Edge computing applications: using a linear MEMS microphone array for UAV position detection through sound source localization

Andrii V. Riabko[1], Tetiana A. Vakaliuk[2,3,4,5], Oksana V. Zaika[1], Roman P. Kukharchuk[1] and Valerii V. Kontsedailo[6]

[1]Oleksandr Dovzhenko Hlukhiv National Pedagogical University, 24 Kyivska Str., Hlukhiv, 41400, Ukraine
[2]Zhytomyr Polytechnic State University, 103 Chudnivsyka Str., Zhytomyr, 10005, Ukraine
[3]Institute for Digitalisation of Education of the NAES of Ukraine, 9 M. Berlynskoho Str., Kyiv, 04060, Ukraine
[4]Kryvyi Rih State Pedagogical University, 54 Universytetskyi Ave., Kryvyi Rih, 50086, Ukraine
[5]Academy of Cognitive and Natural Sciences, 54 Universytetskyi Ave., Kryvyi Rih, 50086, Ukraine
[6]Inner Circle, Nieuwendijk 40, 1012 MB Amsterdam, Netherlands

Abstract

This study explores the use of a microphone array to determine the position of an unmanned aerial vehicle (UAV) based solely on the sound of its engines. The accuracy of localization depends crucially on the arrangement of the microphones. The study also considers a mathematical model of pulse density modulation for a digital MEMS microphone. It demonstrates the frequency dependence of the efficiency of a differential array of first-order microphones. Based on this frequency dependence of directivity and the instability model of the microphone parameters, a rational operating frequency range for the normal functioning of the microphone array can be established. The study proposes a model of a linear microphone array based on MEMS omnidirectional microphones. With a specific geometrical arrangement, this array produces a bidirectional pattern, which can be easily transformed into a unidirectional pattern using specialized algorithms or hardware (e.g., ADAU1761 codecs).

Keywords

edge computing, UAV, sound source localization, MEMS microphone, microphone array, frequency, directivity

1. Introduction

Determining the position of a UAV (Unmanned Aerial Vehicle) by the sound of its engines can be important for several reasons. In military or security applications, being able to identify and locate UAVs by their engine sounds can help in detecting potential threats, including hostile drones or unauthorized surveillance. Sound-based localization can aid in the development of countermeasures to mitigate the risks posed by UAVs in sensitive areas.

Sound-based UAV detection can complement existing air traffic management systems, providing additional situational awareness for managing airspace and preventing collisions with manned aircraft. In search and rescue operations or in case of lost or malfunctioning drones, sound-based tracking can assist in locating and recovering UAVs.

In conservation efforts, it can help monitor UAVs used for illegal activities like poaching or wildlife disturbance. In urban areas or regions with dense UAV traffic, sound-based tracking can be useful for enforcing regulations related to UAV flight paths, altitudes, and no-fly zones. For protecting privacy,

doors-2024: 4th Edge Computing Workshop, April 5, 2024, Zhytomyr, Ukraine
ryabko@meta.ua (A. V. Riabko); tetianavakaliuk@acnsci.org (T. A. Vakaliuk); ksuwazaika@gmail.com (O. V. Zaika); kyxap4yk1@ukr.net (R. P. Kukharchuk); valerakontsedailo@gmail.com (V. V. Kontsedailo)
http://irbis-nbuv.gov.ua/ASUA/0051396 (A. V. Riabko); https://acnsci.org/vakaliuk/ (T. A. Vakaliuk); http://pfm.gnpu.edu.ua/index.php/struktura1/2015-04-01-14-50-26 (O. V. Zaika); http://irbis-nbuv.gov.ua/ASUA/0076404 (R. P. Kukharchuk)
0000-0001-7728-6498 (A. V. Riabko); 0000-0001-6825-4697 (T. A. Vakaliuk); 0000-0002-8479-9408 (O. V. Zaika); 0000-0002-7588-7406 (R. P. Kukharchuk); 0000-0002-6463-370X (V. V. Kontsedailo)

sound-based detection can help identify UAVs flying near private properties, providing a means to take legal action against intrusive drones.

Studying the acoustic signatures of UAVs can aid in research and development efforts to design quieter and more environmentally friendly drones. During natural disasters or emergencies, knowing the positions of UAVs, such as those used for aerial surveys or damage assessment, can assist in coordinating response efforts. Sound-based UAV detection can be employed in border control to monitor and respond to unauthorized drone incursions.

As drone delivery and urban air mobility concepts develop, sound-based localization can contribute to managing UAV traffic in urban environments. While sound-based UAV localization offers several advantages, it also has limitations, such as accuracy challenges in noisy environments and the need for specialized equipment. Therefore, it is often used in conjunction with other tracking and detection methods, such as radar, visual recognition, and GPS, to provide comprehensive situational awareness and enhance safety and security in various applications.

The goal of our work is to develop a software and hardware system for capturing hardware-synchronized sound using digital MEMS microphones (Microelectromechanical Systems, MEMS) for further use in sound source localization systems. This system is intended for further use in sound source localization systems, marking a significant advancement in the field of edge computing.

