

Article

An Acoustic Source Localization Method Using a Drone-Mounted Phased Microphone Array

Yeong-Ju Go  and Jong-Soo Choi *

Department of Aerospace Engineering, Chungnam National University, 99 Daehak-ro, Yuseong-gu, Daejeon 34134, Korea; yjgo@cnu.ac.kr

* Correspondence: jchoi@cnu.ac.kr

Abstract: Currently, the detection of targets using drone-mounted imaging equipment is a very useful technique and is being utilized in many areas. In this study, we focus on acoustic signal detection with a drone detecting targets where sounds occur, unlike image-based detection. We implement a system in which a drone detects acoustic sources above the ground by applying a phase difference microphone array technique. Localization methods of acoustic sources are based on beamforming methods. The background and self-induced noise that is generated when a drone flies reduces the signal-to-noise ratio for detecting acoustic signals of interest, making it difficult to analyze signal characteristics. Furthermore, the strongly correlated noise, generated when a propeller rotates, acts as a factor that degrades the noise source direction of arrival estimation performance of the beamforming method. Spectral reduction methods have been effective in reducing noise by adjusting to specific frequencies in acoustically very harsh situations where drones are always exposed to their own noise. Since the direction of arrival of acoustic sources estimated from the beamforming method is based on the drone's body frame coordinate system, we implement a method to estimate acoustic sources above the ground by fusing flight information output from the drone's flight navigation system. The proposed method for estimating acoustic sources above the ground is experimentally validated by a drone equipped with a 32-channel time-synchronized MEMS microphone array. Additionally, the verification of the sound source location detection method was limited to the explosion sound generated from the fireworks. We confirm that the acoustic source location can be detected with an error performance of approximately 10 degrees of azimuth and elevation at the ground distance of about 150 m between the drone and the explosion location.

Keywords: acoustic source localization; phased microphone array; spectral subtraction; beamforming



Citation: Go, Y.-J.; Choi, J.-S. An Acoustic Source Localization Method Using a Drone-Mounted Phased Microphone Array. *Drones* **2021**, *5*, 75. <https://doi.org/10.3390/drones5030075>

Academic Editors: Arianna Pesci, Giordano Teza, Massimo Fabris and Dimitrios Zarpalas

Received: 23 June 2021

Accepted: 4 August 2021

Published: 6 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

One way to cope with disasters and security situations is to use drones to search for where events occur [1,2], which can provide visual information by taking images from the sky. In general, drone imaging has been used for military purposes, and it is also the most commercially used method for video production. The usability of drones has already been demonstrated, and as the operating area gradually gets closer to humans, there is a demand for new detection methods beyond existing mission equipment that only provides images. One new method is to detect sounds.

Acoustic-based detection methods in disaster and security situations can complement the limitations of image-based target detection methods. In addition, acoustic-based detection methods are needed in situations where image identification is needed at night, in bad weather, and over complex terrain with features that can be difficult to identify [3,4]. This technology extends detection methods by adding human hearing-like sensations to drones. In particular, it is considered to be a necessary technology for detecting the location of explosions in disaster situations such as fires or detecting distress signals for lifesaving situations.

The technology of detecting targets using acoustic signals has been previously studied in a variety of fields [5–7], from implementing a robot that detects sounds by mimicking human ears to noise source instrumentation systems for the analysis of mechanical noise. Prior to the development of radar, this technique was used to detect the invasion of enemy aircraft by sound to defend against anti-aircraft. In recent years, it has also been applied to systems that track where guns are fired. Recently, it has become an essential technology for low noise research and has been used to detect and quantify acoustic sources of noise generated by transport vehicle systems such as automobiles and aviation. This technique is based on the principle of detecting acoustic sources by reconstructing acoustic pressure measured by the concept of phase difference using microphone arrays.

The most significant problem associated with detecting external noise by mounting microphones on a drone is the noise of the drone when it operates. The main noise generated by a drone is aeroacoustics noise generated mainly by rotating rotors [8]. Depending on the mechanism of occurrence, a significant factor that affects the tonal or broadband frequency band is related to the blade passing frequency [9,10]. A typical drone has more than four rotors, which creates a more complex noise field environment [11,12]. Furthermore, the operating noise of a drone in close proximity to a microphone can cause the sound to be masked and indistinguishable. Therefore, robust denoising techniques are essential for identifying external sounds in highly noisy environments. It is ideal for a drone to reduce its own noise in terms of operability, but this is difficult to implement immediately with the current technology; therefore, signal processing that reduces noise from signals acquired from microphones is realistic. Spectrum subtraction is a method of eliminating noise and obtaining clean sound to improve the signal-to-noise ratio in a noisy environment. This is one of the methods of background noise suppression and has been studied for the purpose of clarifying voices in the field of voice recognition [13,14]. In general, background noise is removed by using statistical characteristics of background noise, but applicable methods are also being studied in non-stationary noise environments [15–17]. In addition, there are cases that are effective in improving the signal-to-noise ratio in sound detection such as impact sound [18].

Beamforming methods are well known as methods for estimating the direction of arrival of acoustic sources [19–22]. Beamforming is the principle of calculating beam power by using multiple microphones to correct the phase difference of signals according to the geometric positional relationships among the microphones and estimating the direction of arrival of the source from that intensity [23]. The direction of arrival estimation performance of the beamforming method is determined by the number of microphones, the form of the array, and an improved algorithm. Beamforming methods can also reduce the strength of uncorrelated signals by phase difference correction and increase the strength of the signal of interest to enhance the signal-to-noise ratio by reducing noise. Generally, the greater the number of microphones and the higher the caliber, the better the performance; however, limited number and size microphone arrays are used considering real-time processing and hardware performance.

Recently, studies have been conducted to detect sound sources by installing microphone arrays in drones. Although the use of microphone arrays to detect or distinguish sound sources is common, there are various results depending on the hardware that makes up the array and the array signal processing algorithms. In [24,25], the authors studied how to attach microphone arrays to drones and embedded systems for signal acquisition. In [26–28], the authors presented an arrival angle detection algorithm for sound sources based on beamforming. In addition, valid results for the detection of sound sources were shown through verification experiments on near-field sound sources in a well-refined indoor environment. There are also sound source detection studies on the actual operation of drones in outdoor environments. In [29,30], detection studies were conducted on whistles and voices. These studies confirmed the performance of sound source detection in terms of the signal-to-noise ratio of sound source and background noise, and showed that it is possible to detect nearby sound sources located about 20 m away from drones.

In this study, a method is proposed for detecting sound sources generated above the ground using a drone-mounted phased microphone array. Especially, we tried to effectively remove noise caused by rotors from the microphone signal and to confirm the expectation of the accuracy of the sound source localization in detection mission that occurs during flight. Each of the techniques needed to detect the sound source suggested a major perspective that must be addressed. The goal is to confirm the expected performance of the location detection of sound sources by actually implementing the process of connecting techniques. Since microphones are so close to the drone, the sound of interest is distorted by noise generated when the drone operates, and the spectral subtraction method is used to improve the signal-to-noise ratio. The application of the general spectral subtraction method applies the average model in the entire frequency band. However, in order to effectively remove the drone noise that is an obstacle to us, we separated the spectral band with clear spectral characteristics and applied the subtraction method through different models. The separated bands are divided into bands due to the main appearance of BPF and other noises, including turbulent flow, and are designed to be effective in reducing drone noise.

We distinguished impact sound through spectral subtraction. Spectral reduction has been shown to be effective in restoring acoustic signals of impact sounds. Using this denoised signal, we detected the direction of arrival for ground impact sounds by beamforming methods based on microphone arrays. We applied a method of representing the direction of arrival as absolute coordinates for the ground source, measured on the basis of the microphone array coordinates mounted on a drone, and a data fusion method that can detect the location of sound sources by correcting the changing posture in real time using drone flight information. The entire algorithm for detecting sound sources experimentally identifies detection performance using a 32-channel microphone array. To verify the proposed methods, a clear detection sound source was needed. We focused on identifying the localization error for point sound sources with a clear signal such as impulse. In the verification experiment, the localization performance was confirmed by limiting the impact sound using a firecracker.

