drones

MDPI

Article

Proposal of Practical Sound Source Localization Method Using Histogram and Frequency Information of Spatial Spectrum for Drone Audition

Kotaro Hoshiba *, Izumi Komatsuzaki and Nobuyuki Iwatsuki

Department of Mechanical Engineering, School of Engineering, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan
* Correspondence: hoshiba@rmsv.mech.e.titech.ac.jp

Abstract: A technology to search for victims in disaster areas by localizing human-related sound sources, such as voices and emergency whistles, using a drone-embedded microphone array was researched. One of the challenges is the development of sound source localization methods. Such a sound-based search method requires a high resolution, a high tolerance for quickly changing dynamic ego-noise, a large search range, high real-time performance, and high versatility. In this paper, we propose a novel sound source localization method based on multiple signal classification for victim search using a drone-embedded microphone array to satisfy these requirements. In the proposed method, the ego-noise and target sound components are extracted using the histogram information of the three-dimensional spatial spectrum (azimuth, elevation, and frequency) at the current time, and they are separated using continuity. The direction of arrival of the target sound is estimated from the separated target sound component. Since this method is processed with only simple calculations and does not use previous information, all requirements can be satisfied simultaneously. Evaluation experiments using recorded sound in a real outdoor environment show that the localization performance of the proposed method was higher than that of the existing and previously proposed methods, indicating the usefulness of the proposed method.

Keywords: drone; drone audition; search and rescue; sound source localization; microphone array; multiple signal classification; spatial spectrum; histogram

check for updates

Citation: Hoshiba, K.; Komatsuzaki, I.; Iwatsuki, N. Proposal of Practical Sound Source Localization Method Using Histogram and Frequency Information of Spatial Spectrum for Drone Audition. *Drones* **2024**, *8*, 159. https://doi.org/10.3390/drones8040159

Academic Editor: Sanjay Sharma

Received: 25 March 2024
Revised: 10 April 2024
Accepted: 16 April 2024
Published: 18 April 2024

Correction Statement: This article has been republished with a minor change. The change does not affect the scientific content of the article and further details are available within the backmatter of the website version of this article.

1. Introduction

Rescuing the victims of natural disasters such as earthquakes is one of the most important tasks. The 2023 Turkey-Syria earthquake killed more than fifty thousand people [1,2]. The key factor in search and rescue operations for victims is promptness. It is said that survival rates decrease beyond 72 h after a disaster, which is known as the "Golden 72 h". In recent years, various disaster robots have been developed for search and rescue operations [3]. In particular, drones, which can move quickly regardless of ground conditions, are attracting attention for prompt search and rescue operations, and search and rescue methods that utilize the characteristics of drones are being researched. Many of these methods are vision-based [4,5]. However, vision-based methods face difficulty in searches under poor lighting conditions or when victims are buried in rubble. In fact, in the 1995 Great Hanshin Earthquake, 77% of the fatalities were crushed victims who were buried in collapsed buildings and other structures [6], and the development of technology to search for victims who cannot be searched for with vision is needed.

Therefore, instead of visual information, using auditory information as an acoustic signal obtained from a drone-embedded microphone was considered. The research on handling acoustic signals with a drone-embedded microphone is called "Drone Audition", and various studies on drone audition have been reported. In drone audition research, sound source localization (SSL) and sound source separation/enhancement (SSS/SSE) are often addressed. For example, Go et al. attached four microphone arrays to the underside of a drone and localized the source of impulse sounds using beamforming [7]. Wang et al. mounted a microphone array to the top of a drone and proposed a deep neural network (DNN)-based speech enhancement method [8]. Strauss et al. produced and published a dataset of acoustic signals recorded via a drone-embedded microphone array to advance drone audition research [9].

As part of research on drone audition, we investigated a technology to search for victims by localizing human-related sound sources (target sound sources), such as voices and emergency whistles, using a drone-embedded microphone array. This technology localizes the direction of arrival (DOA) of the target sound based on the time difference of arrival (TDOA) from the target to each microphone, and it searches for the location of the victim. Using acoustic signals, searching operations can be carried out even in poor lighting conditions or when victims are buried in rubble, which is a weakness of vision-based methods. One of the challenges of SSL using drone-embedded microphone arrays is achieving high SSL performance. Some approaches have been studied to overcome this challenge. The first is the development of microphone arrays. To increase the noise tolerance for SSL, microphone arrays consisting of different types and arrangements of microphones have been proposed [7,10–12]. The second is the development of SSL methods. To suppress the decline of localization performance due to the ego-noise of drones, low-noise propellers have been developed [13–15]. The third is the development of high-performance SSL methods. As described below, to improve various performances such as noise tolerance, real-time performance, and so on, SSL methods have been proposed.

In this paper, we addressed to develop a novel SSL method for practical use. The performance requirements for SSL methods in drone audition are as follows.

1. High resolution: High-resolution SSL of the direction of the target sound source is required to accurately locate a victim.
2. High noise tolerance: Because drones generate a large amount of ego-noise, high noise tolerance is required. In addition, ego-noise changes dynamically over time, so tolerance for dynamic noise is also required.
3. Large search range: For prompt search operations, the range that can be localized in a single measurement must be large.
4. High real-time performance: Real-time performance is important for prompt search operations. For high real-time performance, the computational cost must be low.
5. High versatility: The method must be able to localize in any situation and with any drone.

To locate victims using acoustic signals, algorithms based on multiple signal classification (MUSIC) A common method is the SEVD-MUSIC (MUSIC based on Standard EigenValue Decomposition) [16], which is the original MUSIC method. The SEVD-MUSIC algorithm does not have noise suppression calculations, so it has a low noise tolerance. However, it achieves high real-time performance and versatility. To ensure noise tolerance, Nakamura et al. proposed GEVD-MUSIC (MUSIC based on Generalized EigenValue Decomposition) [17]. The GEVD-MUSIC models pre-records ego-noise and whiten noise through generalized eigenvalue decomposition. However, its tolerance for dynamic noise is low, and its computational cost is high. It also has low versatility because the ego-noise of the used drone must be recorded beforehand. To reduce the computational cost of the GEVD-MUSIC, Nakamura et al. proposed GSVD-MUSIC (MUSIC based on Generalized Singular Value Decomposition) [18] using generalized singular value decomposition instead of generalized eigenvalue decomposition. However, the computational cost is still high for real-time search operations. To improve the tolerance for dynamic noise and the versatility of GEVD-MUSIC and GSVD-MUSIC, Okutani et al. and Ohata et al. also proposed iGEVD-MUSIC (incremental GEVD-MUSIC) [19] and iGSVD-MUSIC (incremental GSVD-MUSIC) [20], which model noise by assuming the recorded sound a few seconds before the current time to be noise. These methods are able to cope with slowly changing dynamic noise. However, the ego-noise of a drone often changes quickly, and these methods cannot cope with this. As another approach to improving noise tolerance and reducing computational costs, Hoshiba et al. proposed the angle-limited SEVD-MUSIC, for which the DOA of ego-noise is specified and excluded from the SEVD-MUSIC search range [21]. Although this method ensures the tolerance for static noise and real-time performance, it could not cope with dynamic noise or different drones' ego-noise with different noise characteristics because the excluded range is a fixed value. It also reduces the search range. Hoshiba et al. also proposed AFRF-MUSIC (MUSIC with Active Frequency Range Filtering), which dynamically creates a frequency filter by simply using the four arithmetic operations to eliminate the effect of ego-noise [22]. This method has a low computational cost and high versatility, and it can cope with slowly changing dynamic noise. However, similar to iGEVD-MUSIC and iGSVD-MUSIC, it cannot cope with quickly changing noise because the filter is created from the recorded sound a few seconds before the current time. The characteristics of each method are shown in Table 1.

Table 1. Pros and cons of MUSIC-based sound source localization methods. ○ indicates high performance, and × indicates poor performance for each evaluation item.

Method	Tolerance for Static Noise	Tolerance for Dynamic Noise (Slowly Changing)	Tolerance for Dynamic Noise (Quickly Changing)	Search Range	Real-Time Performance	Versatility
SEVD-MUSIC	×	×	×	○	○	○
GEVD-MUSIC GSVD-MUSIC	○	×	×	○	×	×
iGEVD-MUSIC iGSVD-MUSIC	○	○	×	○	×	○
Angle-limited SEVD-MUSIC	○	×	×	×	○	×
AFRF-MUSIC	○	○	×	○	○	○

The aim of this paper is to develop an SSL method with a drone-embedded microphone array that satisfies all of the following requirements: a tolerance for quickly changing dynamic noise, a large search range, high real-time performance, and high versatility. We propose HIST-MUSIC (MUSIC with HISTogram information), which dynamically estimates the ego-noise and target sound components of the spatial spectrum calculated via SEVD-MUSIC using the histogram information at the current time, and it separates them using continuity [23]. This method theoretically satisfies the above four requirements. However, depending on the DOA of the target sound, the localization performance decreases because the ego-noise and target sound components cannot be accurately separated. Therefore, we propose HIST-MUSIC-3D (three-dimensional HIST-MUSIC), which adds frequency information to the HIST-MUSIC method to accurately separate the ego-noise and target sound components. It is expected to improve localization performance by exploiting the difference in frequency characteristics between ego-noise and the target sound. The remainder of this paper is organized as follows: Section 2 describes the algorithms of HIST-MUSIC and HIST-MUSIC-3D. Section 3 evaluates and discusses the performance of the proposed method. Section 4 presents the conclusion.

2. Methods

This section describes the algorithms of the basic SSL method, SEVD-MUSIC, the proposed method, HIST-MUSIC, and its extended version, HIST-MUSIC-3D.