2. Theoretical background

Over the past few decades, acoustic source localization has emerged as a focal point of interest within the research community [1, 2]. Most studies of sound source identification are based on the analysis of the physiological mechanism of human hearing [3, 4]. It is common practice to use arrays of microphones [5]. An actual problem is acoustic beam formation for sound source localization and its application [6].

Microphone array processing represents a well-established methodology employed in the estimation of sound source direction. In a groundbreaking contribution by Yamada et al. [7], they introduce an innovative approach referred to as Multiple Triangulation and Gaussian Sum Filter Tracking (MT-GSFT). This advanced technique adeptly derives the precise location of sound sources through triangulation, utilizing microphone arrays seamlessly integrated into a fleet of multiple drones [7]. The domain of speech signal processing encompasses several critical areas, and among them, multiple sound source localization (SSL) stands out as a notable and relevant field. A notable contribution to this field comes from Firoozabadi et al. [8], who introduced a two-step approach for the localization of multiple sound sources in three dimensions (3D). This method relies on the precise estimation of time delays (TDE) and strategically leverages distributed microphone arrays (DMA) to enhance the accuracy and effectiveness of the localization process [8].

Sasaki et al. [9] present a method designed to map the 3D coordinates of a sound source by leveraging data gathered from an array of microphones, with each microphone providing an autonomous directional estimate. Additionally, LiDAR technology is employed to create a comprehensive 3D representation of the surroundings and accurately determine the sensor's position with six degrees of freedom (6-DoF).

Catalbas et al. [10] conduct a comparative analysis, assessing the effectiveness of generalized cross-correlation techniques in contrast to noise reduction filters concerning the estimation of sound source trajectory. Throughout the entire movement, they calculate the azimuth angle between the sound source and the receiver. This calculation relies on the parameter of Interaural Time Difference (ITD) to determine the azimuth angle. They then evaluate the accuracy of the estimated delay using various types of Generalized Cross-Correlation (GCC) algorithms for comparison.

It is possible for unmanned aerial vehicles (UAVs) to use audio information to compensate for poor visual information. Hoshiba et al. [11] developed a microphone array system built into the UAV to localize the sound source in flight. They developed the Spherical Microphone Array System (SMAS), consisting of a microphone array, a stable wireless network communication system, and intuitive visualization tools.

Tachikawa et al. [12] introduced an innovative approach that involves estimating positions by utilizing

a modified variant of the convex clustering method in conjunction with sparse coefficients estimation. Additionally, they put forth a technique for constructing a well-suited monopole dictionary, which is based on coherence, ensuring that the convex clustering-based method can accurately estimate the distances of sound sources. The study involved conducting a series of numerical and measurement experiments aimed at assessing the effectiveness and performance of this novel methodology.

When dealing with multiple sound sources, establishing a reliable data association between localization information and the corresponding sound sources becomes paramount for achieving optimal performance. To address the challenges posed by data association uncertainty, Wakabayashi et al. [13] extended the Global Nearest Neighbor (GNN) approach, introducing a modified version known as GNN-c, specifically tailored to meet the real-time and low-latency requirements of drone audio applications. The outcome of their efforts showcases a system capable of accurately estimating the positions of multiple sound sources, achieving an impressive accuracy level of approximately 3 meters.

Many acoustic image-based sound source diagnosis systems suffer from spatial stationary limitations, making it challenging to integrate information from various capture positions, thereby leading to unreliable and incomplete diagnostics. In their paper, Carneiro and Berry [14] introduce a novel measurement methodology called Acoustic Imaging Structure From Motion (AISFM). This approach utilizes a mobile spherical microphone array to create acoustic images through beamforming, seamlessly integrating data from multiple capture positions. Their method is not only proposed but also meticulously developed and rigorously validated, offering a promising solution to enhance the accuracy and comprehensiveness of sound source diagnostics.

In a research conducted by Kita and Kajikawa [15] a sound source localization (SSL) technique is introduced, specifically designed for the localization of sources situated within structures, including mechanical equipment and buildings.

The registration of acoustic signals with cross-shaped antennas is widely discussed in the literature [16].

Advanced signal processing methods involving multiple microphones can enhance noise resilience. However, as the quantity of microphones employed escalates, the computational overhead rises concomitantly. This, in turn, curtails response time and hinders their extensive adoption across various categories of mobile robotic platforms [17]. Within the realm of robot audition, sound source localization (SSL) holds a pivotal role, serving as a fundamental component. SSL empowers a robotic platform to pinpoint the origin of sound using auditory cues exclusively. Its significance extends beyond mere sound localization, as it significantly influences other facets of robot audition, including source separation. Moreover, SSL contributes to elevating the quality of human-robot interaction by augmenting the robot's perceptual prowess [18].

In general, machine learning is widely used in acoustics [19, 20]. In the realm of human-robot interaction, He et al. [21] have introduced a pioneering approach. Their proposal involves harnessing neural networks for the simultaneous detection and localization of multiple sound sources. This innovative method represents a departure from conventional signal processing techniques by offering a distinct advantage: it necessitates fewer stringent assumptions about the environmental conditions, thereby enhancing its adaptability and effectiveness [21]. Ebrahimkhanlou and Salamone [22] have put forth an advanced methodology for localizing acoustic emissions (AE) sources within metallic plates, especially those with intricate geometric features like rivet-connected stiffeners. This innovative approach leverages two deep learning techniques: a stack of autoencoders and a convolutional neural network (CNN), strategically employed to enhance the accuracy and precision of the localization process [22].