In Section 2, we describe how to detect the location of a ground acoustic source. We describe a spectral reduction method to improve the signal-to-noise ratio for acoustic sources of interest, a beamforming method to estimate the direction of arrival of acoustic sources, and a method to represent the direction of arrival of acoustic sources detected by drones with geographical information above the ground. In Section 3, we describe the experiments to verify the location detection performance for ground acoustic sources. We describe the microphone array system and experimental environment mounted on a drone. In Section 4, we address the analysis of data measured through the experiments and the estimation performance of ground acoustic sources.

2. Materials and Methods

The method for detecting ground acoustic sources can be divided into three stages: (1) the measured signal in a microphone array uses spectral subtraction to improve the signal-to-noise ratio, (2) the beamforming method uses time domain data to estimate the angle of arrival for the acoustic source of interest, (3) the flight data of the drone are fused to estimate the location of the source on the ground.

2.1. Spectral Subtraction

Spectral subtraction is a method of restoring a signal by subtracting estimates of noise from a noisy signal [31]. In a noisy environment, the measured signals are represented by the sum of the signals of interest, $s(t)$, and noise, $n(t)$. Here, the signal of interest is the signal we want to measure, and noise is the signal we want to remove, not the signal of interest:

$$y(t) = s(t) + n(t) \quad (1)$$

Spectral subtraction is performed by all computations in the frequency domain. Thus, signals in the time domain are extracted sequentially from block to block through window functions and analyzed as frequency signals through Fourier transformations.

Noise spectra are typically estimated from time intervals where only steady noise exists without signals of interest. This assumes the amplitude of the estimated spectrum as a noise model and computes the difference from the amplitude of the measured signal spectrum to eliminate the averaged noise. Amplitude in the subtracted spectrum applies the phase of the measured signal, and the noise-reduced time domain signal, $s'(t)$, is extracted through the inverse Fourier transform (Figure 1). An important view of the spectral subtraction method applied in this study is that, independently, noise reduction for each microphone signal preserves the phase of the signal of interest, so that the phase relationship between the signals of each microphone is not distorted. Preserving the phase is directly related to estimation error, as the method of estimating the location of acoustic sources by array signal processing is based on the phase relationship of the signals measured on each microphone. The noise model assumes that the signal is consistently and continuously affecting the interval at which it is measured. As it is difficult to estimate sufficient averaged spectra when noise properties change over time, in such cases, a sufficiently averaged spectrum can be obtained by applying a smoothing filter.

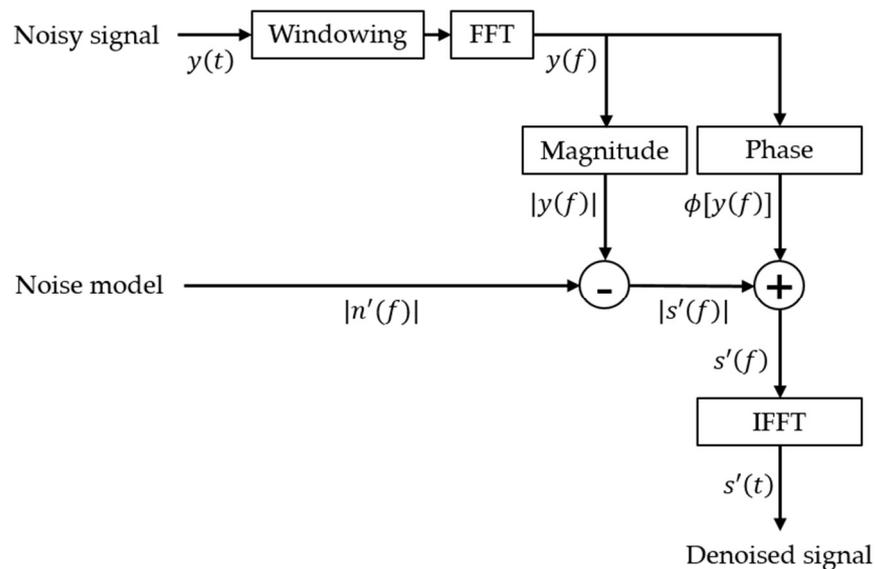


Figure 1. Block diagram of the spectral subtraction method.

2.2. Beamforming

A delay-and-sum beamforming calculated in the time domain was applied. In the time domain, beam power, b , can be expressed by the arrival vector of acoustic waves defined in the three-dimensional space geometry, x , and time, t , as follows:

$$b(x, t) = \frac{1}{M} \sum_{i=0}^{M-1} w_i p_i(t - \tau_i) \quad (2)$$

where p is the acoustic pressure measured on the microphone and relative time delay, τ is the relative time delay caused by the position of each microphone, w is the weight and does not give a differentiated weight, M represents the number of stagnant microphones used, and i represents the microphone index. Beam power is calculated for each direction of arrival of a hypothetical acoustic source. The magnitude of the calculated beam powers is determined for the direction of arrival of the acoustic source by making a relative comparison in the entire virtual source space and estimating the direction of arrival with the maximum value. The phase of the sound pressure signal measured between the

reference point of the microphone, P_{ref} , and the position between each microphone, P_i , is caused by a time delay, τ_i ; τ_i is analytically related by speed of sound, c_0 , and delay distance, dl_i , as in Equation (3) [32,33], and dl_i is calculated from the position vector \mathbf{r}_m defined as P_{ref} and P_i and the incident direction of the plane wave, i.e., the normal vector of the plane wave, \mathbf{n} (Figure 2). The normal vector of the plane wave is the relation between the azimuth angle, Φ , and elevation angle, Θ , which can be represented in the spherical coordinate system. The speed of sound was corrected by atmospheric temperature, T [34], as follows:

$$\tau_i = dl_i / c_0 \quad (3)$$

$$dl_i = \mathbf{n} \cdot \mathbf{r}_m / |\mathbf{n}| \quad (4)$$

$$\mathbf{r}_m = P_i - P_{ref} \quad (5)$$

$$\mathbf{n} = (\cos \Phi \sin \Theta, \sin \Phi \sin \Theta, \cos(\Theta)) \text{ for } 0 \leq \Theta \leq 2\pi, 0 \leq \Phi \leq \pi/2 \quad (6)$$

$$c_0 = 331.3 \sqrt{(1 + T/273.15)} \quad (7)$$

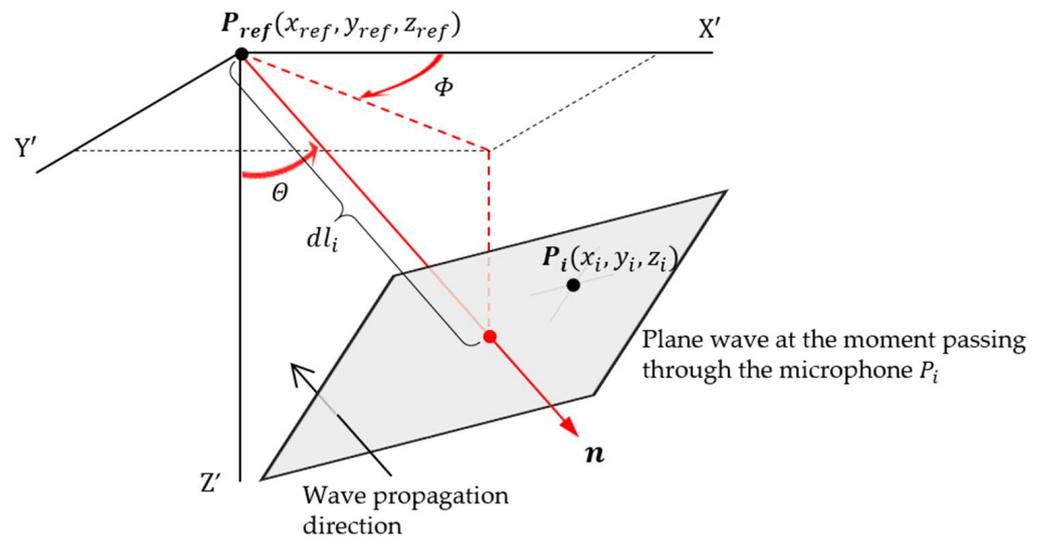


Figure 2. The relationship between the incident direction of the plane wave and the location of the microphone.