2.1. SEVD-MUSIC

The M-channel input acoustic signal of the f-th frame is Fourier transformed to $Z(\omega, f)$ via the short-time Fourier transform (STFT). Here, ω is the frequency bin number. Using $Z(\omega, f)$, the correlation matrix $R(\omega, f)$ is defined as follows.

$$R(\omega, f) = \frac{1}{T_R} \sum_{\tau=f}^{f+T_R-1} Z(\omega, \tau) Z^*(\omega, \tau) \tag{1}$$

Here, T_R is the number of frames used to average the correlation matrix, and Z^* is the complex conjugate transpose of Z. SEVD-MUSIC calculates eigenvectors via the standard eigenvalue decomposition of the $R(\omega, f)$.

$$R(\omega, f) = E(\omega, f)\Lambda(\omega, f)E^*(\omega, f) \tag{2}$$

Here, $\Lambda(\omega, f)$ is the matrix with diagonal components that are eigenvalues in descending order, and $E(\omega, f)$ is the matrix containing eigenvectors corresponding to $\Lambda(\omega, f)$. Using E and the transfer function $G(\omega, \psi)$ corresponding to the sound source direction ψ in the drone's coordinates, the spatial spectrum $P(\omega, \psi, f)$ is calculated as follows.

$$P(\omega, \psi, f) = \frac{|G^*(\omega, \psi)G(\omega, \psi)|}{\sum_{m=L+1}^{M} |G^*(\omega, \psi)e_m(\omega, f)|}. \tag{3}$$

Here, L is the number of target sound sources, and e_m is the m-th eigenvector in E. ψ is defined as $\psi = (\theta, \phi)$, which are the azimuth and elevation angles of the target sound direction in the drone's coordinates, respectively. $P(\omega, \psi, f)$ is summed over the ω direction to estimate the target sound direction.

$$\bar{P}(\psi, f) = \sum_{\omega=\omega_L}^{\omega_H} P(\omega, \psi, f) \tag{4}$$

Here, ω_H and ω_L are indices corresponding to the upper and lower limits of the frequency bin used for calculation, respectively. $P(\omega, \psi, f)$ and $\bar{P}(\psi, f)$ represent the power of the sound arriving from the direction ψ at the f-th frame. To detect the target sound direction ψ_{target}, peak detection is performed for $\bar{P}(\psi, f)$ as follows.

$$\psi_{target}(f) = \underset{\psi}{\mathrm{argmax}}(\bar{P}(\psi, f)) \tag{5}$$

2.2. HIST-MUSIC

In order to improve the noise tolerance, search range, real-time performance, and versatility of SSL compared to the MUSIC-based methods introduced in Section 1, HIST-MUSIC was proposed. This method excludes the ego-noise component using only sound recorded at the current time, not pre-recorded or previously recorded sound, in order to increase the tolerance for quickly changing noise and versatility. It also performs only simple calculations to improve real-time performance. The algorithm is described below.

This method considers the case in which a microphone array is installed outside the rotors, as shown in Figure 1. In this case, ego-noise arrives the microphone array from one direction. An example of the spatial spectrum $\bar{P}$ calculated via SEVD-MUSIC, plotted with the axes given in Figure 2, is shown in Figure 3.The circumferential and radial directions represent the azimuth and elevation angles, respectively. The sound power arriving from each direction is depicted in a color map. The high values on the left side of Figure 3 are the components of ego-noise, and the peak around $(\theta, \phi) = (0$ deg., -45 deg.$)$ is the component of the target sound.

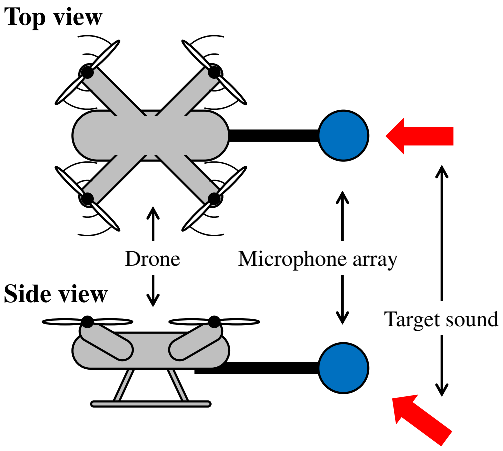


Figure 1. Example of microphone array installation on a drone.

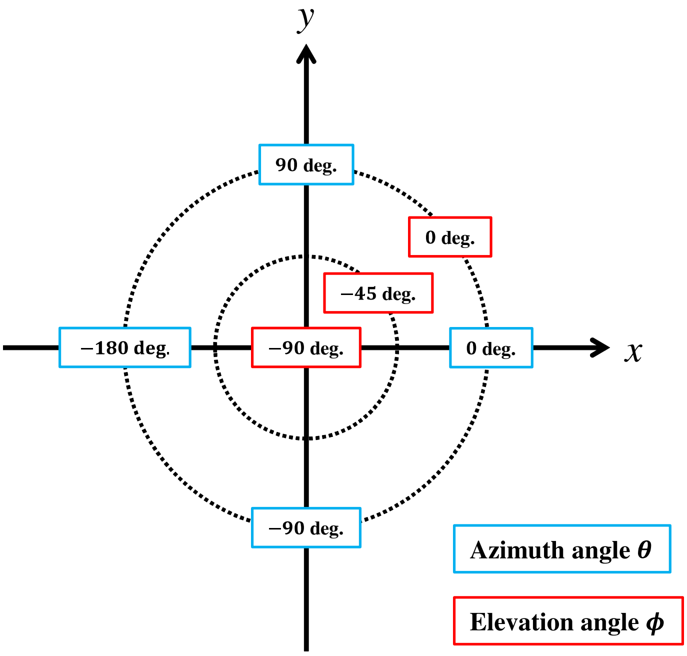


Figure 2. Setting of the coordinate system of the azimuth angle θ and the elevation angle ϕ when plotting spatial spectra.

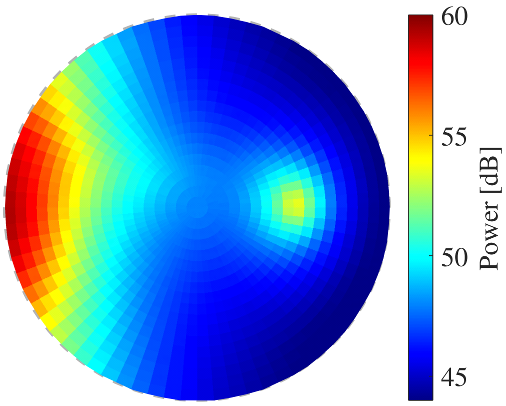


Figure 3. Example of spatial spectrum $\bar{P}$ calculated via SEVD-MUSIC. ψ_{target} is (0 deg., −45 deg.).

If the peak of the target sound component is lower than the value of the ego-noise component, the target sound cannot be detected via Equation (5). To deal with such a case, HIST-MUSIC obtains $H(p, f)$, the histogram of $\bar{P}$ for all ψ calculated via SEVD-MUSIC.

$$H(p, f) = \text{histogram}(\{\bar{P}(\psi, f)\}_{\text{all } \psi}) \tag{6}$$

Here, p is the power bin of the spatial spectrum. A schematic diagram of the obtained H is shown in Figure 4. The horizontal axis represents the power bin p, and the vertical axis represents the probability density. The solid black line is the obtained histogram (H), the red dotted line is the histogram of only the target sound component (H_t), and the blue dotted line is the histogram of only the ego-noise component (H_n). The obtained histogram H is, thus, represented as the sum of the histograms of the target sound component H_t and the ego-noise component H_n. This representation was inspired by the work of Mori et al. [24], who represented the echo images of the body as the sum of the histograms of normal and diseased areas. Since most of the spatial spectrum consists of an ego-noise component other than the target sound, the maximum peak of H is due to ego-noise. If a target sound has a lower power than the maximum peak of H, a target sound cannot be detected at all. Therefore, the target sound is assumed to have a higher power than the maximum peak of H. To detect target sounds using Equation (5), the ego-noise component above a reference value lower than the maximum value of the target sound component p_{tmax} (the cyan area in Figure 4) must be excluded. Then, when focusing on the right side of the maximum peak of H, it is found that the target sound component generates an inflection point, p_i. p_i is always lower than p_{tmax}, so we adopt this point as the reference value.

$$p_i(f) = p \mid \frac{d^2}{dp^2} H(p, f) = 0 \tag{7}$$

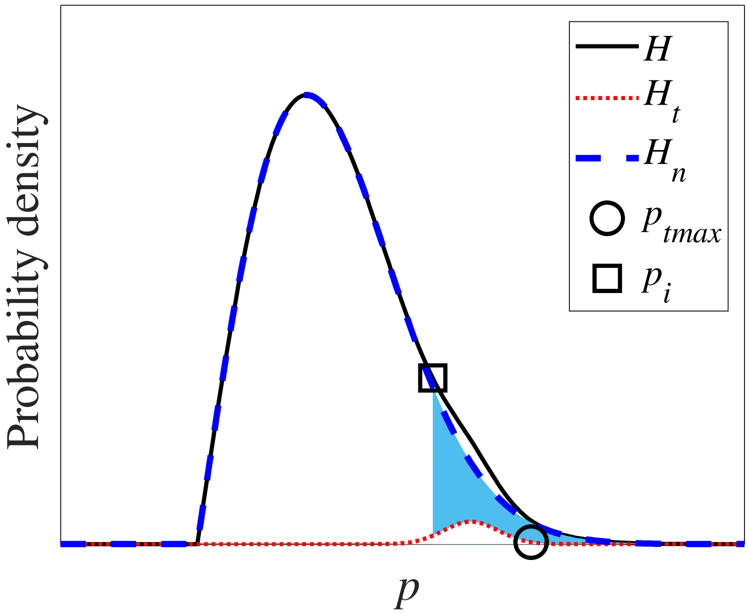


Figure 4. Schematic diagram of histograms. The solid black line is the obtained histogram, H, the red dotted line is the histogram of the target sound component, H_t, the blue dotted line is the histogram of the ego-noise component, H_n, the circle mark is the maximum value of the target sound component, p_{tmax}, and the square mark is the inflection point of H, p_i.

An example of the results of calculating p_i using actual experimental data is shown in Figure 5. It can be seen that a p_i lower than p_{tmax} can be obtained.

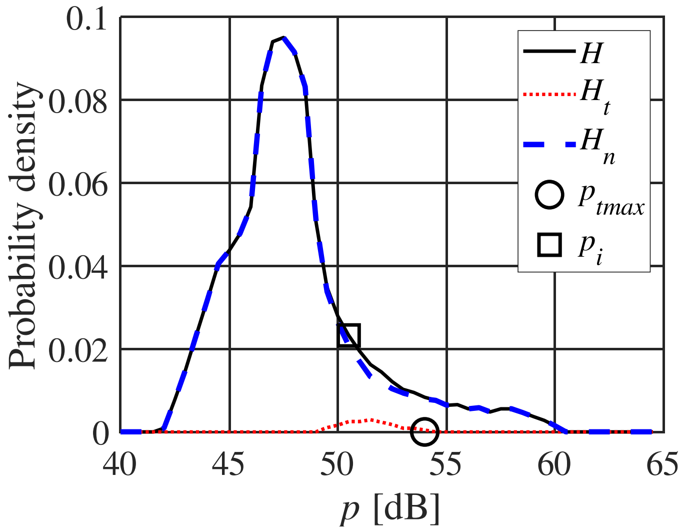


Figure 5. Histograms of actual experimental spatial spectrum. The solid black line is the obtained histogram, H, the red dotted line is the histogram of the target sound component, H_t, the blue dotted line is the histogram of the ego-noise component, H_n, the circle mark is the maximum value of the target sound component, p_{tmax}, and the square mark is the inflection point of H, p_i.

