}}, \tag{9}$$

where $\mathbf{I}_{ij}$ is the estimated 1D active intensity vector calculated from the *i*-th, *j*-th microphones, $\mathrm{I}_{x,ij},\mathrm{I}_{y,ij},\mathrm{I}_{z,ij}$ means 1D intensity vector component of the x, y, z-axis, respectively, $\chi_{x,ij},\chi_{y,ij},\chi_{z,ij}$ the coefficient of the position vector of microphones, $\mathrm{Im}\left(\hat{\mathbf{G}}_{ij}\right)$ the imaginary part of cross power spectral density function, ρ_0 the air density, and d_{ij} spacing between the microphones.
The three-dimensional (3D) active sound intensity vector can be calculated as the sum of the 1D vectors. Therefore, the 3D intensity vector estimated from the acoustic center of the microphone array, which consists of *N* microphones, can be written as

$$\mathbf{I}(\omega)=\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\mathbf{I}_{ij}(\omega)=\begin{bmatrix}\mathrm{I}_x\\ \mathrm{I}_y\\ \mathrm{I}_z\end{bmatrix}, \tag{10}$$

where $\mathrm{I}_x,\mathrm{I}_y,\mathrm{I}_z$ is the 3D sound intensity vector component of x, y, z-axis, respectively. The estimated azimuth angle and elevation angle by using sound intensimetry can be calculated by

$$\phi_{3DAI}=\tan^{-1}\left(\frac{\mathrm{I}_y}{\mathrm{I}_x}\right),\ \theta_{3DAI}=\tan^{-1}\left(\frac{\mathrm{I}_z}{\sqrt{\mathrm{I}_x^2+\mathrm{I}_y^2}}\right), \tag{11a,b}$$

where ϕ_{3DAI}, θ_{3DAI} denotes the azimuth angle and elevation angle estimated by sound intensimetry, respectively.

2.3 Localization test

The localization test is conducted using the simulated moving sound source signal, and the test condition is shown in Fig. 5. A tetrahedral-shaped microphone array configured by 4 microphones is used for the test. The size of the microphone array and latency are important factors to implement the actual localization system. Therefore, the localization results are compared for sampling rates and array sizes. The test conditions and parameters are indicated in Table 1.

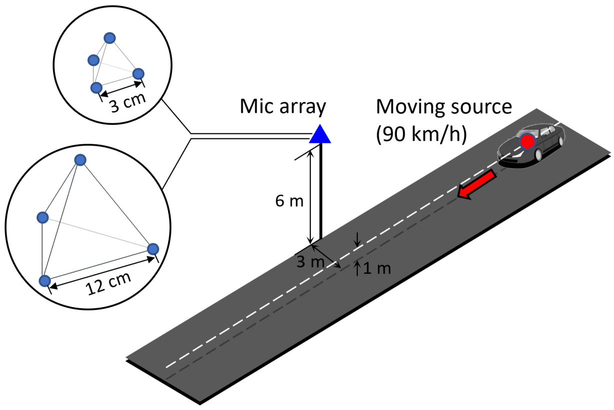

Figure 5. Illustration of the test condition.

Table 1. Test conditions and parameters

Microphone array shape	Tetrahedron
Number of microphones	4
Microphone spacing, m	0.03, 0.12
Signal to noise ratio, dB	>100
Total measurement time, s	8
Time window size, s	0.1
Sampling rate, kHz	6, 100
Moving source velocity, km/h	90 (constant)
Road surface to source distance, m	1
Nearest distance between the source and array, m	5.8 (@4 s)