A virtual source for calculating beam power was assumed to be a ground source, and based on the drone's headings, the azimuth angle, Φ , was detected from 0 to 360° and the elevation angle, Θ , from 0 to 90°. The definition of detection angles defined between the drone's heading vector and the direction of arrival of the detected acoustic source is shown in Figure 3. The beam power calculated from the lower hemisphere range of the drone explored the maximum beam power and estimated its orientation as the angle of arrival at which the acoustic source enters.

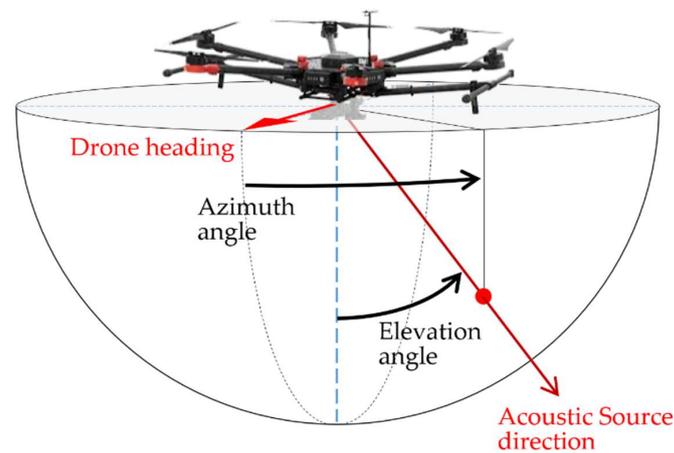


Figure 3. DOA estimation coordinate system of the microphone array based on drone heading direction.

2.3. Acoustic Source Localization on the Ground

The location detection of acoustic sources generated on the ground is determined by matching the state information the drone is flying with the direction of arrival of acoustic sources estimated from the drone-mounted microphone array. The direction of arrival of an acoustic source is estimated to be the azimuth and elevation angles with the maximum beam power, based on the heading direction of the drone's body frame coordinate system. The body frame coordinate is calculated by the roll, pitch, yaw angle, and heading vector based on the magnetic northward and navigation frame coordinate system from the GPS [35–37]. The drone's posture and position changing in real time is updated to 30 Hz on the data processing board with signals from the microphone. The drone-mounted microphone array estimates the angle of arrival based on the drone's heading direction. The conversion of the heading direction to the navigation frame coordinate system coordinates converts the estimated angle of arrival to the same criterion to geometrically derive the point of intersection with the ground with geometric information above the ground, and estimates this intersection as an acoustic source (Figure 4).

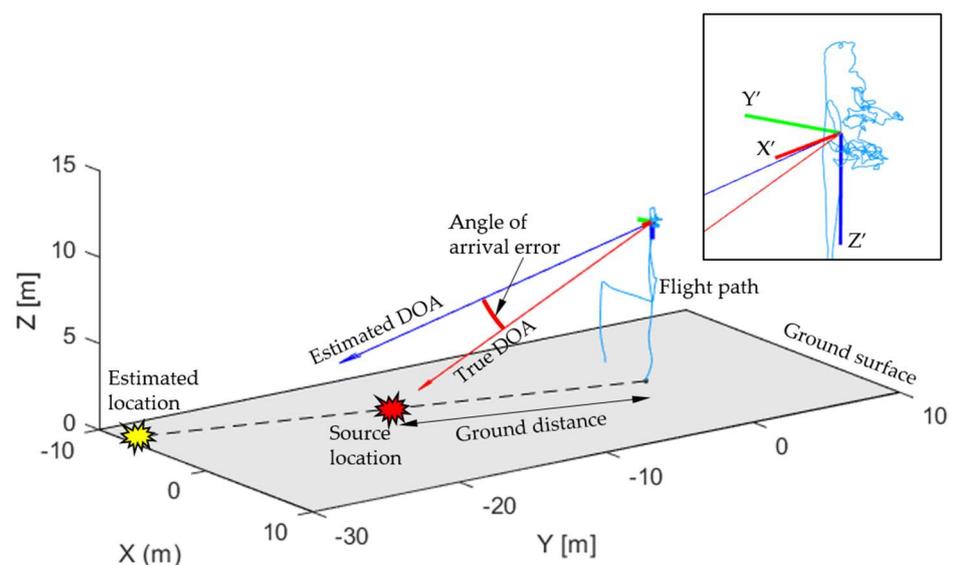


Figure 4. The concept of estimating the location of the acoustic source on the ground based on the DOA measured by the microphone array.

2.4. Phased Microphone Array Mounted on the Drone

The microphone array system consisted of 4 microphone modules with support mounts, and a signal processing board. One microphone module had a radius of 0.025 m, and 8 microphones were arranged in a circular equidistant array (Figure 5). Each microphone module was attached to the underside of the hemispherical dome to avoid the flow-induced pressure caused by the propeller wake. As shown in Figure 6, four microphone modules were attached to each end of the X-shaped mount with a radius of 0.54 m, and thus an array of 32 microphones was implemented. The 32-channel microphone array is a prototype built for the purpose of studying signal acquisition and location detection for acoustic sources of interest outside the drone, designed to verify the performance of several signal processing algorithms. The x-shaped support allows for adjustments of angle and height, as well as the array shape. All the acoustic pressure signals measured on each microphone are time synchronized and can be collected via a separate data acquisition board and stored in external memory in accordance with the acoustic event or trigger switch from the external controller. For verification experiments, signals were manually recorded using the trigger function of an external controller in time for acoustic source generation, and acoustic signals stored on memory cards were post-processed. The microphone array specifications are listed in Table 1.



Figure 5. Microphone module consisting of 8 microphones.



Figure 6. Microphone array system mounted on the drone.

Table 1. Microphone array specifications.

| Parameter | Specification |
|------------------------|----------------|
| Number of channels | 32 |
| Sensitivity | −26 dBFS |
| Signal-to-noise ratio | 64 dB(A) |
| Maximum pressure level | 122.5 dB |
| Sampling rate | 25.6 kS/s |
| Frequency range | 25 Hz~12.8 kHz |

3. Experiment

Experiments to verify the acoustic source localization performance were conducted in a stadium at Chungnam National University. The experimental environment is shown in Figure 7, indicating the locations of the acoustic sources, i.e., firecrackers, and where the drone was hovering in place. To quantify the locations of the sources, a line in the soccer stadium was referred to and the absolute location was measured using GPS. The Universal Transverse Mercator Coordinate System (UTM) using the WGS84 ellipsoid was used to derive plane coordinates based on the central position (Location 4) of the soccer stadium. The locations of the acoustic sources on this plane coordinate and the location of the drone are defined as absolute coordinates, and the results of estimating the location of the acoustic source based on the drone's GPS information can also be expressed.

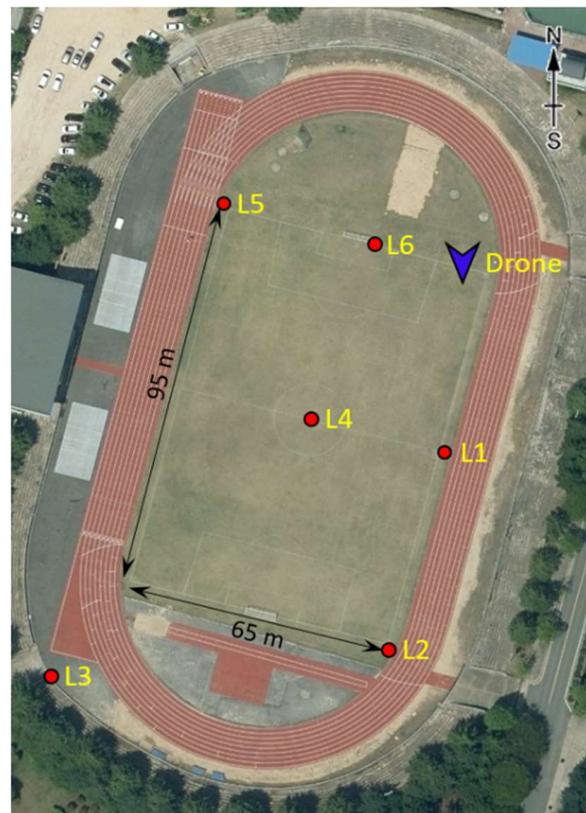


Figure 7. Experimental site. Firecracker locations (red circles) and drone location (blue arrow) in hovering position.