Next, the algorithm to exclude only the ego-noise component using p_i is presented. Using p_i, the components with a power higher than p_i, $\bar{P}_{cut}$, are extracted from $\bar{P}$.

$$\bar{P}_{cut}(\psi, f) = \bar{P}(\psi, f) \mid \bar{P} > p_i \tag{8}$$

Figure 6 shows the extracted result of P_{cut}. It can be seen that the ego-noise and target sound components are extracted. If $\bar{P}_{cut}$ is simply determined as ego-noise, then the target sound component is also determined as an ego-noise component. Then, the ego-noise component and the target sound component are separated from $\bar{P}_{cut}$. Considering that ego-noise arrives the microphone array from one direction, the reference direction ψ_0 is set in the direction from which the ego-noise arrives. In $\bar{P}_{cut}$, a continuous component including ψ_0 is determined as the ego-noise component $\bar{P}_{noise}$, and the other components are determined as the target component $\bar{P}_{target}$.

$$\bar{P}_{noise}(\psi, f) = \bar{P}_{cut}(\psi, f) \mid \psi \ni \psi_0 \tag{9}$$

$$\bar{P}_{target}(\psi, f) = \bar{P}_{cut}(\psi, f) \mid \psi \not\ni \psi_0 \tag{10}$$

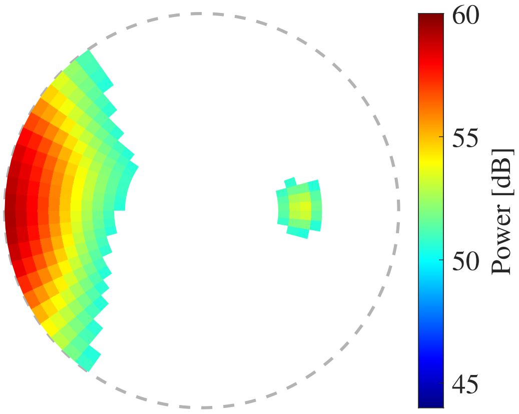


Figure 6. Components with a higher power than the reference value p_i in the spatial spectrum shown in Figure 3, $\bar{P}_{cut}$.

Figures 7 and 8 show the separated results of $\bar{P}_{noise}$ and $\bar{P}_{target}$ from Figure 6. In this case, ψ_0 is set to $(-180$ deg., 0 deg.$)$.

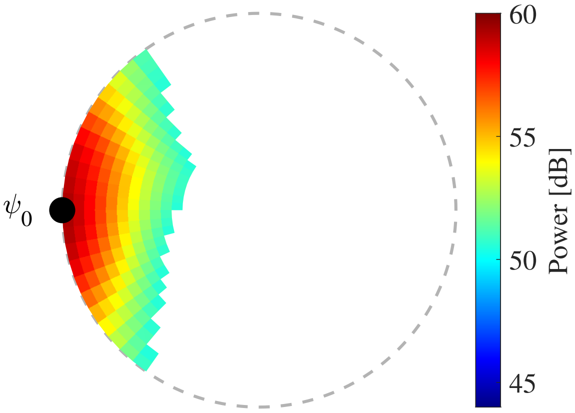


Figure 7. Separated result of the ego-noise component $\bar{P}_{noise}$ from Figure 6. ψ_0 is the reference direction.

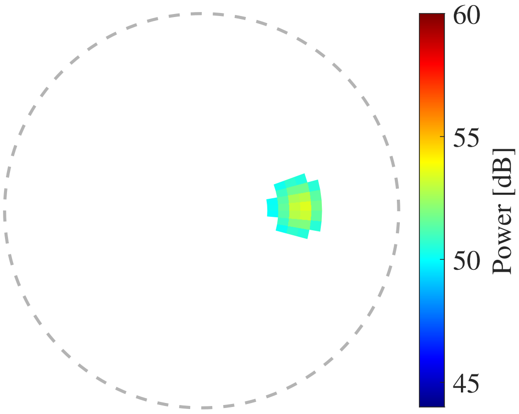


Figure 8. Separated result of the target component $\bar{P}_{target}$ from Figure 6.

Now, only the target sound component can be extracted. Using $\bar{P}_{target}$, the target sound direction can be detected via Equation (5) as follows.

$$\psi_{target}(f) = \underset{\psi}{\text{argmax}}(\bar{P}_{target}(\psi, f)) \tag{11}$$

2.3. HIST-MUSIC-3D

In previous work, we reported that the performance of HIST-MUSIC, introduced in Section 2.2, decreases when the target sound direction is close to the ego-noise direction [23]. Figure 9 shows an example of the spatial spectrum $\bar{P}$ when $\psi_{target} = (-90$ deg., -45 deg.$)$. If this spatial spectrum is processed via HIST-MUSIC, because the directions of the target sound and ego-noise are so close, the target sound component is determined as the ego-noise component, as shown in Figure 10. As a result, the target sound direction cannot be detected.

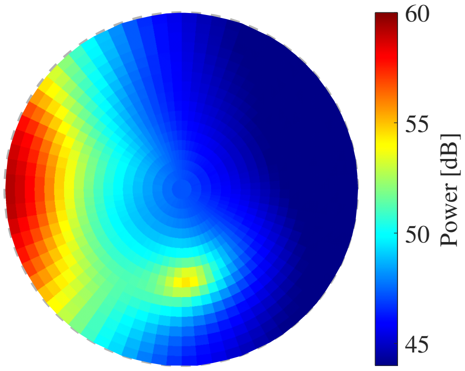


Figure 9. Example of spatial spectrum $\bar{P}$ calculated via SEVD-MUSIC. ψ_{target} is $(-90$ deg., -45 deg.$)$.

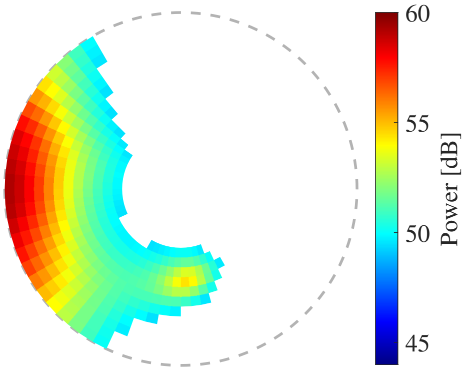


Figure 10. Separated result of the ego-noise component $\bar{P}_{noise}$ from Figure 9 using HIST-MUSIC.

To solve this problem, we propose HIST-MUSIC-3D, which can perform sound source detection even when the direction of the target sound and ego-noise are close to each other by using the difference in frequency characteristics between the target sound and ego-noise and determining the target sound and ego-noise through considering the frequency information. The algorithm is described below.

Instead of $\bar{P}$, HIST-MUSIC-3D uses P in Equation (4), which is a spatial spectrum, before summing in the frequency direction. This method obtains $H_{3D}(p, f)$, the histogram of P for all ω and ψ, as in Equation (6).

$$H_{3D}(p, f) = \text{histogram}(\{P(\omega, \psi, f)\}_{\text{all } \omega,\psi}) \tag{12}$$

Then, the inflection point p_{i3D} of the histogram H_{3D} is determined as the reference value, as in Equation (7).

$$p_{i3D}(f) = p|\frac{d^2}{dp^2}H_{3D}(p, f) = 0 \tag{13}$$

Using p_{i3D}, the components with power higher than p_{i3D} are extracted from P.

$$P_{cut}(\omega, \psi, f) = P(\omega, \psi, f)|P > p_{i3D} \tag{14}$$

P_{cut} in the case of Figure 9, plotted with the axes given in Figure 11, is shown in Figure 12.The sound power arriving from each direction and each frequency is depicted with a color of point. As shown in the figure, the spatial spectrum is obtained in three dimensions (azimuth θ, elevation ϕ, and frequency ω). Due to the difference in frequency characteristics, the components extracted via p_{i3D} can be divided into a large cluster (ego-noise component) and a small cluster (target sound component) when viewed as three-dimensional data.

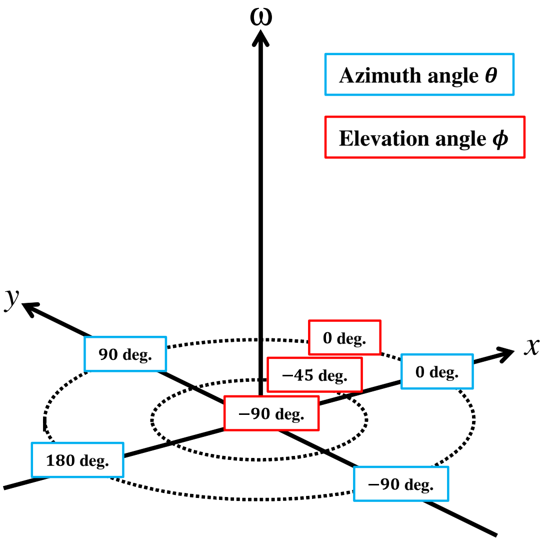


Figure 11. Setting of the coordinate system of azimuth angle θ, elevation angle ϕ, and frequency ω when plotting three-dimensional spatial spectra.

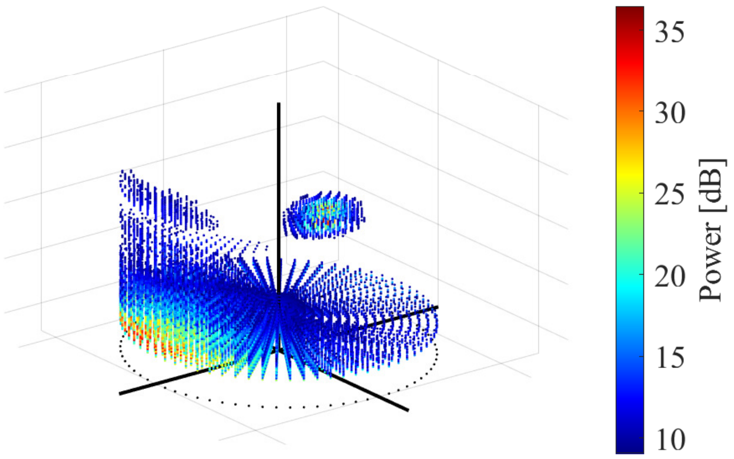


Figure 12. Obtained P_{cut} in three dimensions. ψ_{target} is $(-90 \text{ deg.}, -45 \text{ deg.})$.