In their pioneering work, Adavanne et al. [23] have introduced an innovative solution – a convolutional recurrent neural network (CRNN) – designed to address the intricate task of joint sound event localization and detection (SELD) within three-dimensional (3-D) space. This method represents a significant advancement in the field, enabling the simultaneous identification and spatial localization of multiple overlapping sound events with remarkable precision.

Let's summarize the theoretical review. Localizing a sound source means determining the direction or location from which a sound is emanating. There are several algorithms and techniques used for

sound source localization, and the choice of method often depends on the specific application and available hardware. Here are some commonly used algorithms. Time Difference of Arrival (TDOA) is based on measuring the time it takes for a sound to reach multiple microphones. By comparing the time differences, it's possible to triangulate the source's position. Cross-correlation or beamforming techniques are often used to calculate the time differences accurately. Generalized Cross-Correlation (GCC) is a technique used in conjunction with TDOA. It involves cross-correlating the signals from two or more microphones to find the delay between them. GCC-PHAT (GCC with Phase Transform) is a commonly used variant that works well in reverberant environments. Steering Vector Methods are commonly used in microphone arrays or beamforming applications [24]. They estimate the direction of arrival (DOA) by analyzing the phase differences between signals received by different microphones. Popular algorithms include Multiple Signal Classification (MUSIC) and Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT). Acoustic Intensity Methods measure the sound intensity at multiple microphone positions and use this information to estimate the source direction. The Steered Response Power (SRP) algorithm is an example of this approach. Machine learning and deep learning techniques, such as neural networks and support vector machines, can be used to train models for sound source localization. These models can take input from multiple microphones and learn to predict the source location based on training data. Particle filtering is a probabilistic method that estimates the source location using a Bayesian filtering approach. It is useful when dealing with complex and dynamic environments. Some methods use time-frequency analysis techniques like the Short-Time Fourier Transform (STFT) or Wavelet Transform to analyze the spectral content of audio signals and infer the source location. In mobile sound source localization, the Doppler effect can be used to estimate the source's speed and direction based on the frequency shift in the received signal. Many practical systems use a combination of the above techniques to improve accuracy and robustness, especially in real-world scenarios with noise and reverberation.

The choice of algorithm depends on factors like the number and arrangement of microphones, environmental conditions, computational resources, and the desired level of accuracy. Different applications, such as robotics, audio conferencing, surveillance, and hearing aids, may employ different algorithms tailored to their specific requirements.

Determining the position of a UAV (Unmanned Aerial Vehicle) based solely on the sound of its engines can be challenging but is feasible using a combination of sound source localization techniques and signal processing. Here's a high-level overview of the process:

1. Microphone Array Setup: Set up a microphone array on the ground. The microphones should be strategically placed to capture the UAV's sound from different angles. The arrangement of microphones plays a crucial role in accurate localization. The response of microphone arrays depends, first of all, on the number of microphones working on the array [25].
2. Sound Data Collection: Record the sound generated by the UAV's engines as it flies overhead. Ensure that the recording system has a high sampling rate to capture the sound accurately.
3. Time Delay of Arrival (TDOA): Analyze the recorded audio data to calculate the time delay of arrival (TDOA) of the sound at each microphone. TDOA is the time difference between when the sound reaches different microphones. This information is critical for triangulation.
4. Triangulation: Use the TDOA data from multiple microphones to triangulate the UAV's position. Several algorithms, such as multilateration or beamforming, can help estimate the UAV's coordinates based on the TDOA information.
5. UAV Sound Signature: To improve accuracy, consider using machine learning techniques to create a database of UAV sound signatures. This involves training a model to recognize the unique sound characteristics of different UAVs. When a new sound recording is obtained, the model can help identify the specific UAV type.
6. Integration with Other Sensors: For real-time tracking, integrate sound-based localization with other sensors like GPS, radar, or visual cameras. This fusion of data sources can provide more accurate and robust positioning.

7. Calibration and Testing: Regularly calibrate and test the microphone array and signal processing algorithms to ensure accurate and reliable results.

It's important to note that the accuracy of sound-based UAV localization depends on various factors, including the UAV's altitude, speed, engine type, and background noise. Additionally, environmental conditions, such as wind and temperature, can affect sound propagation and localization accuracy. Therefore, this method may work best in controlled environments or in conjunction with other tracking methods for enhanced precision and reliability.

3. Research methods

The goal of our work is to develop a software and hardware system for capturing hardware-synchronized sound using digital MEMS microphones (Microelectromechanical Systems, MEMS) for further use in sound source localization systems.

Despite the fact that the use of radar equipment has become part of everyday practice when monitoring UAVs, there is some interest in assessing the possibility of using airborne acoustic signals for this purpose. The above applies mainly to receiving hydroacoustic antennas, i.e. to conditions when the speed of the source is much lower than the speed of sound $M = v/c \ll 1$. In the case of receiving air-acoustic signals propagating at a speed of sound significantly lower than that of hydroacoustic signals in water and created by fairly fast moving sources (passenger cars on autobahns, racing cars, the movement of airliners along runways during takeoff and landing, UAVs), there is a different situation. Research into the features of recording these signals with phased arrays remains relevant, since on their basis data can be obtained on the current coordinates and speed of movement of a moving object. The purpose of this work is to analyze the angular dependencies in the signal at the output of a receiving air-acoustic antenna and those qualitative changes in their nature that are introduced due to a combination of factors such as the Doppler effect and the sharp directivity of the antenna array.