The acoustic sources used for the experiment were commercial firecrackers, which use gunpowder to generate short impulse signals after a certain period of time when ignited. The acoustic pressure signal from the explosion of the firecracker used in the experiment is shown in Figure 8. This signal was collected separately without a drone and measured at a distance of about 20 m away from the location of the firecracker near

the center of Figure 6. That is, the signal was completely unrelated to the operation of the drone and was intended to observe pure acoustic signal features. Strong impulse signals, such as explosions, are easy to distinguish in domains because of their strong amplitude and clear features. In the practical problem of finding acoustic sources, impulses are only one example of the detection targets, but they are important underlying experiments in the goal of assessing the detection performance that this study wants to identify. Additionally, it is possible to identify whether spectral subtraction methods work well for a specific frequency or frequency band that wants to denoise. The impulse signal has the constant response characteristic across all frequencies. In order to restore the energy of pure impulse signals as much as possible in a state of noise mixing, it is important to subtract only a specific frequency for the noise to be removed. The impulse is one of the useful targets for providing basic clues to determine the acoustic frequency of the drone to be removed. It was confirmed that the explosion signal was characterized by a short and strong shock sound in an instant. The firecrackers were fixed just above the ground. The location estimation results analyzed in the Results section were performed for approximately 20 min during 1 sortie flight. The firecrackers were detonated three times at each point in the order of location numbers from L1 to L6. The interval between each three repeated explosions was about 10 s. A straight line from the hovering position to the detonating location is described in Table 1, up to 151.5 m. During this flight, the drone maintained a hovering position at an altitude of about 150 m. The drone's position was near the right-hand corner north of the soccer stadium, heading southwards, and it maintained a stable hovering position due to the stabilizer mode.

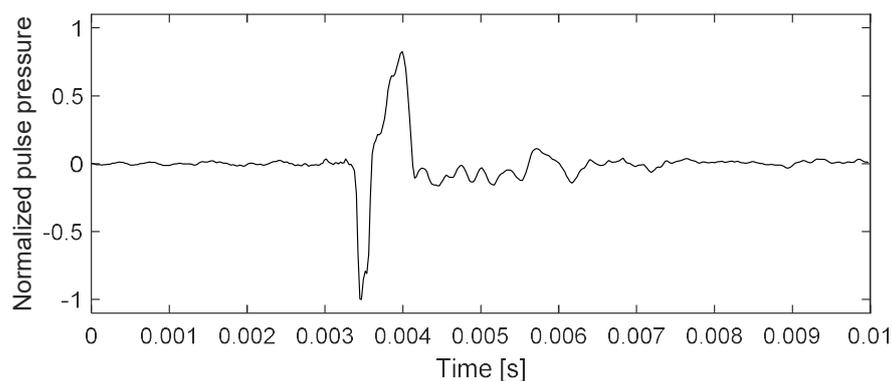


Figure 8. Acoustic pressure signal from the explosion of a firecracker.

4. Results and Discussion

Figure 9 shows the acoustic pressure signals measured by the drone-mounted microphone system, in a hovering position. It is a spectrum that averages acoustic pressure signals measured over 3 s using a total of 32 microphones. The frequency decomposition was 1 Hz, and a 5 Hz high pass filter was applied. This spectrum is meaningful for identifying noise characteristics that actually operate underneath the drone. Noise can be observed together with atmospheric flow and background noise present in the atmosphere when the actual drone is maneuvering. In particular, it is a very poor measurement environment in which fluctuating pressure caused by the propellers' wakes at the bottom of the drone directly affects the microphones. Analysis of the mean spectrum is the basis for determining the frequency bands and amplitude that are necessary to determine the noise generated by drones when flying, and that should be deducted when applying spectral subtraction.

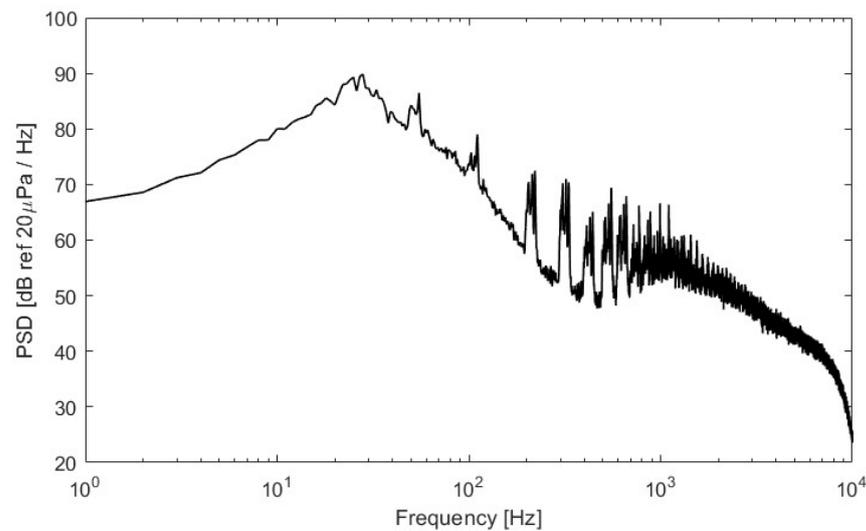


Figure 9. Averaged hovering noise of the drone measured by the mounted microphones.

The drone operates with six rotors, each with two blades. The average motor rotation at hovering was about 3000 rpm. In order to maintain the drone's position during hovering, the number of motor rotations changed by about 200 rpm. Changes in motor rotation are directly related to the blade passing frequency and affect the harmonic components that appear as tonal component features. In the average spectrum, six motors remained hovering for three seconds, outputting slightly different rotations, resulting in peak frequencies scattered for each harmonic. Accurately predicting the blade passing frequency and its amplitude should be able to measure each motor's rotation speed and rotation phase, as it is efficient to match the same number of rotations in each motor from the point of view of drone control.

With respect to the average motor rotation speed, the shaft rate frequency characteristic was shown at about 50 Hz, the first harmonic of blade passing frequency at 100 Hz, and harmonic frequencies for its multiples were observed. Harmonics are observed for rotational frequencies up to 2 kHz, and the tonal frequencies shown here were caused by propellers and motors. Broadband noise above 800 Hz is characterized by turbulence around the propellers. Strong broadband components were observed in bands below 200 Hz. It was observed that this frequency band was mainly caused by atmospheric flow and propeller wake at fluctuating pressures that directly affected the microphones, resulting in high amplitude.

Spectral reduction methods were treated differently for each of the three frequency band intervals in the frequency domain. The first band is the low frequency band below 200 Hz, mainly affected by atmospheric flow and rotor wake. Here, the fluctuating pressure components caused by the flow around the microphones have the greatest effect on the acoustic pressure of the entire frequency characteristics. In the signal we were trying to detect, this band was completely eliminated because it contributed less and was generally associated with high noise amplitude. The second band is from 200 to 1000 Hz, which features the blade passing frequency of the rotor. Blade passing frequency is directly affected by rotor rotation speed, as shown in Figure 9. Since six rotors were operating at different rotations for maneuvering during flight, peak frequencies continued to change over time even in hover flight. Since this peak frequency is difficult to detect accurately in real time, the excess of amplitude as compared with the spectral reduction model was deducted in the same amount as the model. The third band is over 1000 Hz, containing motor noise, which makes it difficult to specify the frequency characteristics of this section. As the processing method is not clear, we simply subtracted the amount of amplitude the subtraction model had from the measured amplitude. The subtraction of amplitude is calculated for each frequency in the spectrum and is in Pascal units, not dB scales. The

analysis of impulse sources requires short-time spectral analysis to process signals in real time. Short-time spectral analysis does not show a clear spectrum curve as compared with the averaged spectrum. Small peak amplitudes were removed through the spectrum smoothing filter to derive amplitude envelope similar to the mean properties and applied as a subtraction model. For the subtraction model, the negative pressure data used signals from the negative pressure signal window of interest, which took into account real-time processing. The signals for the subtraction model were used one second before the window frame of the acoustic pressure signal of interest, considering real-time processing. The frequency and acoustic pressure level models to be subtracted calculate the spectrum of acoustic pressure data measured one second in advance from real-time measured pressure signals and derive the mean characteristics through a smoothing filter. It was not appropriate to use a model to deduct long-time averaged signals because the impact of the number of propeller rotations and wake changed rapidly as the drone was maneuvered.