Using the obtained three-dimensional spatial spectrum, the ego-noise and target sound components are separated. As in HIST-MUSIC, the reference direction ψ_0 in which the ego-noise arrives is set. The continuous component including ψ_0 is separated as the ego-noise component, and the other components are separated as the target sound component.

$$P_{noise}(\omega,\psi,f) = P_{cut}(\omega,\psi,f)|\psi \ni \psi_0 \tag{15}$$

$$P_{target}(\omega,\psi,f) = P_{cut}(\omega,\psi,f)|\psi \not\ni \psi_0 \tag{16}$$

The separated results from Figure 12 are shown in Figure 13. As in Figure 7, ψ_0 was set to $(-180 \text{ deg.}, 0 \text{ deg.})$. The red cluster is the separated ego-noise component P_{noise}, and the blue cluster is the target sound component P_{target}. As shown in the figure, even if the ego-noise and target sound components cannot be separated by analyzing the two-dimensional spatial spectrum, they can be separated by analyzing the three-dimensional spatial spectrum. P_{target} is summed over the ω direction to estimate the target sound direction.

$$\bar{P}'_{target}(\psi,f) = \sum_{\omega} P_{target}(\omega,\psi,f) \tag{17}$$

Figure 14 shows the result of $\bar{P}'_{target}$, plotted with the axes given in Figure 2. As shown in the figure, a two-dimensional spatial spectrum of only the target sound component can be obtained in the end.

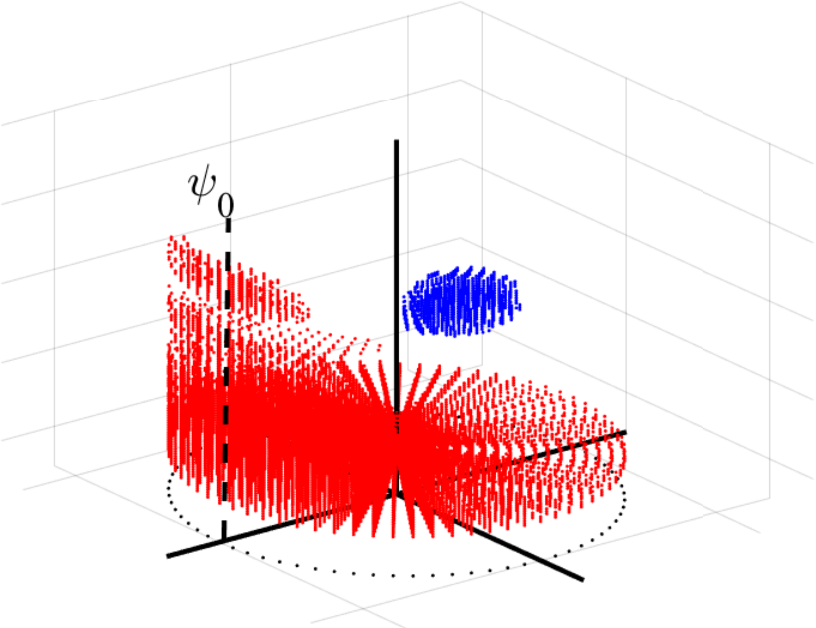


Figure 13. Separated result of the ego-noise component P_{noise} and the target sound component P_{target} from Figure 12. The red cluster is P_{noise}, and the blue cluster is P_{target}. ψ_0 is the reference direction.

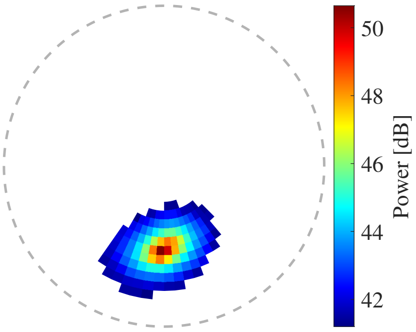


Figure 14. Two-dimensional spatial spectrum $\bar{P}'_{target}$, which is the sum of P_{target} over the ω direction.

Then, to detect the target sound direction ψ_{target}, peak detection is performed for $\bar{P}'_{target}$.

$$\psi_{target} = \underset{\psi}{\operatorname{argmax}}(\bar{P}'_{target}(\psi, f)) \tag{18}$$

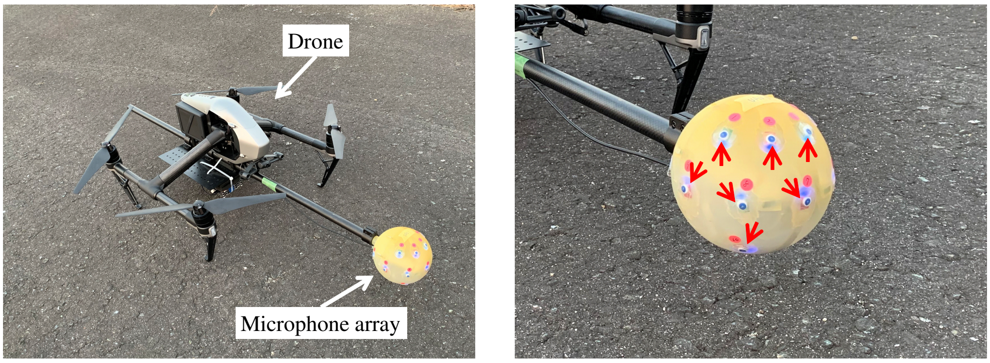
3. Evaluation Experiments

To evaluate the performance of the proposed method, evaluation experiments were performed.

3.1. Experimental Procedure

In the experiments, evaluation signals were created by adding ego-noise recorded in a real outdoor environment and target sounds that arrived from various directions generated via numerical calculation. The performance was evaluated by processing SSL using the evaluation signals.

The ego-noise was recorded using DJI (Shenzhen, China) Inspire 2 with an embedded microphone array, as shown in Figure 15a. The drone and the microphone array were connected with a single pipe. The distance between the center of the drone and the microphone array was 600 mm. As the microphone array, a 16-channel spherical microphone array [3], shown in Figure 15b, was used. The microphones are marked with red arrows in Figure 15b. The microphone array consists of 12-channel MEMS microphones in the lower hemisphere and 4-channel MEMS microphones in the upper hemisphere of a 110-mm-diameter body. The position of each microphone is shown in Figure 16. Using this microphone array, ego-noise was recorded at a sampling frequency of 16 kHz and a quantization bit rate of 24 bits while the drone was hovering at an altitude of 10 m and flying at speeds of 1, 2, and 3 m/s.



(**a**) DJI Inspire 2 with microphone array.

(**b**) 16-channel microphone array. Microphones are marked with red arrows.

Figure 15. Drone and microphone array used in the experiments.

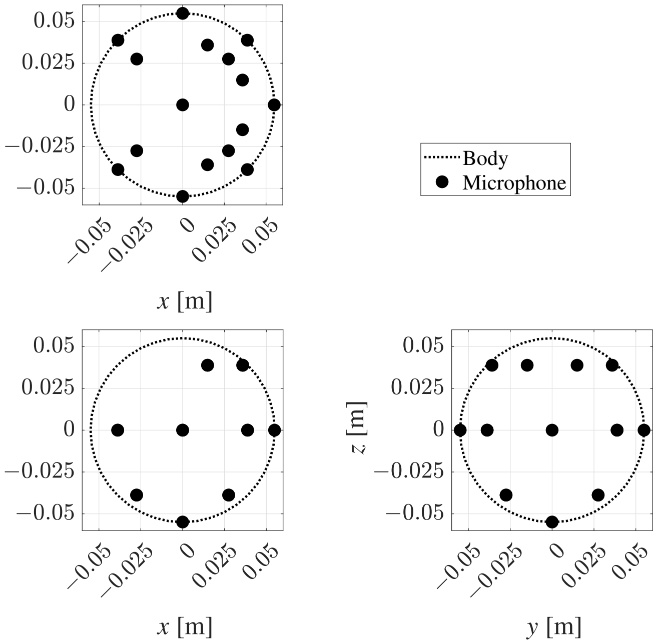
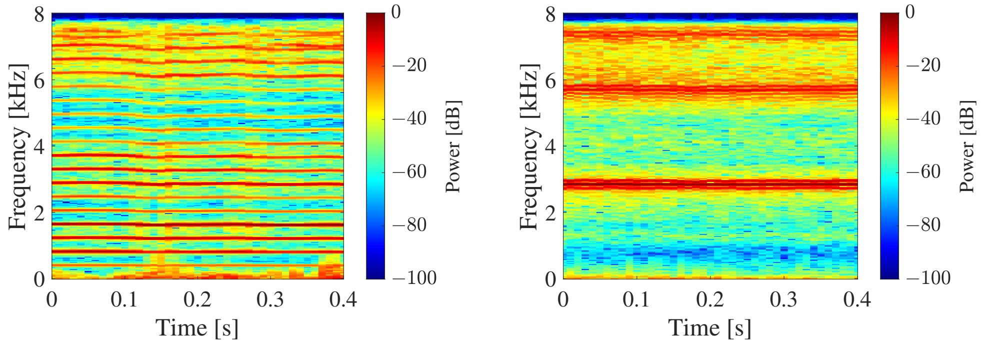


Figure 16. Microphone positions in the microphone array.

Target sounds were generated using a one-channel sample recorded in an outdoor environment. Considering the positions of 16 microphones, numerical calculations were performed to simulate 16-channel acoustic signals when the target sound arrived from an arbitrary direction. The DOA was set to 5 deg. increments within the range of $-180 \sim 180$ deg. for azimuth angle θ and $-90 \sim 0$ deg. for elevation angle ϕ. As target sound samples, voice and whistle sounds were used. Their spectrograms are shown in Figure 17.