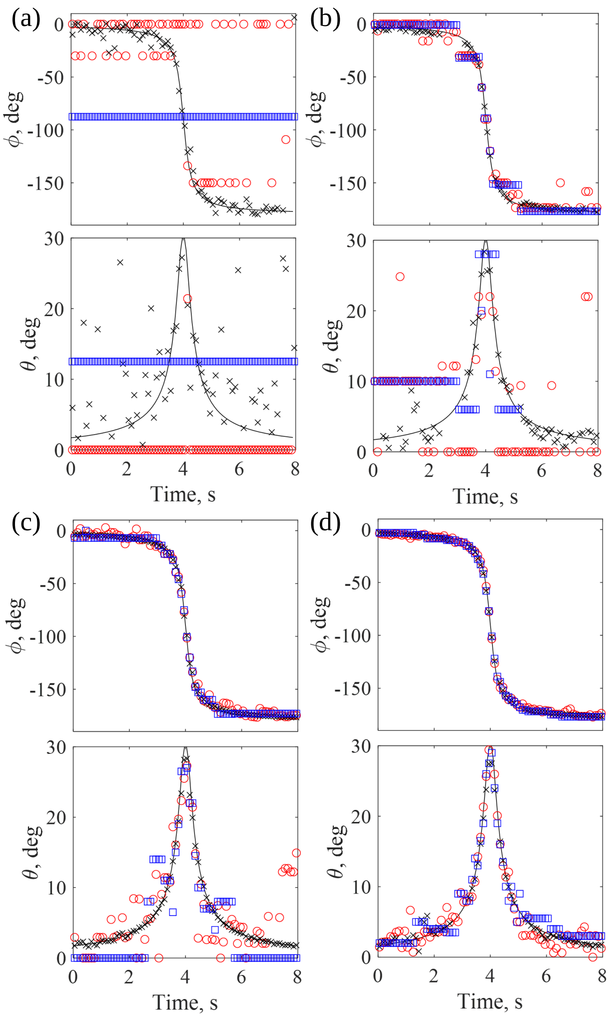

Figure 6. Localization test results for a moving sound source using beamforming, TDoA, and sound intensimetry for different sampling rates and array sizes. (a) f_s=6 kHz, d=3 cm, (b) f_s=6 kHz, d=12 cm, (c) f_s=100 kHz, d=3 cm, (d) f_s=100 kHz, d=12 cm. ○: TDoA, □: DAS, ×: 3DAI, _: real path of the moving source.

The localization test results are shown in Figs. 6-7. The microphone arrays that have microphone spacing of 3 cm and 12 cm are used, and the estimated results according to the localization algorithm are compared when the sampling rate is 6 kHz and 100 kHz, respectively. The localization error and latency are compared. The RMSE and mean latency are defined as

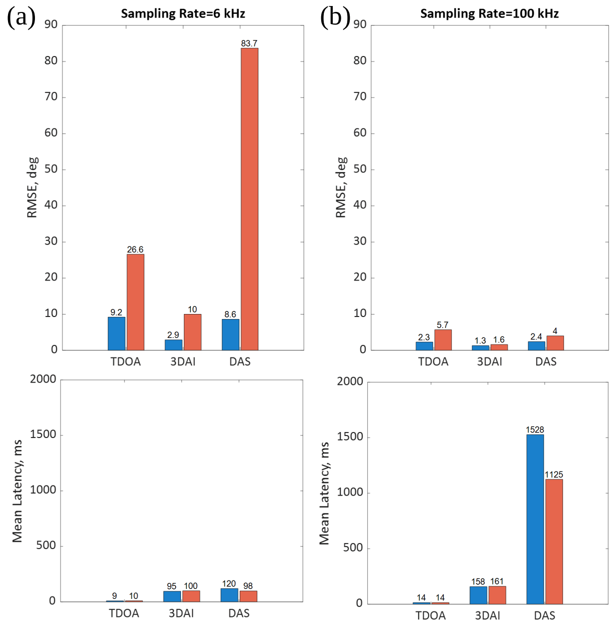

Figure 7. Localization test results for a moving sound source using beamforming, TDoA, and sound intensimetry for different sampling rates and array sizes. (a) Test result when the sampling rate is f_s=6 kHz, (b) f_s=100 kHz. ▬, d=12 cm; ▬, d=3 cm.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\varepsilon(i)}, \tag{12}$$

$$Mean\ Latency = \frac{1}{n}\sum_{i=1}^{n} l(i), \tag{13}$$

where $\varepsilon(i)$, $l(i)$ are DoA estimation error and latency for i^{th} time segment, i.e. n=80 when 8 s of total measurement time with 0.1 s of the time window size. Latency is corresponding to the calculated time using MATLAB via a desktop with a 3.6 GHz CPU and 32 GB RAM.

For the same module size, the sound intensimetry has high accuracy for the source localization, however, it takes more than 10 times the latency compared to the TDoA method. Notwithstanding, the intensity method can be adopted for real-time localization because the latency is near to the measurement time window size.