Figures 10 and 11 show the data that indicate before and after the application of the spectral subtraction method and are measured on microphone Channel 0 for the first of three explosions at Location 1. Explosive sounds represent strong impulse signal characteristics, and the ideal impulse signal has a flat amplitude in the whole band in the frequency domain. The explosive sound of a firecracker affects amplitude over the entire band of frequencies because it produces impulse waves due to rapid pressure changes. Therefore, in order to effectively preserve the amplitude of the impulse signal, spectral subtraction must be performed with the correct focus on the noise frequency. Figure 11 shows the result of applying spectral subtraction to extract frequencies for explosives. By comparing the spectra, it can be observed that the low frequency band below 200 Hz has been eliminated and that the strong tonal component between 100 and 1000 Hz has been reduced. The 32-channel signal, which passed the spectral reduction method for each channel, was restored back to the time domain signal by a 0.5 s, 50% overlap window function. This process made it possible to specify the time the impulse signal and the time the signal occurred in the time domain data.

The results of spectral subtraction on the impulse signals of firecrackers can be observed more clearly in Figure 12. We compared the subtraction model obtained through the smoothing filter, the measured signal, and the spectrum after applying the subtraction method. Different subtractions can be distinguished in bands below 200 Hz, between 200 Hz and 1000 Hz, and above 1000 Hz. From this setting, we were able to effectively extract the acoustic pressure of the impulse signal of the firecracker. Defining frequency band discrimination was empirically distinguished here by already measured drone acoustic data, but in the future, it is considered that automatic detection will be possible if the blade passing frequency is detected through the peak finder of the spectrum or if a learned real-time adaptive discrimination filter is used. In addition, the process of spectral subtraction applied in this study can be applied to clarify other acoustic sound sources such as voices. However, this study only confirmed the analysis of the impulse signal and it is necessary to verify the performance of other sounds separately. The spectral subtraction method, implemented separately for each frequency band, was used to eliminate frequencies affected by the drone's own noise and found to be effective in enhancing the signal-to-noise ratio of acoustic pressure. Especially, it is effective in situations where it is difficult to accurately estimate the harmonic components of blade passing frequency that constantly change during flight. However, if a tonal sound source such as a whistle is detected in a band with a smoothing filter to remove the harmonics of blade passing frequency, it may be considered as blade passing frequency, which may cause a decrease in the signal-to-noise ratio. If this situation is to be detected, it is necessary to consider other ways to only subtract blade passing frequencies.

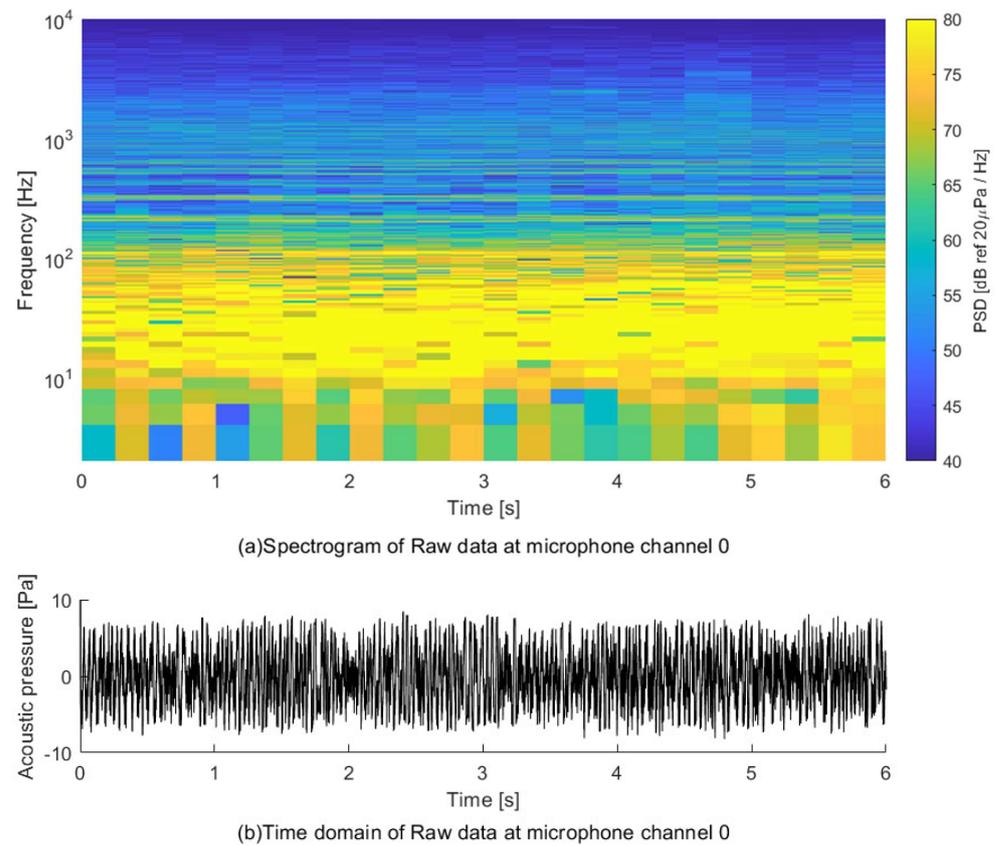


Figure 10. Measured acoustic signal at Location 1 for the first explosion.

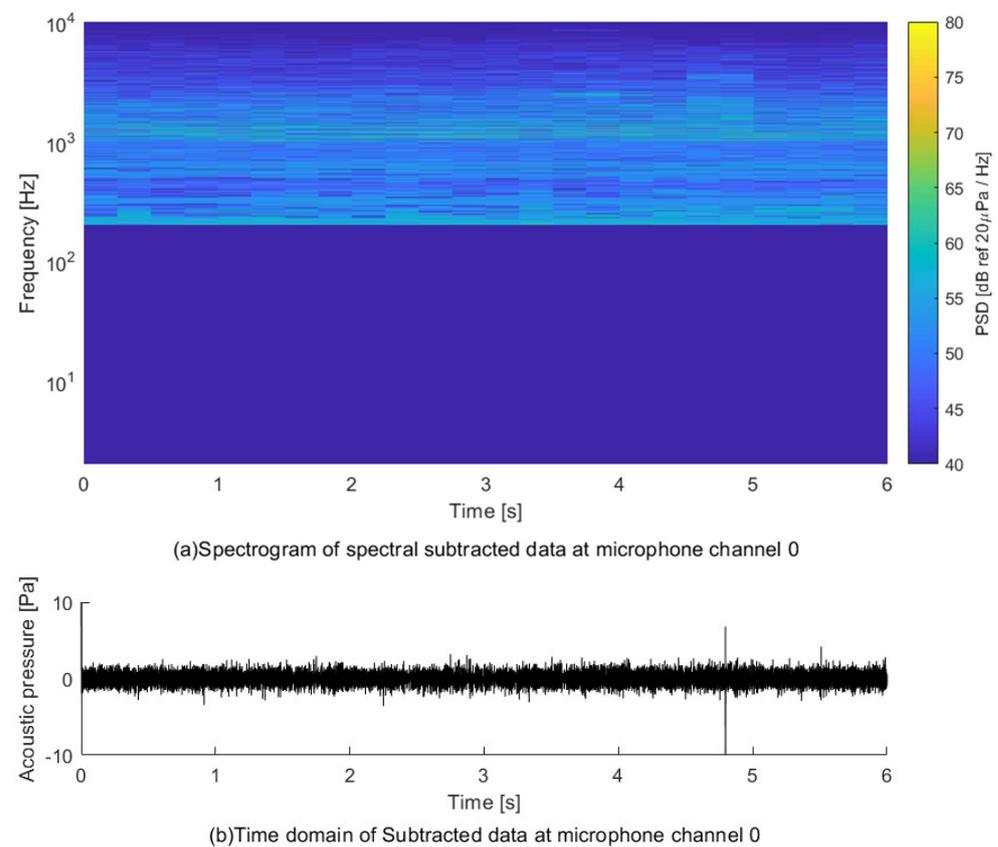


Figure 11. Spectral subtracted acoustic signal at Location 1 for the first explosion.