A special case is considered, which is widespread in everyday practice, when the trajectory of an object is rectilinear, lies in a horizontal plane, close and parallel to the Earth's surface, and the speed of its movement is constant.

As previously stated, the arrangement of microphones plays a crucial role in accurate localization. The purpose of the study is to find the optimal configuration of a microphone array for localizing a moving sound source (UAV).

Directivity is the sensitivity of a microphone to sound depending on the direction or angle from which the sound is coming. Directionality or sound pickup angle is considered to be the area of possible location of the sound signal source, within which there is no significant loss of microphone efficiency. Microphones use different directivity characteristics. They are most often depicted as polar diagrams. This is done to graphically display sensitivity variations around the microphone over a 360-degree range, where the microphone is the center of the circle and the angular reference point is placed in front of the microphone. The polar pattern shows how a microphone's sensitivity to a sound signal depends on the location of its source.

Microphone arrays are an array of several microphones combined by joint digital signal processing. Microphone arrays provide the following advantages over single-channel systems: 1) directionality of sound reception; 2) noise suppression of point sources; 3) suppression of non-stationary environmental noise; 4) partial weakening of reverberation; 5) the possibility of spatial localization of the sound source; 6) the ability to accompany a moving point sound source.

A microphone array is one of the types of directional microphones, implemented as a set of sound receivers operating in concert (in phase or with certain phase delays). Geometrically, gratings can be implemented in different configurations – one-dimensional (linear, arc-shaped), two-dimensional (flat, spherical), three-dimensional, spiral, with uniform or non-equidistant pitch. The array's radiation pattern is created by changing the ratio of phase delays for different channels (in the simplest case, an in-phase array with a fixed position of the main lobe; in more complex and expensive implementations,

a scanning system). The implementation of phase delays can be hardware (for example, on analog delay lines) or software (digital).

The basic microphone array structures are Broadside and Endfire (figure 1).

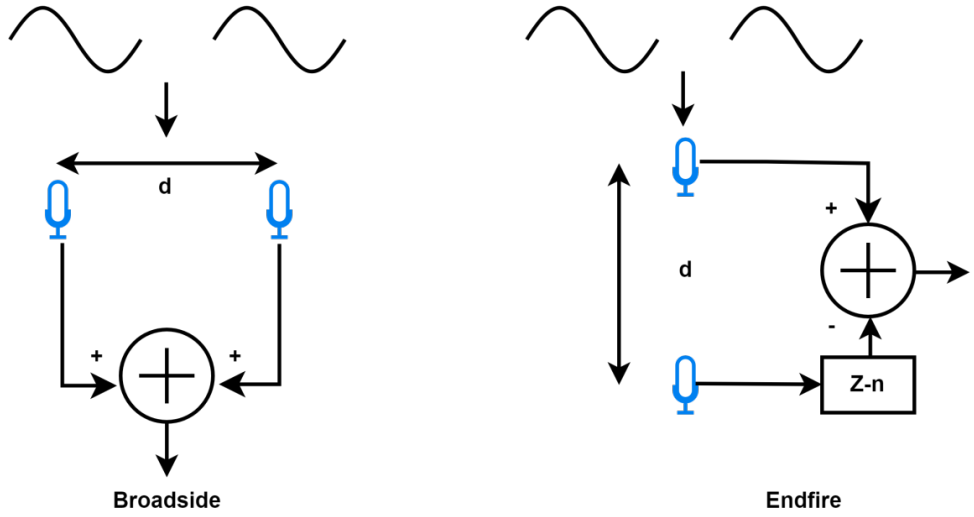

Figure 1: Basic array structures.

These structures use omnidirectional microphones (microphones that, regardless of their orientation, receive signals from any direction). The figure 2 shows signal reception versus direction for various frequencies with a single omnidirectional microphone. For one microphone, frequency invariance is observed.

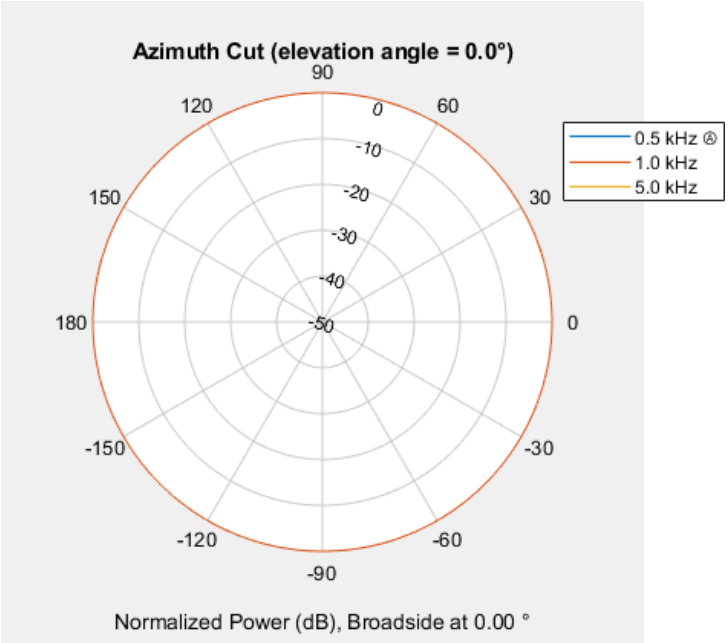

Figure 2: Dependence of signal reception on direction by one omnidirectional microphone for frequencies 500 Hz, 1 and 5 kHz.