Using the obtained ego-noise and the target sound, evaluation signals were created via addition such that the SNR (Signal-to-Noise Ratio) was -20 to 0 dB in 4 dB increments. The evaluation signals were processed through the conventional and proposed SSL methods and evaluated. The methods used in the evaluation are shown in Table 2. As common analysis conditions for these methods, the length of FFT and overlap in STFT were 512 and 352 points, the number of frames used to average the correlation matrix was $T_R = 50$ (that is, a signal with the length of 0.5 s was used in the process for SSL), the number of target sources was $L = 2$, the analyzed frequency range was $\omega_L = 500$ Hz and $\omega_H = 4000$ Hz, and the transfer function, G, used in Equation (3) was derived from geometric calculation. As described in Section 1, GEVD-MUSIC uses pre-recorded ego-noise to whiten noise. For this, the ego-noise of the same drone recorded on a different day was used. iGEVD-MUSIC performed the same whitening by assuming that the recorded sound 2 s before the current time was noise. AFRF-MUSIC creates a frequency filter from the frequency spectrum. The filter range was set to 500 Hz. The angle-limited SEVD-MUSIC limits the search range of the azimuth angle in the spatial spectrum. Angle-limited SEVD-MUSIC (1) was set to $-90 \leq \theta \leq 90$ deg., and angle-limited SEVD-MUSIC (2) was set to $-135 \leq \theta \leq 135$ deg. In HIST-MUSIC and HIST-MUSIC-3D, the reference direction was set to $\psi_0 = (-180 \text{ deg.}, 0 \text{ deg.})$.



(**a**) Voice. (**b**) Whistle.

Figure 17. Spectrograms of target sound samples.

Table 2. SSL methods used for evaluation.

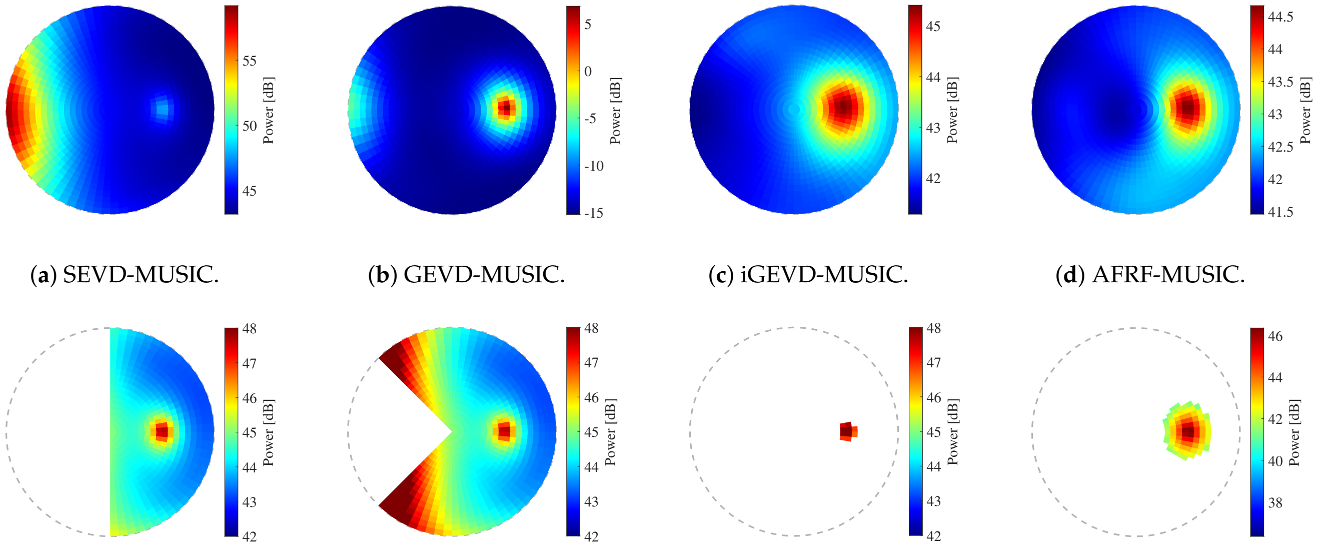
1. SEVD-MUSIC
2. GEVD-MUSIC
3. iGEVD-MUSIC
4. AFRF-MUSIC
5. Angle-limited SEVD-MUSIC (1) (Azimuth range: $-90 \leq \theta \leq 90$ deg.)
6. Angle-limited SEVD-MUSIC (2) (Azimuth range: $-135 \leq \theta \leq 135$ deg.)
7. HIST-MUSIC
8. HIST-MUSIC-3D

3.2. Results

The two-dimensional spatial spectra calculated via the methods shown in Table 2 for various evaluation signals under different conditions are presented.

Figure 18 shows the calculated spatial spectra when the SNR was −16 dB, the target sound was the whistle sound, ψ_{target} was (0 deg., −45 deg.), and the flight speed was 0 m/s (hovering). Figure 19 shows the spatial spectra when the flight speed was 3 m/s. Figure 20 shows the spatial spectra when ψ_{target} was (−135 deg., −45 deg.) and the flight speed was 3 m/s. These figures are plotted with the axes given in Figure 2. The sound power arriving from each direction is depicted in a color map.

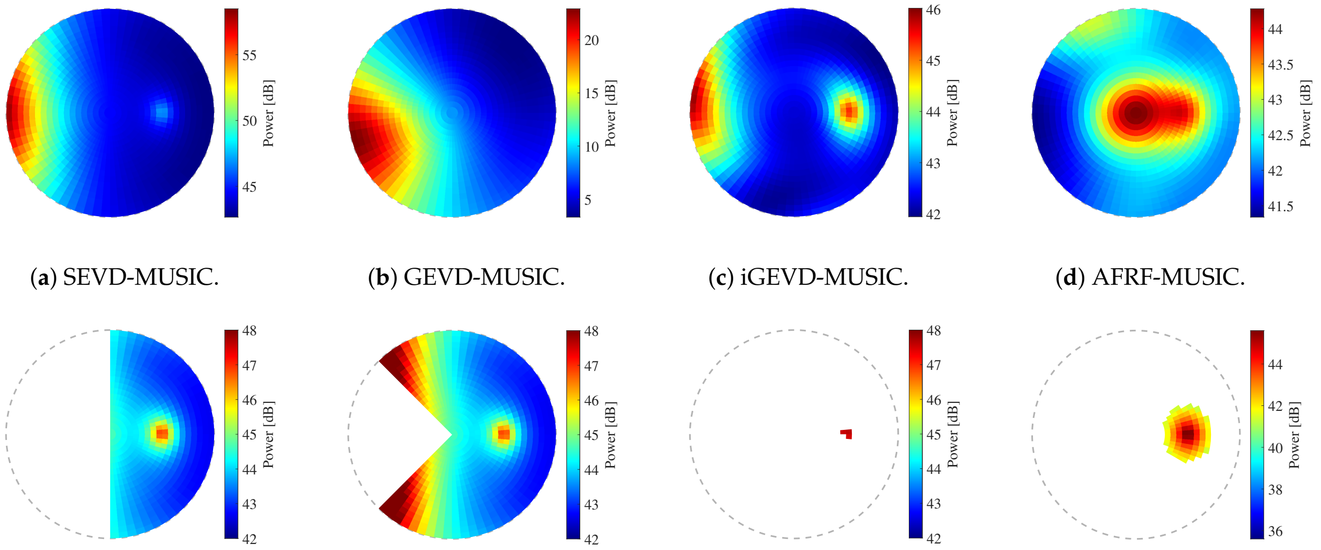
The SEVD-MUSIC results show that the ego-noise component exists around $(\theta, \phi) = (-180 \text{ deg.}, 0 \text{ deg.})$. Although the target sound component exists in the direction away from the ego-noise component, the power of the ego-noise component is larger than the target sound component in all three situations, and the DOA of the target sound cannot be detected through peak detection. The GEVD-MUSIC and AFRF-MUSIC results show that the ego-noise component is removed through each method, and the target sound component has the maximum peak when the drone is hovering, because the changes in the ego-noise over time are small. However, when the drone flies at a speed of 3 m/s, the ego-noise component cannot be removed because the ego-noise changes quickly over time. Therefore, the target sound component cannot be detected. Compared to GEVD-MUSIC and AFRF-MUSIC, the iGEVD-MUSIC results show that the target sound component can be seen relatively well even when flying at 3 m/s; however, the ego-noise component still has the maximum peak. The angle-limited SEVD-MUSIC results show that the target sound components at (0 deg., −45 deg.) can be detected by setting the search range to be narrow. However, the target sound component at (−135 deg., −45 deg.) is not included in the search range and cannot be detected. On the other hand, when the search range is set wider, the ego-noise component is included in the search range, and the detection of the target sound component fails. If the range at which the ego-noise component exists could be known in advance, it would be possible to maximize the search range and improve the tolerance for noise; however, the range of the ego-noise component changes due to flight conditions, as shown in Figures 18f, 19f and 20f, and it is difficult to deal with these changes. These results indicate that there is a trade-off relationship between the tolerance for dynamic noise and the search range, and it is difficult to satisfy both at the same time. The HIST-MUSIC results show that the target sound component is successfully detected at $\psi_{target} = (0 \text{ deg.}, -45 \text{ deg.})$. However, when $\psi_{target} = (-135 \text{ deg.}, -45 \text{ deg.})$, which is close to the ego-noise direction, the target sound component is determined as the ego-noise component, and the detection fails. In the results of the proposed method, HIST-MUSIC-3D, it was found that HIST-MUSIC-3D solves all the above problems and successfully detects the target sound component in all three situations.



(**a**) SEVD-MUSIC. (**b**) GEVD-MUSIC. (**c**) iGEVD-MUSIC. (**d**) AFRF-MUSIC.

(**e**) Angle-limited SEVD-MUSIC (1). (**f**) Angle-limited SEVD-MUSIC (2). (**g**) HIST-MUSIC. (**h**) HIST-MUSIC-3D.