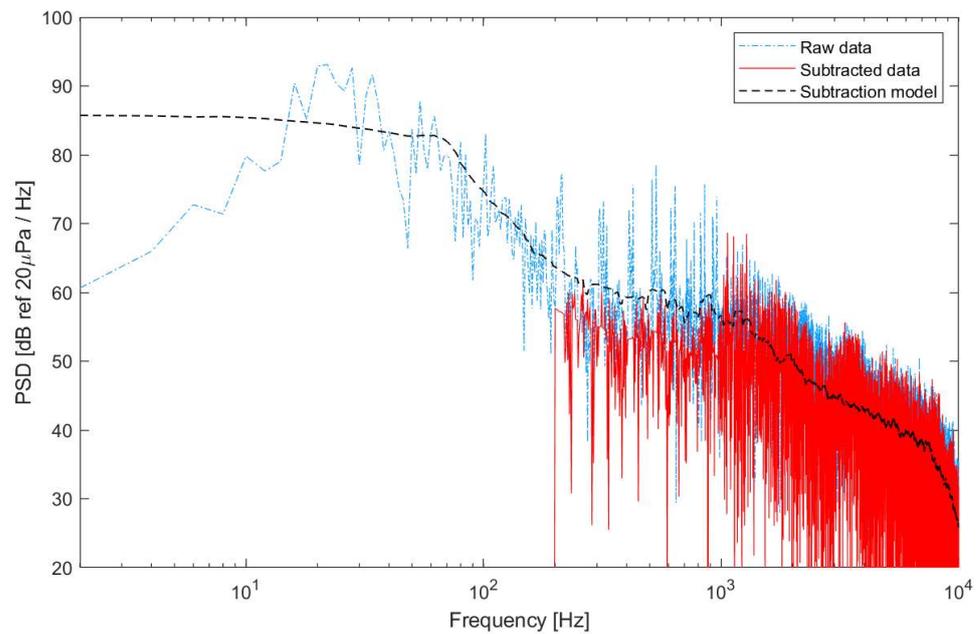


Figure 12. Comparison of the application results of spectral subtraction method to impulse signal (spectra for 4.8 s frames in the spectrogram of Figures 10 and 11).

Time-domain signals with improved signal-to-noise ratio by spectral subtraction methods become input signals that calculate beam power for detecting arrival angles.

Although the beamforming method has the effect of improving the signal-to-noise ratio according to the correlation by phase difference calibration, it is difficult to detect target acoustic sources due to reduced beamforming performance if there is a strong correlation in the proximity field, such as propeller noise. Figure 13 shows the beam power calculated using signals measured from 32 microphones mounted on the drone. The beam power represents the magnitude relative to the lower directions of the drone based on the center point of the drone.

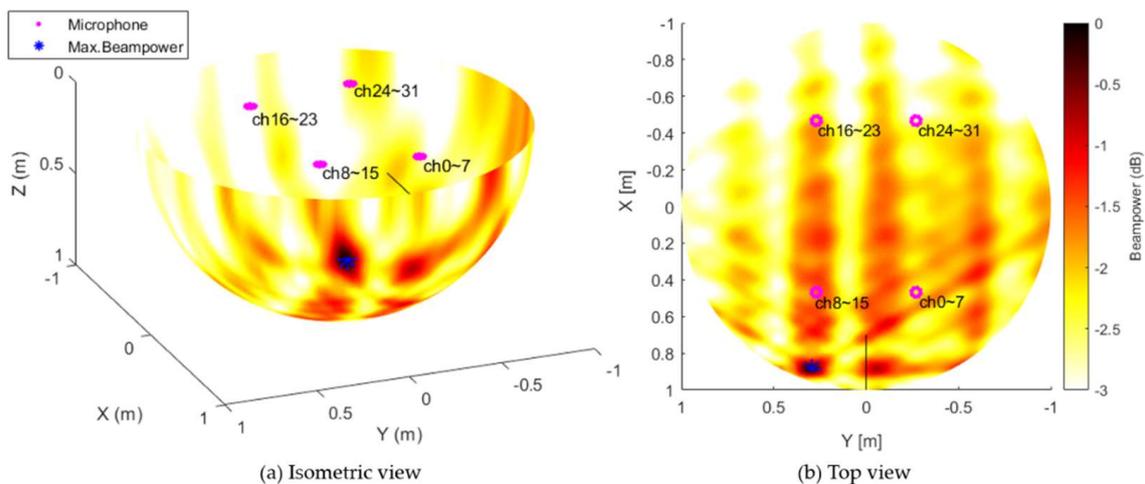


Figure 13. Beam power results at Location 1 for the first explosion.

In general, -3 dB is judged as the effective range as compared with the calculated maximum beam power. Since the 32-channel microphone array used in this study was not considered to be optimized for beamforming methods, sidelobes were also prominent, but we found that the direction of arrival calculated with maximum beam power was the direction of arrival for the actual acoustic source. Beamforming methods are also

computable in frequency domains, where signals that are characterized across broadband, such as impulse signals, are useful for finding features that are calculated in the time domain. Improving beam power performance requires the optimization of the number of available microphones, the maximum implementable aperture size, and the microphone spacing that matches the frequency characteristics of the target acoustic source. Generally, the greater the number of microphones and the larger the aperture size, the better the performance tends to be, but additional consideration is needed to minimize the operation and acoustic measurement interference of drones to be mounted on drones.

The direction of arrival of acoustic waves was estimated by dividing the maximum beam power into the azimuth and elevation angles based on the heading of the flying drone. The drone's headings are related to the body coordinate system, and the roll, pitch, and yaw angle are output from the drone's flight control system in the reference direction. The pose angle of the drone and the estimated arrival angle of the acoustic wave were corrected relative to the ground. Figure 14 shows the record of flight posture and positional data collected from the time of takeoff until the drone lands after the completion of the measurement and maintains hovering during experiments measuring the sound of the explosion. The flight data showed the amount of change based on the average value of the hovering time interval. The time data also showed when the first explosion was detected at each explosion location from Locations 1 to 6. The remaining firecrackers exploded sequentially, seconds apart after the first explosion.