The Broadside structure is an array of omnidirectional microphones positioned perpendicular to the direction of the desired signal. Such arrays have an axis of symmetry, relative to which the sound is released without attenuation both “in front” of the array and “behind”. Such structures are widely used in applications where sound pressure waves enter the sensor array from one side. Consider a Broadside structure consisting of two microphones spaced 7.5 cm apart. The minimum response is

observed when the signal is incident at an angle of 90° or 270° (in this case, the angle between the direction of the useful signal and the normal to the line of elements is taken as 0°). But this response strongly depends on the frequency of the received signal. Theoretically, such a system has a perfect zero at a frequency of 2.3 kHz. Above this frequency, depending on the direction of arrival, there are zeros at other angles (figure 3). The microphone array shows a clear directional characteristic at 4 kHz, and at 1 kHz its pattern is essentially omnidirectional. As a result, at lower frequencies the array cannot achieve significant spatial filtering.

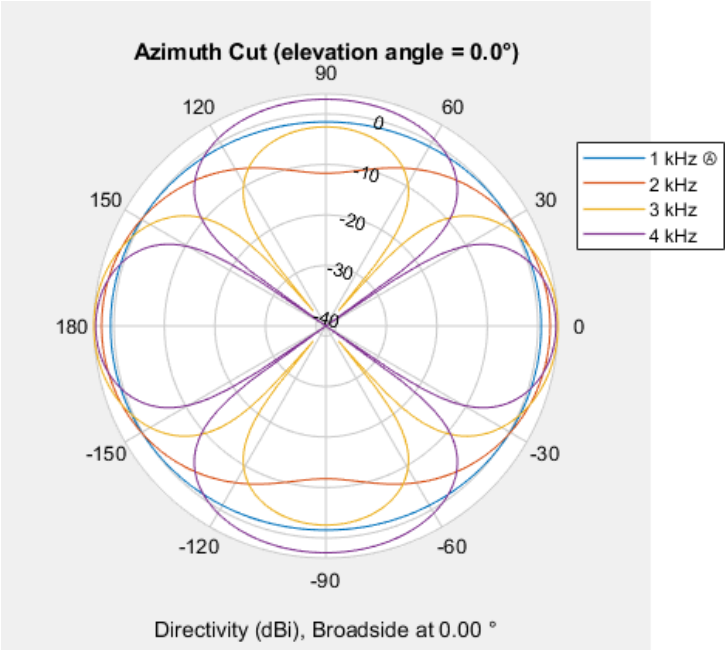

Figure 3: Dependence of signal reception on direction by a Broadside structure of two omnidirectional microphones for frequencies of 1 kHz, 2 kHz, 3 kHz and 4 kHz.

The Endfire structure consists of several microphones located in the direction of the useful acoustic signal. This design is called a differential array of microphones. The delayed signal from the first microphone is summed with the signal from the next microphone. To create a cardioid polar pattern, the signal from the rear microphones must be delayed by the same amount of time that the sound waves travel between the two microphone elements. Such structures are used to produce cardioid, hypercardioid or supercardioid directional response and theoretically completely eliminate sound incident on the array at an angle of 180°. A unidirectional microphone is more sensitive to sound coming from one direction and less sensitive to sounds from other directions. The most typical for such microphones is the cardioid characteristic, representing a peculiar diagram in the shape of a heart. At the same time, the peak of sensitivity is reached in the direction along the axis of the microphone, and the decline is in the opposite direction (figure 4).

To generate a cardioid response in direction, the signal from the omnidirectional microphones must be delayed for a time equal to the propagation of the acoustic wave between the two elements. Developers of such systems have two degrees of freedom to change the output signal of the speaker system: changing the distance between microphones and changing the delay time. Figure 5 shows the signal reception versus direction for various frequencies by the Endfire structure with two elements and a distance between them of 2.1 cm.

The distance between the microphones is crucial for the formation of a cardioid response. Figure 6 shows the same microphones, but placed at a distance of 15 cm.

The structures considered have the following advantages and disadvantages. Advantages of Broadside: flat geometry, simple processing implementation, ability to control the direction of the beam. Disadvantages of the Broadside: less off-axis rejection, close microphone spacing, and a large number of microphones needed to prevent spatial leakage.