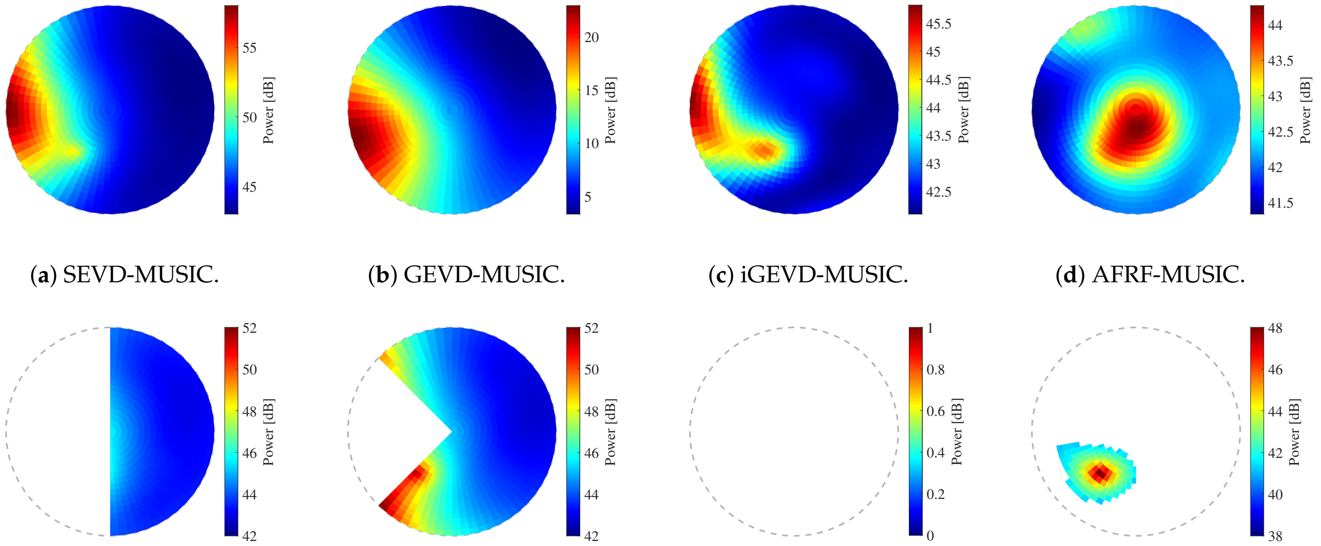
Figure 18. Calculated spatial spectra. SNR: −16 dB, target sound: whistle, ψ_0: (0 deg., −45 deg.), flight speed: hovering.



(**a**) SEVD-MUSIC. (**b**) GEVD-MUSIC. (**c**) iGEVD-MUSIC. (**d**) AFRF-MUSIC.

(**e**) Angle-limited SEVD-MUSIC (1). (**f**) Angle-limited SEVD-MUSIC (2). (**g**) HIST-MUSIC. (**h**) HIST-MUSIC-3D.

Figure 19. Calculated spatial spectra. SNR: −16 dB; target sound: whistle; ψ_0: (0 deg., −45 deg.); flight speed: 3 m/s.



(**a**) SEVD-MUSIC. (**b**) GEVD-MUSIC. (**c**) iGEVD-MUSIC. (**d**) AFRF-MUSIC.

(**e**) Angle-limited SEVD-MUSIC (1). (**f**) Angle-limited SEVD-MUSIC (2). (**g**) HIST-MUSIC. (**h**) HIST-MUSIC-3D.

Figure 20. Calculated spatial spectra. SNR: −16 dB; target sound: whistle; ψ_0: (−135 deg., −45 deg.); flight speed: 3 m/s.

3.3. Discussion

To discuss the performance of the proposed method, the obtained results were evaluated in terms of their noise tolerance, searchable range, real-time performance, and versatility.

3.3.1. Tolerance for Dynamic Noise and Searchable Range

To evaluate the noise tolerance and searchable range, the localization of the target sound was performed using each SSL method for a total of 3,112,800 evaluation signals, including 50 frames (25 s) each of ego-noise while flying at 0 (hovering), 1, 2, and 3 m/s, two target sound samples (voice and whistle), 1297 DOAs, and six SNRs. Noise tolerance was evaluated based on the success rate of localization for all trials as follows.

$$\text{Success rate} = \frac{\text{Success counts}}{\text{All evaluation signals } (3,112,800)} \times 100\ [\%] \tag{19}$$

Localization success was defined as a localization within 5 deg. of the true value in DOA of the target sound. In addition, the success rate was calculated for each DOA, and the percentage of DOAs with a success rate higher than 50% was defined and evaluated as the searchable range rate as follows.

$$\text{Searchable range rate} = \frac{\text{Direction counts (over 50\% success rate)}}{\text{All DOAs } (1,297)} \times 100\ [\%] \tag{20}$$

Figures 21 and 22 show the calculated success rate and searchable range rate. The horizontal axis represents the SNR, and the vertical axis represents the success rate and the searchable range rate, respectively.

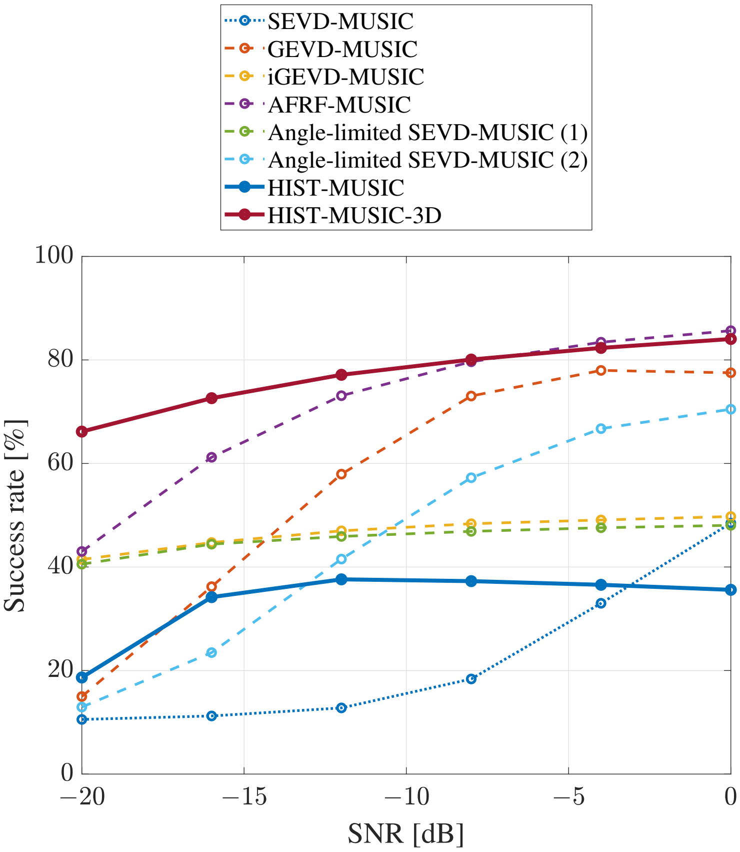


Figure 21. Success rate of SSL for each SSL method.

The success rate and searchable range rate of SEVD-MUSIC were less than 50% because SEVD-MUSIC does not perform for noise suppression processing. GEVD-MUSIC, angle-limited SEVD-MUSIC (2), and AFRF-MUSIC were all high when the SNR was high; however, they also decreased as the SNR decreased. In particular, AFRF-MUSIC had significantly lower searchable ranges at a low SNR. iGEVD-MUSIC and angle-limited SEVD-MUSIC (1) also had low success rates and searchable range rates at all SNRs. These results are due to the insufficient tolerance for quickly changing dynamic ego-noise. These results indicate that existing methods are not able to achieve a high success rate or a large searchable range at low SNRs. In particular, even methods with theoretically large search ranges have small searchable ranges in a real environment. When considering operations in a real environment, the SNR was assumed to be less than −20 dB, and the ego-noise was assumed to change quickly over time. Therefore, the practicality of these methods is insufficient.

HIST-MUSIC was proposed to satisfy the tolerance for quickly changing dynamic noise and a large search range; however, both were less than 40% in a real environment. This is because the target sound component was determined as the ego-noise and excluded from the search range when the DOA of the target sound was close to the ego-noise, as described in Section 2.3. Figure 23a shows the success rate for each set DOA when the SNR was −16 dB. Each set DOA follows the axes given in Figure 2, and the success rate for each set DOA is depicted in a color map. As shown in the figure, the success rate was 0% in the range where ego-noise exists and is close to the ego-noise, which caused the success rate and the searchable range to decrease.

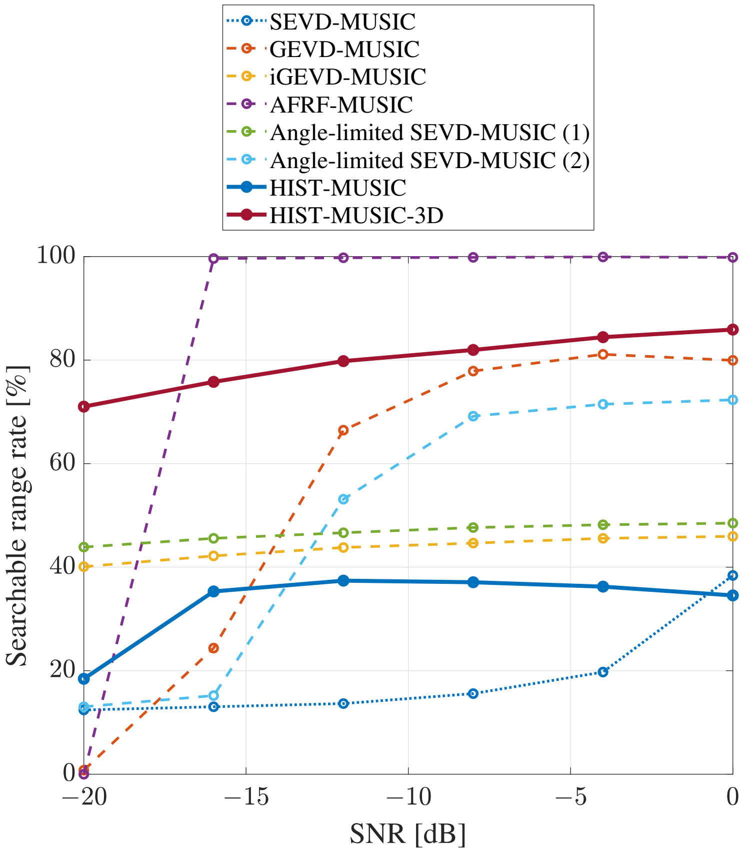
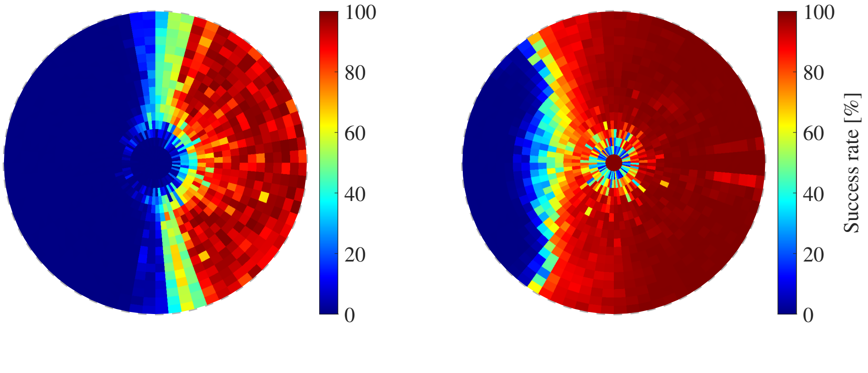


Figure 22. Searchable range rate of SSL for each SSL method.