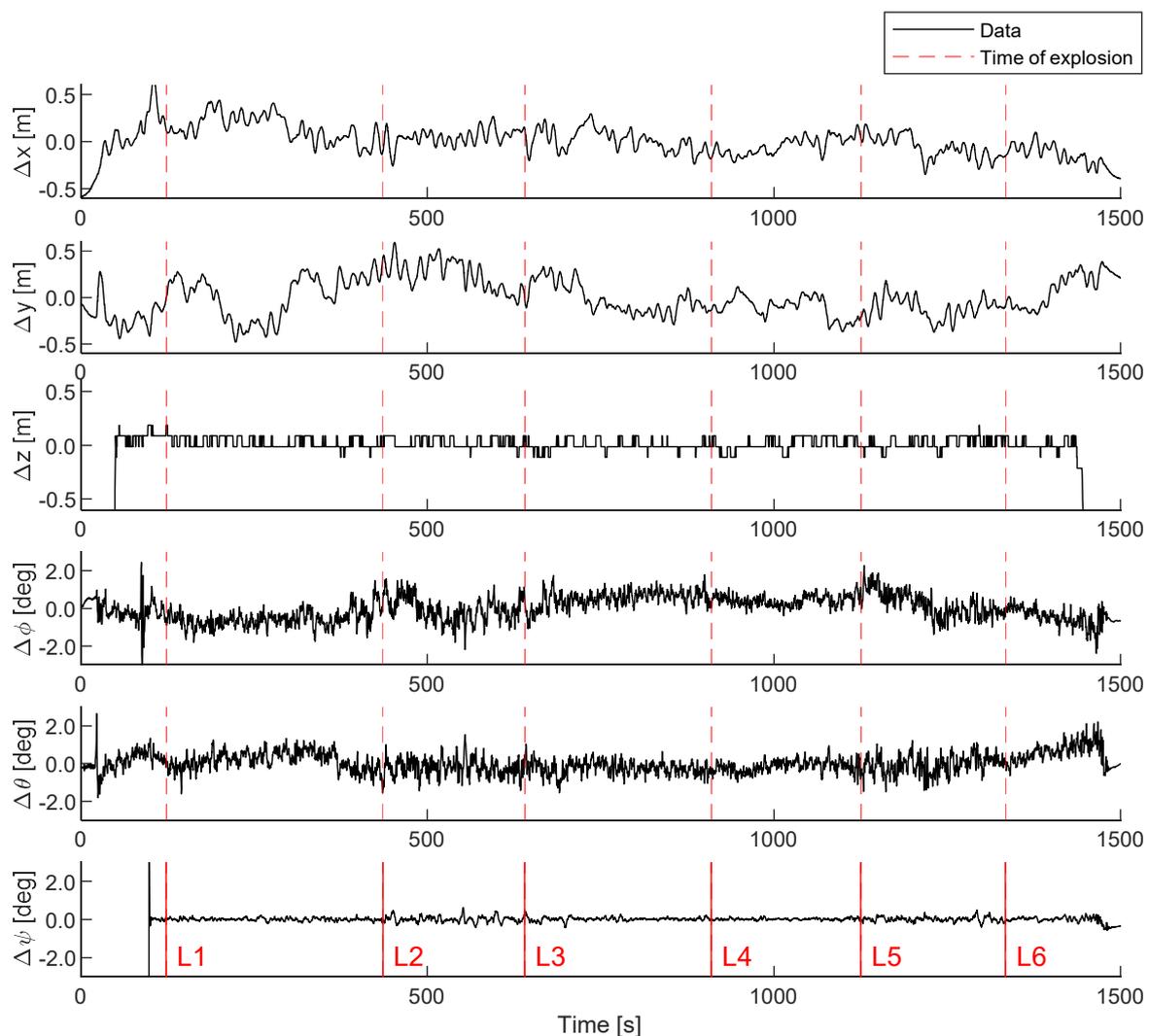


Figure 14. The amount of change in position and tilt angle during flight.

The estimated angle of arrival at the time an acoustic source was detected and the position and posture data of the drone could be corrected to specify the source location on the ground. Figure 15 shows the results of determining the estimated accuracy of the acoustic sources above the ground by the error of the horizontal and vertical angles. In this graph, the angle error of 0 degrees implies the true direction for the actual position based on GPS localization, and the measurement error represents the directional error angle that occurred from the true direction. In this experiment, acoustic source detection was performed up to a maximum ground distance of 151.5 m. The mean estimation error for three repeated impact sounds confirmed the detection performance of 8.8 degrees for the horizontal angle error and 10.3 degrees for the vertical angle error (Table 2). The estimated direction error for the entire location tended to increase both horizontally and vertically as the ground distance increased; the vertical angle error tended to be slightly higher than the horizontal angle error. The estimation error was the most sensitive during the drone's positional information correction process. In other words, we experimentally confirm that the location and posture of the drone at the time of detecting an acoustic source of interest must be applied to enable accurate localization.

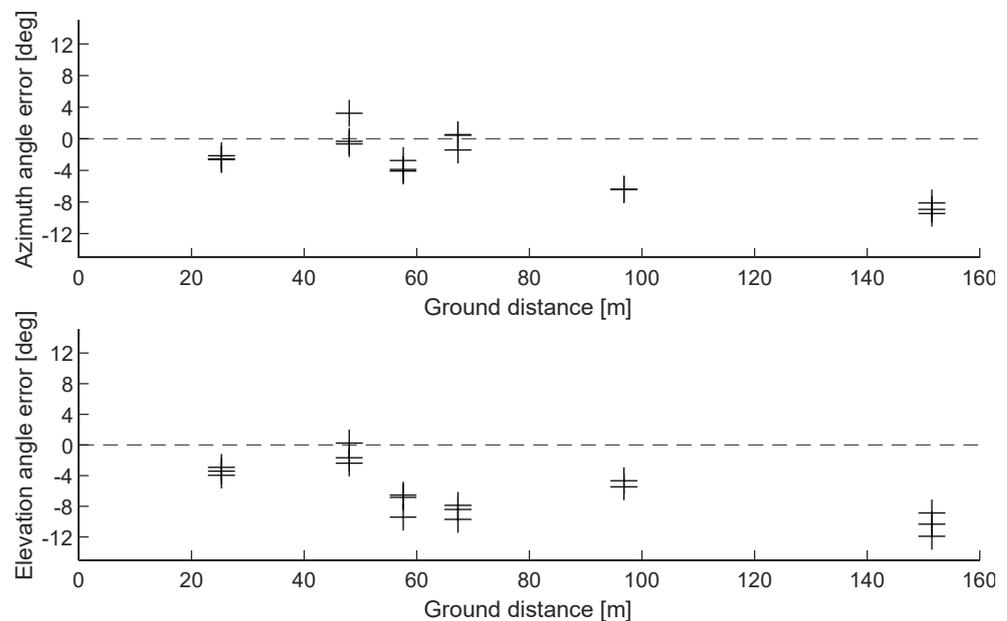


Figure 15. Estimated angle of arrival error for azimuth and elevation angles.

Table 2. Estimated angle of arrival error.

| Ground Distance (m) | Explosion Location | Mean Azimuth Angle Error (Degree) | Mean Elevation Angle Error (Degree) |
|---------------------|--------------------|-----------------------------------|-------------------------------------|
| 25.3 | L6 | 2.42 | 3.42 |
| 48.0 | L1 | 1.40 | 1.42 |
| 57.6 | L4 | 3.55 | 7.61 |
| 67.3 | L5 | 0.80 | 8.67 |
| 96.8 | L2 | 6.40 | 5.07 |
| 151.5 | L3 | 8.81 | 10.39 |

5. Conclusions

Using an array of microphones mounted on a drone, we described how to detect the location of acoustic sources generated on the ground and confirmed the performance of the system through experiments. The spectral subtraction method was applied to enhance the signal-to-noise ratio of the acoustic source of interest and was effective in eliminating noise generated by the drone. It was able to remove the fluctuating pressure and turbulence noise

affected by the blade passing frequency of drone propellers and the wake, and to preserve the impact sound that was intended to be detected. Furthermore, we confirmed that the direction of arrival could be estimated when applied to beamforming methods that detected the direction of arrival of acoustic sources using phase difference relationships, and that the phase difference relationship could be effectively restored even with spectral subtraction methods. It is important to highlight that the spectral subtraction method implemented in this study was able to improve the signal-to-noise ratio, and also to effectively preserve the phase for each microphone. A drone mounted with a 32-channel microphone array was used to detect acoustic sources, and flight data could be used to detect the locations on the ground. We implemented a valid acoustic source detection method by fusing flight data at the same time as when the acoustic source was detected. The detection performance of acoustic sources with a ground distance of 151.5 m was confirmed with a horizontal angle error of 8.8 degrees and a vertical angle error of 10.3 degrees.

In this study, we proposed and implemented the necessary elements to implement a method for exploring external sound sources of interest detected by drones. Although each element of the technique is not entirely state-of-the-art, it is meaningful to specify and implement what must be performed when detecting acoustic sources using a drone, and to describe the process of connecting them. Language recognition is a detection target that should immediately become of interest. It can be used to detect people or to strengthen the voice of a person targeted from the air. Drones, especially for lifesaving activities, can be a very useful technology. There have already been several studies on language recognition technology using deep learning, and in the near future, applications of this technology are expected. Deep learning can also be applied to noise cancellation technology for the drone's self-noise. From a hardware perspective, there should also be additional studies on optimal arrays that can effectively fit drones with fewer microphones to improve location detection performance.