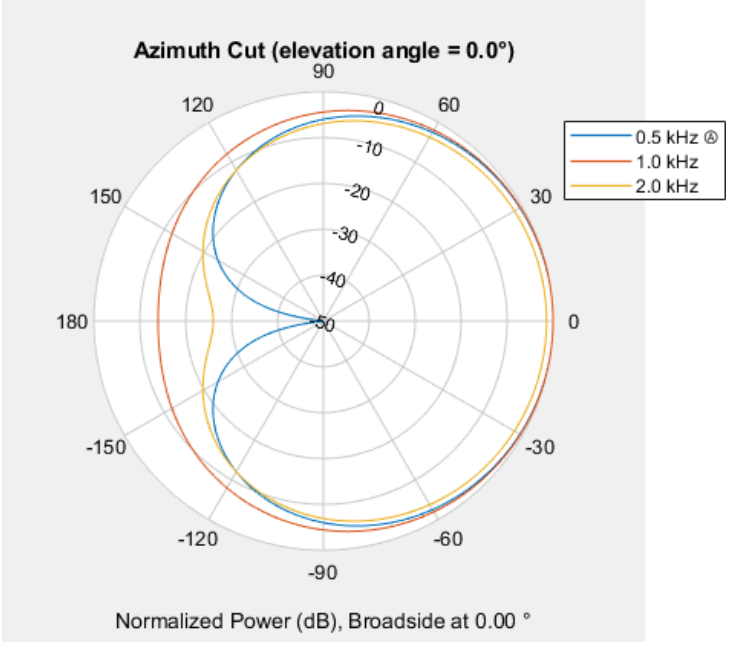

Figure 4: Dependence of signal reception on a unidirectional microphone for frequencies of 500 Hz, 1 kHz and 2 kHz.

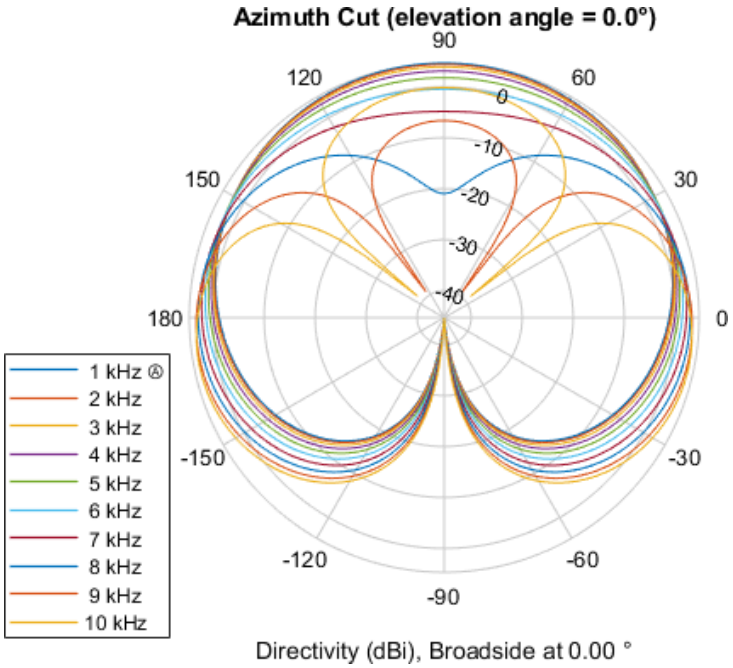

Figure 5: Dependence of signal reception on direction by an Endfire structure of two omnidirectional microphones for frequencies from 1 to 10 kHz, which are located at a distance of 21 cm.

Advantages of Endfire: Better off-axis suppression, smaller overall size. Disadvantages of Endfire: non-flat (volumetric) geometry, more complex processing, suppression of the useful signal in the low frequency range, the direction of the source of the useful signal must coincide with the axis of the microphone array; For two-dimensional gratings, beam formation is possible only in the horizontal direction (the grating array).

To form a differential array of higher orders, you need to add additional microphones. Since the petals will deviate more back and to the side in the directional diagram, the distance between the microphones will have to be increased. The figure 7 shows an array of 4 microphones (third order), which forms

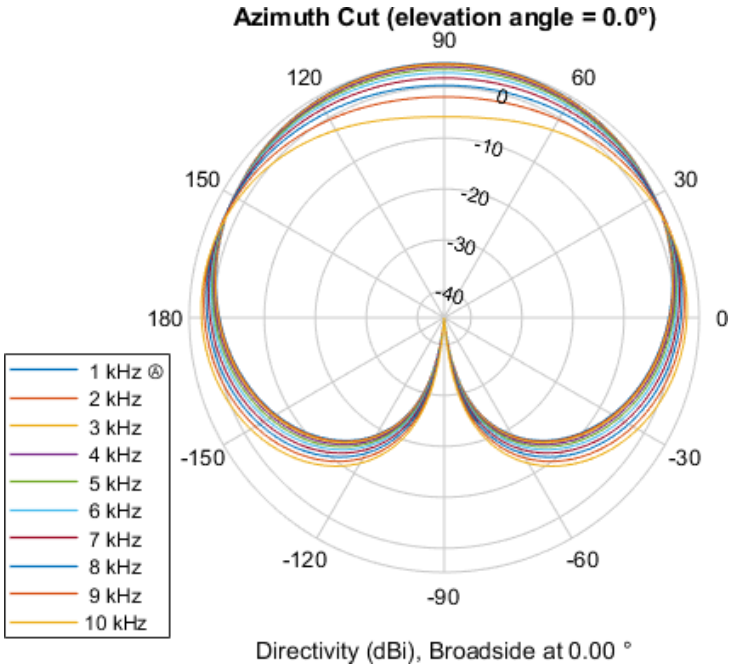

Figure 6: Dependence of signal reception on direction by an Endfire structure of two omnidirectional microphones for frequencies from 1 to 10 kHz, which are located at a distance of 15 cm.

a supercardioid pattern. Consider how beam formation depends on the number of microphones and the distance between them. It is worth noting that the sensitivity and frequency response of all array microphones must be precisely matched.