On the other hand, HIST-MUSIC-3D had a high success rate and a searchable range rate at all SNRs. In particular, they were the highest in all methods at low SNRs. As in HIST-MUSIC, Figure 23b shows the success rate at each set DOA. As shown in the figure, the success rate improved in the range close to the ego-noise, where HIST-MUSIC failed localization. Then, the success rate and the searchable range rate increased.

These results show that HIST-MUSIC-3D can simultaneously satisfy high tolerance for quickly changing dynamic noise and a large search range.



(**a**) HIST-MUSIC. (**b**) HIST-MUSIC-3D.

Figure 23. Success rate for each set DOA at the SNR of −16 dB.

3.3.2. Real-Time Performance

To evaluate real-time performance, the processing time for evaluation signals using each method was measured. A virtual machine with a four-core CPU, a 1.5 GHz CPU clock, and 4 GB of memory was used for processing, assuming a Jetson Nano. The RTF (Real Time Factor) was used as the evaluation index, which was calculated as follows.

$$\mathrm{RTF} = \frac{\text{Processing time}}{\text{Time length of evaluation signal}} \tag{21}$$

That is, if the RTF < 1, processing is completed within the time length of the signal, which means that SSL can be performed in real time without delay accumulation. Two hundred evaluation signals of 0.5 s were processed, and the average value of the RTF was measured.

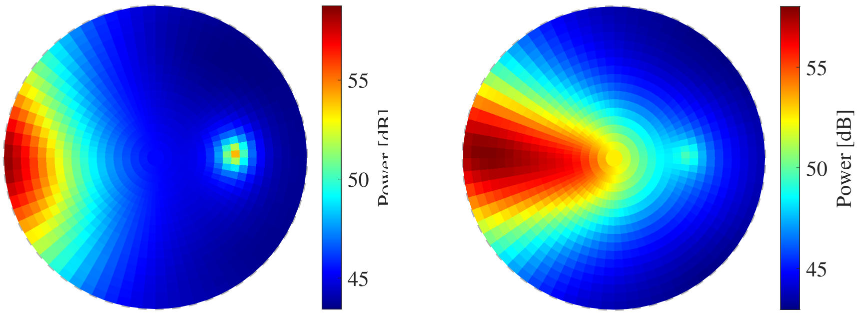
Table 3 shows the measurement results of the RTF for each method. SEVD-MUSIC does not include the processing of noise suppression; therefore, the RTF is lower than 1. GEVD-MUSIC and iGEVD-MUSIC whiten the noise by calculating the inverse of the noise correlation matrix. Therefore, they have a high calculation cost, and the RTF was higher than 1. This means that they cannot process in real time. The RTFs of AFRF-MUSIC and the angle-limited SEVD-MUSIC were significantly lower than that of SEVD-MUSIC. SEVD-MUSIC calculates the spatial spectrum for each direction and each frequency, while AFRF-MUSIC limits the calculated frequency range, and the angle-limited SEVD-MUSIC limits the calculated direction. This greatly reduced the calculation cost. HIST-MUSIC removes noise by adding simple processing to SEVD-MUSIC; therefore, the RTF is almost the same as that of SEVD-MUSIC. HIST-MUSIC-3D repeats the same simple processing as HIST-MUSIC for each frequency bin, resulting in an increase in the RTF of 0.05 compared to SEVD-MUSIC. However, the RTF of HIST-MUSIC-3D is still lower than 1, indicating that HIST-MUSIC-3D can perform real-time SSL.

Table 3. Measurement results of RTF.

Method	RTF
SEVD-MUSIC	0.92
GEVD-MUSIC	1.09
iGEVD-MUSIC	1.23
AFRF-MUSIC	0.18
Angle-limited SEVD-MUSIC (1)	0.54
Angle-limited SEVD-MUSIC (2)	0.73
HIST-MUSIC	0.93
HIST-MUSIC-3D	0.97

3.3.3. Versatility

To evaluate versatility, success rates and searchable ranges were calculated for different ego-noises, as in Section 3.3.1. The ego-noise of DJI Inspire 2 with a high-altitude propeller and ACSL (Tokyo, Japan) MS-06LA were used for different ego-noises, and evaluation signals were created. Figure 24 shows examples of spatial spectra obtained by processing the evaluation signals with SEVD-MUSIC. As can be seen, the ego-noise has different characteristics from that in Section 3.2. The created evaluation signals were processed with HIST-MUSIC and HIST-MUSIC-3D, and the results were compared.



(**a**) DJI Inspire 2 with a high-altitude propeller.

(**b**) ACSL MS-06LA.

Figure 24. Examples of spatial spectra processed via SEVD-MUSIC.

Figures 25 and 26 show the calculated success rate and searchable range rate for different ego-noises. The horizontal axis represents the SNR, and the vertical axis represents the success rate and the searchable range rate, respectively. The solid lines show the results for DJI Inspire 2 with the normal propeller shown in Section 3.3.1, the dashed lines show the results for DJI Inspire 2 with a high-altitude propeller, and the dotted lines show the results for ACSL MS-06LA. Figures 27 and 28 show the success rate for each set DOA at the SNR of −16 dB when the ego-noise of DJI Inspire 2 with a high-altitude propeller and ACSL MS-06LA was used. Each set DOA follows the axes given in Figure 2, and the success rate for each set DOA is depicted in a color map. The results of HIST-MUSIC show that the success rate and the searchable range differ, depending on the ego-noise. This is because HIST-MUSIC does not consider frequency information, so the range where localization fails changes when the ego-noise characteristics change, as shown in Figures 27a and 28a. On the other hand, the results of HIST-MUSIC-3D do not change significantly even if the ego-noise is different, and a high success rate and large search range are achieved. This is because HIST-MUSIC-3D considers frequency information, so even if the ego-noise characteristics change, the ego-noise component is accurately excluded, and the range in which localization fails does not change, as shown in Figures 27b and 28b. Therefore, it was found that HIST-MUSIC-3D has high versatility.

These results confirm that the proposed method, HIST-MUSIC-3D, has a high tolerance for quickly changing dynamic noise, a large searchable range, high real-time performance, and high versatility and that it is practical in real environments.

The performance of the conventional methods used for comparison is not necessarily consistent with that introduced in each method [16,17,19,21–23], and lower performance was observed in this evaluation experiment. The decline in performance was considered to occur because the previous evaluation experiments were not closer to a real environment, such as using ego-noise while hovering. In this paper, the evaluation experiments were performed under more severe conditions assumed in a real environment; therefore, the performance of the conventional method declined. The proposed method was developed using data obtained under such severe conditions, and thus, the performance was improved compared to the conventional method. On the other hand, it is necessary to know the performance limitations of the proposed method for its practical use in search and rescue operations, for example, for how low an SNR the target sound can be localized, whether it is possible to localize sounds other than voice and whistle sounds, whether it can handle a variety of drone models, and so on. These performance limitations need to be clarified in future work.

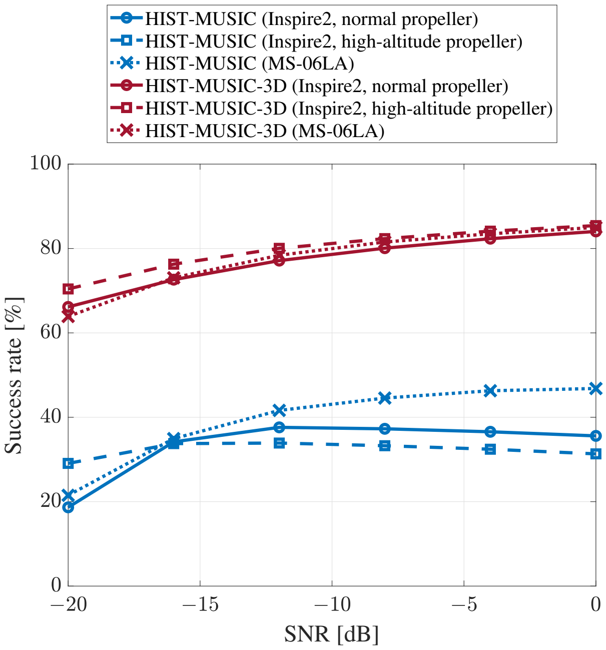


Figure 25. Success rate of SSL for different ego-noises.

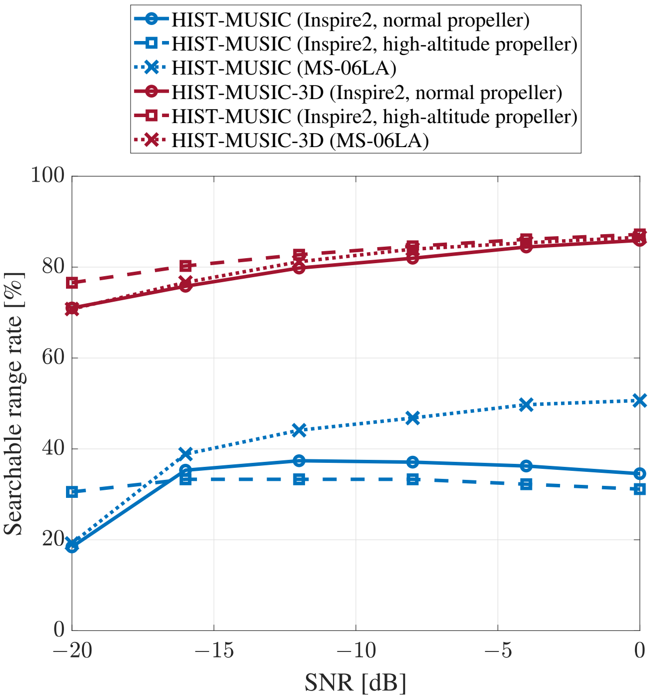
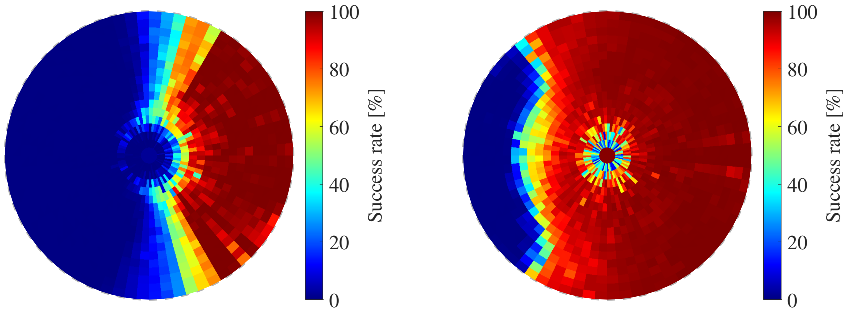
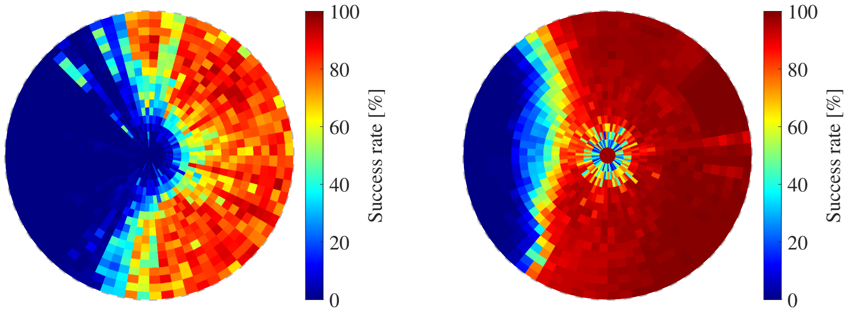


Figure 26. Searchable range rate of SSL for different ego-noises.