TDoA method is the fastest one for source localization, so one can be useful for making a node-based module for source localization with a compact integrated system because such a system needs effectively share the memory. However, the array size should be large to increase the localization accuracy for the low frequencies because the estimation error in the TDoA method increases as the Helmholtz number decreases.

Beamforming is advantageous for sound source visualization because it can calculate the spatial response in all directions. Also, one can simply adopt the machine learning algorithm because of the characteristic of the output dataset of beamforming. However, it takes a lot of time for calculation when the grid searching area is too large. Also, one can have poor spatial resolution when the sampling rate is low. One can say that beamforming can be greatly affected by hardware performance and localization parameters.

3. CONCLUSIONS

The sound source localization for a moving source is conducted using a simulated source signal. Beamforming, TDoA, and sound intensimetry are used to estimate a fastly moving source on the paved road. The test result shows that the intensity method has a small estimation error when using a small-sized array having low Helmholtz numbers, the TDoA method has a low computational cost, and the beamforming method is advantageous for visualizing the sound source location. One can think that the sound source localization algorithm needs to be selected in consideration of the frequency range of interest and the specification of the hardware system to be implemented.

4. ACKNOWLEDGEMENTS

This work was partially supported by the research project of KRISS (KRISS-2023-GP2023-0002, KRISS-2023-GP2023-0004-05).

5. REFERENCES

[1] M. Basner, W. Babisch, A. Davis, M. Brink, C. Clark, S. Janssen, S. Stansfeld, "Auditory and non-auditory effects of noise on health," *The Lancet*, 383, 9925, pp. 1325-1332, 2014.

[2] K. Paunović, B. Jakovljević, R. Mirčić, D. Pajić, M. Konatarević, "*The new soundscape after the Covid-19 lockdown*," *in Proc. Inter-noise 2022*, (Glasgow, Scotland), pp. 1374-1382, 2022.

[3] L. Maffei, M. Masullo, "Electric vehicles and urban noise control policies," *Archives of Acoustics*, 39, pp. 333-341, 2014.

[4] H. Bian, Q. Tan, S. Zhong, X. Zhang, "Assessment of UAM and drone noise impact on the environment based on virtual flights," *Aerosp. Sci. Technol.*, 118, pp. 1-6, 2021.

[5] Atkins, Jacobs, *Roadside vehicle noise measurement: study, enforcement and technology*, Dept. Transport, UK, 2019.

[6] F. Meng, Y. Li, B. Masiero, M. Vorländer, "Signal reconstruction of fast moving sound sources using compressive beamforming," *Appl. Acoust.*, 150, pp. 236-245, 2019.

[7] A sound source of vehicle sound, https://youtu.be/A6jGWKnQ1kE

[8] M. Brandstein, D. Ward, *Microphone Arrays*, Springer, 2001.

[9] G. DeMuth, "Frequency domain beamforming techniques," in *Proc. IEEE ICASSP*, (Hartford, USA), pp. 713-715, 1977.

[10] C. Knapp, G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust.*, 24, pp. 320-327, 1976.

[11] I.-J. Jung, J.-G. Ih, "Comparison of the sound source localization methods appropriate for a compact microphone array," *J. Acoust. Soc. Kr.*, 31, pp. 47-56, 2020.

[12] J.-C. Pascal, J.-F. Li, "A systematic method to obtain 3D finite-difference formulations for acoustic intensity and other energy quantities," *J. Sound Vib.*, 310, pp. 1093-1111, 2008.

[13] I.-J. Jung, J.-G. Ih, "Design of a compact omnidirectional sound camera using the three-dimensional acoustic intensimetry," *Mech. Syst. Signal Process.*, 172, pp. 1-13, 2022.

[14] I.-J. Jung, J.-G. Ih, "Combined microphone array for precise localization of sound source using the acoustic intensimetry," *Mech. Syst. Signal Process.*, 160, pp. 1-18, 2021.

[15] D.C. Thomas, B.Y. Christensen, K.L. Gee, "Phase and amplitude gradient method for the estimation of acoustic vector quantities," *J. Acoust. Soc. Am.*, 137, pp. 3366-3376, 2015.

[16] A. H. Moore, C. Evers, P. A. Naylor, D. L. Alon, B. Rafaely, "Direction of arrival estimation using pseudo-intensity vectors with direct-path dominance test," *in Proc. the 23rd European Signal Process.*, (Nice, France), pp. 2341-2345, 2015.