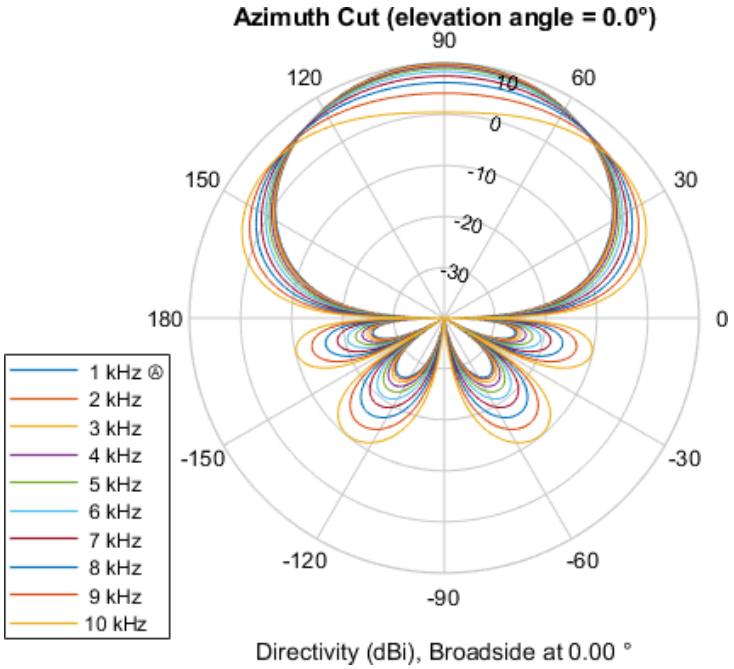

Figure 7: Differential array of 4 microphones (third order).

Differential microphone arrays make it possible to obtain high directivity characteristics of the system with its small size. But with such a construction, the problem arises of a significant change in the characteristics of the entire system with a slight deviation of the parameters of an individual microphone from its nominal values. If this approach is used for critical applications, measures must be taken to reduce deviations of microphone parameters from nominal values. As stated earlier, a first-order

differential microphone array consists of two omnidirectional sensors separated by d (figure 8).

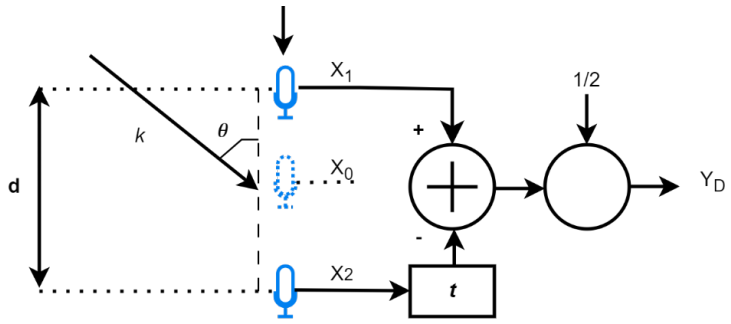

Figure 8: Structure of a first order differential microphone array.

4. Results

When sound arrives from the main direction $\theta = 0$, a delay appears between these sensors:

$$\tau_D = \frac{d}{c}, \tag{1}$$

where c is the speed of sound.

A plane wave, which is characterized by wave number $\overrightarrow{k}$, arrives at the input of the differential grating. Due to radial symmetry, the output signals of sensors $X_1(\omega)$ and $X_2(\omega)$ can be expressed by a function depending on the angle θ and frequency ω. There is a relationship $|\overrightarrow{k}|d = kd = \omega\tau_D$ between the wave number and the frequency of the signal. At the central point of the array, you can place a virtual microphone with an output signal $X_0(\omega)$. A plane wave incident at an angle θ with wave number causes $k = 2\pi/\lambda$ the appearance of signals at the output of microphones X_1 and X_2:

$$X_1(\omega) = X_0(\omega)e^{j\frac{kd}{2}\cos\theta},\ X_2(\omega) = X_0(\omega)e^{-j\frac{kd}{2}\cos\theta}. \tag{2}$$

At the output of the differential lattice we get

$$Y_D(\omega) = \frac{1}{2}(X_1(\omega) - X_2(\omega)e^{j\omega\tau}. \tag{3}$$

The directivity function of the differential array H_D is the ratio of the signal at the output of the array $Y_D(\omega)$ to the signal at the output of the virtual microphone $X_0(\omega)$:

$$H_D(\omega,\theta) = je^{-j\frac{\omega\tau}{2}}\sin\left(\frac{kd}{2}\left(\frac{\tau}{\tau_D} + \cos\theta\right)\right). \tag{4}$$

Usually very small values of $kd \ll 1$ are considered, which makes it possible to use the approximation $\sin\alpha \approx \alpha$. In this case, the idealized directivity function H_D has the form:

$$H_D(\omega,\theta) \approx \widetilde{H}_D(\theta) = j\frac{kd}{2}\left(\frac{\tau}{\tau_D} + \cos\theta\right). \tag{5}$$

With this view, the main characteristics of differential microphone arrays are obvious: 1) the form of $\widetilde{H}_D(\theta)$ is determined by the expression $\tau/\tau_D + \cos\theta$, which does not depend on frequency; 2) due to the subtraction of the signal a phase shift occurs by $\pi/2$; 3) the frequency response of the directivity function $H_D(\omega)$ has the form of a first-order high-pass filter.