(**a**) HIST-MUSIC. (**b**) HIST-MUSIC-3D.

Figure 27. Success rate for each set DOA at the SNR of −16 dB when the ego-noise of DJI Inspire 2 with a high-altitude propeller was used.



(**a**) HIST-MUSIC. (**b**) HIST-MUSIC-3D.

Figure 28. Success rate for each set DOA at the SNR of −16 dB when the ego-noise of ACSL MS-06LA was used.

The issue of the proposed method is the flying speed of the drone that can localize the sound source. In this experiment, the maximum speed of the drone was 3 m/s. Including our proposed method, MUSIC-based SSL methods do not consider that the target sound direction relative to the microphone array changes within the processing frame (0.5 s in this experiment). For example, if the drone is flying at a speed of 3 m/s, the maximum change in the target sound direction in a processing frame is 8.5 deg., and if it is flying at a speed of 10 m/s, the maximum change is 26.6 deg. If there is a large change in the target sound direction within a processing frame, accurate SSL cannot be achieved. Based on this reason and the results of this experiment, the maximum speed of the drone at which drone audition technology is useful is considered to be a few m/s. However, for prompt search and rescue tasks, SSL at faster drone speeds needs to be addressed in future work.

4. Conclusions

In this paper, we have proposed a novel SSL method using a drone-embedded microphone array to satisfy four requirements for drone audition: a high tolerance for quickly changing dynamic noise, a large searchable range, high real-time performance, and high versatility. The proposed method dynamically extracts ego-noise and target sound components from a three-dimensional spatial spectrum with directional and frequency axes at the current time using histogram information. The DOA of the target sound is estimated by separating them using continuity and extracting only the target sound component. These are performed using only simple calculations. Therefore, all the requirements are satisfied simultaneously. Evaluation experiments using the ego-noise recorded in a real outdoor environment show that the localization performance of the proposed method is higher than that of the existing and previously proposed methods, indicating the usefulness of the proposed method.

Author Contributions: Conceptualization, K.H., I.K. and N.I.; methodology, K.H., I.K. and N.I.; software, I.K.; validation, K.H., I.K. and N.I.; formal analysis, I.K.; investigation, K.H., I.K. and N.I.; resources, K.H. and N.I.; data curation, I.K.; writing—original draft preparation, K.H.; writing—review and editing, K.H., I.K. and N.I; visualization, I.K.; supervision, K.H.; project administration, K.H.; funding acquisition, K.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by JSPS (Japan Society for the Promotion of Science) KAKENHI Grant No. JP22K14218 and the Ono Charitable Trust for Acoustics.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. World Health Organization. Türkiye Earthquake: External Situation Report no. 4: 6–12 March 2023. Available online: https://www.who.int/europe/publications/i/item/WHO-EURO-2023-7145-46911-68727 (accessed on 24 March 2023).
2. World Health Organization. Whole of Syria: Situation Report, 3–12 March 2023. Available online: https://www.emro.who.int/images/stories/syria/whole-of-syria-sit-rep-6-12-march-2023.pdf (accessed on 24 March 2023).
3. Tadokoro, S. *Disaster Robotics: Results from the ImPACT Tough Robotics Challenge*; Springer: Berlin, Germany, 2019; pp. 1–528.
4. Quan, A.; Herrmann, C.; Soliman, H. Project Vulture: A Prototype for Using Drones in Search and Rescue Operations. In Proceedings of the 15th International Conference on Distributed Computing in Sensor Systems (DCOSS), Santorini, Greece, 29–31 May 2019; pp. 619–624. [CrossRef]
5. Sambolek, S.; Ivasic-Kos, M. Automatic Person Detection in Search and Rescue Operations Using Deep CNN Detectors. *IEEE Access* **2021**, *9*, 37905–37922. [CrossRef]
6. Tanida, N. What happened to elderly people in the great Hanshin earthquake. *BMJ* **1996**, *313*, 1133–1135. . [CrossRef] [PubMed]
7. Go, Y.-J.; Choi, J.-S. An Acoustic Source Localization Method Using a Drone-Mounted Phased Microphone Array. *Drones* **2021**, *5*, 75. [CrossRef]
8. Wang, L.; Cavallaro, A. Deep Learning Assisted Time-Frequency Processing for Speech Enhancement on Drones. *IEEE Trans. Emerg. Top. Comput. Intell.* **2021**, *5*, 871–881. [CrossRef]
9. Strauss, M.; Mordel, P.; Miguet, V.; Deleforge, A. DREGON: Dataset and Methods for UAV-Embedded Sound Source Localization. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1–5 October 2018; pp. 1–8. [CrossRef]
10. Ishiki, T.; Washizaki, K.; Kumon, M. Evaluation of Microphone Array for Multirotor Helicopters. *J. Robot. Mechatron.* **2017**, *29*, 168–176. [CrossRef]
11. Nonami, K.; Hoshiba, K.; Nakadai, K.; Kumon, M.; Okuno, H.G.; Tanabe, Y.; Yonezawa, K.; Tokutake, H.; Suzuki, S.; Yamaguchi, K.; Sunada, S.; Takaki, T.; Nakata, T.; Noda, R.; Liu, H.; Tadokoro, S. Recent R&D Technologies and Future Prospective of Flying Robot in Tough Robotics Challenge. In *Disaster Robotics*; Tadokoro, S., Ed.; Springer: Cham, Switzerland, 2019; pp. 77–142. [CrossRef]
12. Madokoro, H.; Yamamoto, S.; Watanabe, K.; Nishiguchi, M.; Nix, S.; Woo, H.; Sato, K. Prototype Development of Cross-Shaped Microphone Array System for Drone Localization Based on Delay-and-Sum Beamforming in GNSS-Denied Areas. *Drones* **2021**, *5*, 123. [CrossRef]
13. Noda, R.; Nakata, T.; Senda, K.; Liu, H. Multi-scale morphological effect on noise level and frequency characteristics of drone propellers. In Proceedings of the Quiet Drones, Paris, France, 19–21 October 2020.
14. Hoshiba, K.; Noda, R.; Nakata, T.; Liu, H.; Senda, K.; Nakadai, K.; Kumon, M.; Okuno, H.G. Development of surface-processed low-noise propeller for search and rescue tasks with drone audition. In Proceedings of the Quiet Drones, Paris, France, 19–21 October 2020.
15. Noda, R.; Nakata, T.; Senda, K.; Liu, H. Development of Microstructured Low Noise Propeller for Aerial Acoustic Surveillance. In Proceedings of the 2021 IEEE/SICE International Symposium on System Integration (SII), Iwaki, Japan, 11–14 January 2021; pp. 482–486. [CrossRef]
16. Schmidt, R.O. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propag.* **1986**, *34*, 276–280. [CrossRef]
17. Nakamura, K.; Nakadai, K.; Asano, F.; Hasegawa, Y.; Tsujino, H. Intelligent sound source localization for dynamic environments. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), St. Louis, MO, USA, 10–15 October 2009; pp. 664–669. [CrossRef]
18. Nakamura, K.; Nakadai, K.; Ince, G. Real-time super-resolution Sound Source Localization for robots. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vilamoura-Algarve, Portugal, 7–12 October 2012; pp. 694–699. [CrossRef]
19. Okutani, K.; Yoshida, T.; Nakamura, K.; Nakadai, K. Outdoor auditory scene analysis using a moving microphone array embedded in a quadrocopter. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vilamoura-Algarve, Portugal, 7–12 October 2012; pp. 3288–3293. [CrossRef]

20. Ohata, T.; Nakamura, K.; Mizumoto, T.; Tezuka, T.; Nakadai, K. Improvement in outdoor sound source detection using a quadrotor-embedded microphone array. In Proceedings of the IEEE/RSJ International Conference on Robots and Intelligent Systems (IROS), Chicago, IL, USA, 14–18 September 2014; pp. 1902–1907. [CrossRef]
21. Hoshiba, K.; Washizaki, K.; Wakabayashi, M.; Ishiki, T.; Kumon, M.; Bando, Y.; Gabriel, D.; Nakadai, K.; Okuno, H.G. Design of UAV-Embedded Microphone Array System for Sound Source Localization in Outdoor Environments. *Sensors* **2017**, *17*, 2535. [CrossRef] [PubMed]
22. Hoshiba, K.; Nakadai, K.; Kumon, M.; Okuno, H.G. Assessment of music-Based Noise-Robust Sound Source Localization with Active Frequency Range Filtering. *J. Robot. Mechatron.* **2018**, *30*, 426–435. [CrossRef]
23. Komatsuzaki, I.; Hoshiba, K.; Takeda, Y.; Sugahara, Y. Propopsal of noise-robust sound source localization method using histogram information for quickly changing dynamic noise. In Proceedings of the 40th Annual Conference of the Robotics Society if Japan, Tokyo, Japan, 5–9 September 2022. (In Japanese)
24. Mori, S.; Hirata, S.; Yamaguchi, T.; Hachiya, H. Quantitative evaluation method for liver fibrosis based on multi-Rayleigh model with estimation of number of tissue components in ultrasound B-mode image. *Jpn. J. Appl. Phys.* **2018**, *57*, 07LF17. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.