Author Contributions: Conceptualization, Y.-J.G. and J.-S.C.; methodology, Y.-J.G.; software, Y.-J.G.; validation, Y.-J.G. and J.-S.C.; formal analysis, Y.-J.G.; investigation, Y.-J.G.; resources, Y.-J.G.; data curation, Y.-J.G.; writing—original draft preparation, Y.-J.G.; writing—review and editing, Y.-J.G. and J.-S.C.; visualization, Y.-J.G.; supervision, J.-S.C.; project administration, J.-S.C.; funding acquisition, J.-S.C. Both authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the research fund of Chungnam National University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Erdelj, M.; Natalizio, E.; Chowdhury, K.R.; Akyildi, I.F. Help from the Sky: Leveraging UAVs for Disaster Management. *IEEE Pervasive Comput.* **2017**, *16*, 24–32. [[CrossRef](#)]
2. Vattapparamban, E.; Güvenç, I.; Yurekli, A.I.; Akkaya, K.; Uluğaç, S. Drones for smart cities: Issues in cybersecurity privacy and public safety. In Proceedings of the International Wireless Communications and Mobile Computing Conference (IWCMC), Paphos, Cyprus, 5–9 September 2016; pp. 216–221. [[CrossRef](#)]
3. Ciaburro, G.; Iannace, G. Improving Smart Cities Safety Using Sound Events Detection Based on Deep Neural Network Algorithms. *Informatics* **2020**, *7*, 23. [[CrossRef](#)]
4. Ciaburro, G. Sound Event Detection in Underground Parking Garage Using Convolutional Neural Network. *Big Data Cogn. Comput.* **2020**, *4*, 20. [[CrossRef](#)]
5. Rascon, C.; Ivan, M. Localization of sound sources in robotics: A review. *Robot. Auton. Syst.* **2017**, *96*, 184–210. [[CrossRef](#)]
6. Liaquat, M.U.; Munawar, H.S.; Rahman, A.; Qadir, Z.; Kouzani, A.Z.; Mahmud, M.A.P. Localization of Sound Sources: A Systematic Review. *Energies* **2021**, *14*, 3910. [[CrossRef](#)]
7. Lee, J.; Go, Y.; Kim, S.; Choi, J. Flight Path Measurement of Drones Using Microphone Array and Performance Improvement Method Using Unscented Kalman Filter. *J. Korean Soc. Aeronaut. Space Sci.* **2018**, *46*, 975–985.

8. Hubbard, H.H. *Aeroacoustics of Flight Vehicles: Theory and Practice*; The Acoustical Society of America: New York, NY, USA, 1995; Volume 1, pp. 1–205.
9. Brooks, T.F.; Marcolini, M.A.; Pope, D.S. Main Rotor Broadband Noise Study in the DNW. *J. Am. Helicopter Soc.* **1989**, *34*, 3–12. [[CrossRef](#)]
10. Brooks, T.F.; Jolly, J.R., Jr.; Marcolini, M.A. Helicopter Main Rotor Noise: Determination of Source Contribution using Scaled Model Data. *NASA Technical Paper 2825*, August 1988.
11. Intaratep, N.; Alexander, W.N.; Devenport, W.J.; Grace, S.M.; Dropkin, A. Experimental Study of Quadcopter Acoustics and Performance at Static Thrust Conditions. In Proceedings of the 22nd AIAA/CEAS Aeroacoustics Conference, Lyon, France, 30 May–1 June 2016.
12. Djurek, I.; Petosic, A.; Grubesa, S.; Suhanek, M. Analysis of a Quadcopter’s Acoustic Signature in Different Flight Regimes. *IEEE Access* **2020**, *8*, 10662–10670. [[CrossRef](#)]
13. Boll, S. Suppression of acoustic noise in speech using spectral subtraction. *IEEE Trans. Acoust. Speech Signal Process.* **1979**, *27*, 113–120. [[CrossRef](#)]
14. Berouti, M.; Schwartz, R.; Makhoul, J. Enhancement of speech corrupted by acoustic noise. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Washington, DC, USA, 2–4 April 1979; pp. 208–211. [[CrossRef](#)]
15. Bharti, S.S.; Gupta, M.; Agarwal, S. A new spectral subtraction method for speech enhancement using adaptive noise estimation. In Proceedings of the 2016 3rd International Conference on Recent Advances in Information Technology (RAIT), Dhanbad, India, 3–5 March 2016; pp. 128–132. [[CrossRef](#)]
16. Yamashita, K.; Shimamura, T. Nonstationary noise estimation using low-frequency regions for spectral subtraction. *IEEE Signal Process. Lett.* **2005**, *12*, 465–468. [[CrossRef](#)]
17. Cohen, I.; Berdugo, B. Speech enhancement for non-stationary noise environments. *Signal Process.* **2001**, *81*, 2403–2418. [[CrossRef](#)]
18. Ramos, A.; Sverre, H.; Gudvangen, S.; Otterlei, R. A Spectral Subtraction Based Algorithm for Real-time Noise Cancellation with Application to Gunshot Acoustics. *Int. J. Electron. Telecommun.* **2013**, *59*, 93–98. [[CrossRef](#)]
19. Allen, C.S.; Blake, W.K.; Dougherty, R.P.; Lynch, D.; Soderman, P.T.; Underbrink, J.R. *Aeroacoustic Measurements*; Springer: Berlin/Heidelberg, Germany, 2002; pp. 62–215.
20. Chiariotti, P.; Martarelli, M.; Castellini, P. Acoustic beamforming for noise source localization—Reviews, methodology and applications. *Mech. Syst. Signal Process.* **2019**, *120*, 422–448. [[CrossRef](#)]
21. Michel, U. History of acoustic beamforming. In Proceedings of the 1st Berlin Beamforming Conference—1st BeBeC, Berlin, Germany, 22–23 November 2006.
22. Merino-Martínez, R.; Sijtsma, P.; Snellen, M.; Ahlefeldt, T.; Antoni, J.; Bahr, C.J.; Blacodon, D.; Ernst, D.; Finez, A.; Funke, S.; et al. A review of acoustic imaging methods using phased microphone arrays. *CEAS Aeronaut. J.* **2019**, *10*, 197–230. [[CrossRef](#)]
23. Van Trees, H.L. *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*; Wiley-Interscience: New York, NY, USA, 2002; pp. 231–331.
24. Clayton, M.; Wang, L.; McPherson, A.P.; Cavallaro, A. An embedded multichannel sound acquisition system for drone audition. *arXiv* **2021**, arXiv:abs/2101.06795.
25. Ruiz-Espitia, O.; Martínez-Carranza, J.; Rascón, C. AIRA-UAS: An Evaluation Corpus for Audio Processing in Unmanned Aerial System. In Proceedings of the International Conference on Unmanned Aircraft Systems (ICUAS), Dallas Marriott City Center Dallas, TX, USA, 12–15 June 2018; pp. 836–845.
26. Blanchard, T.; Thomas, J.H.; Raoof, K. Acoustic localization estimation of an Unmanned Aerial Vehicle using microphone array. *J. Acoust. Soc. Am.* **2020**, *148*, 1456. [[CrossRef](#)] [[PubMed](#)]
27. Strauss, M.; Mordel, P.; Miguet, V.; Deleforge, A. DREGON: Dataset and Methods for UAV-Embedded Sound Source Localization. In Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1–5 October 2018; pp. 1–8. [[CrossRef](#)]
28. Misra, P.; Kumar, A.A.; Mohapatra, P.; Balamuralidhar, P. DroneEARS: Robust Acoustic Source Localization with Aerial Drones. In Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, Australia, 21–25 May 2018; pp. 80–85. [[CrossRef](#)]
29. Salvati, D.; Drioli, C.; Ferrin, G.; Foresti, G.L. Acoustic Source Localization from Multirotor UAVs. *IEEE Trans. Ind. Electron.* **2020**, *67*, 8618–8628. [[CrossRef](#)]
30. Hoshiya, 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](#)]
31. Vaseghi, S.V. *Advanced Signal Processing and Digital Noise Reduction*; Springer: Berlin/Heidelberg, Germany, 1996; pp. 242–260.
32. Johnson, D.H.; Dudgeon, D.E. *Array Signal Processing, Concepts and Techniques*; PTR Prentice Hall: Englewood Cliffs, NJ, USA, 1993.
33. Frost, O.L., III. An algorithm for linear constrained adaptive array processing. *IEEE* **2009**, *60*, 925–935.
34. Bies, D.A.; Hansen, C.H. *Engineering Noise Control—Theory and Practice*, 4th ed.; CRC Press: New York, NJ, USA, 2009; pp. 18–19. ISBN 978-0-415-48707-8.

-
35. Hoffmann, G.M.; Huang, H.; Waslander, S.L.; Tomlin, C.J. Quadrotor helicopter flight dynamics and control: Theory and experiment. In Proceedings of the AIAA Guidance, Navigation and Control Conference and Exhibit, Boston, MA, USA, 10–12 August 1998.
 36. Huang, H.; Hoffmann, G.M.; Waslander, S.L.; Tomlin, C.J. Aerodynamics and control of autonomous quadrotor helicopters in aggressive maneuvering. In Proceedings of the IEEE International Conference on Robotics and Automation, Xi'an, China, 30 May–5 June 2021; pp. 3277–3282.
 37. Nelson, R.C. *Flight Stability and Automatic Control*, 2nd ed.; McGraw-Hill Education: New York, NY, USA, 1997; pp. 96–130.