At low frequencies, the output signal $Y_D(\omega)$ becomes highly susceptible to any changes in the shape of the characteristic $H_D(\omega)$. For this reason, the distance d should not be chosen too small, which may lead to a conflict with the condition $kd \ll 1$.

The exact expression for the directivity function (4) contains a sine function that scales the amplitude. It is rational to limit the operating range of the differential grating in the low frequency range to the first maximum of the sine. This first maximum fixes the cutoff frequency ωc:

$$\omega_c = \frac{\pi}{\tau_D + \tau}. \tag{6}$$

For low frequencies, the directivity characteristics are practically independent of frequency. However, as the frequency increases, the shape of the frequency response becomes more and more deformed. In addition, at some frequencies the signal is completely suppressed.

In order to compensate for the high-frequency nature of the behavior $H_D(\omega, \theta)$ it is necessary to develop a filter $W_{eq}(\omega)$. For the main direction $\theta = 0$, the adjusted frequency response$H_D(\omega, \theta = 0)W_{eq}(\omega)$ must be constant and equal to 0 dB, and for frequencies below ω_c:

$$W_{eq}(\omega) = \begin{cases} \frac{1}{\sin\left(\frac{\pi\omega}{2\omega_c}\right)}, & 0 < \omega < \omega_c, \\ 1, & \text{in other cases.} \end{cases} \tag{7}$$

For low frequencies $\omega \to 0$, the filter gain W_{eq} has very large values. This means that any noise present in the input signal will be greatly amplified. The level of this noise is determined by the specific sensor. This circumstance limits the frequency range of the signal for processing using a differential microphone array.

The directional properties of a microphone array are characterized by the directional coefficient (DI). It can be expressed as the ratio of the squared modulus of the directivity function in the main direction to the average value of the squared modulus in all directions:

$$DI(\omega) = \frac{|H(\omega, \theta = 0)|^2}{\frac{1}{4\pi}\int\limits_0^{2\pi}\int\limits_0^{\pi} |H(\omega, \theta)|^2 \sin\theta d\theta d\varphi}. \tag{8}$$

Taking into account the exact expression for the directivity function (4), we can obtain a new expression for the dependence of the directivity on frequency:

$$DI_D(\omega) = \frac{2sin^2\left(\frac{\omega}{2}\left(\tau_D + \tau\right)\right)}{1 - si\left(\omega\tau_D\right)\cos\left(\omega\tau\right)} \tag{9}$$

where $si(x) = \frac{1}{x}\sin(x)$.

The efficiency factor for low frequencies is obtained similarly to the result of approximation $\widetilde{H}_D$ according to expression (5):

$$\lim_{\omega\to 0} DI_D(\omega) = \widetilde{H}_D = \frac{3(\tau_D + \tau)^2}{3\tau_D^2 + 3\tau^2}. \tag{10}$$

Let us study the influence of microphone parameter mismatch for first-order differential arrays. We use a model of instability of microphone parameters in the form of a transfer function $M = M_{ref} + \Delta M$. The nominal transfer function of the sensor M_{ref} in this case is normalized to the value 1. It is assumed that the deviation ΔM is an independent random variable with variance:

$$\sigma_M^2 = E\{|\Delta M|^2\}, \tag{11}$$

where $E\{|\Delta M|^2\}$ expectation operator. Signals from two sensors in figure 8 will then be written as follows:

$$\hat{X}_1(\omega) = X_0(\omega)(1 + \Delta M_1)e^{j\frac{kd}{2}\cos\theta},\ \hat{X}_2(\omega) = X_0(\omega)(1 + \Delta M_2)e^{-j\frac{kd}{2}\cos\theta}. \tag{12}$$

The directivity function $\hat{H}_D$ for a differential array, taking into account the instability of the microphone parameters, can be obtained similarly to expression (4). But now there are additional conditions

that depend on ΔM_i $(i = 1, 2)$. For random numbers, the quadratic terms remain, and the linear ones are set to zero, so we get:

$$E\{|\hat{H}_D(\omega, \theta)|^2\} = |H_D(\omega, \theta)|^2 + 2\sigma_M^2. \tag{13}$$

As a result, we can obtain a modified expression for DI:

$$E\{|\hat{DI}_D(\omega)|\} = \frac{2\sin^2\left(\frac{\omega}{2}\left(\tau_D + \tau\right)\right) + \sigma_M^2}{1 - si\left(\omega\tau_D\right)\cos\left(\omega\tau\right)\sigma_M^2} \tag{14}$$

It is important to understand that in expression (13) the efficiency factor $H_D(\omega, \theta)$ characterizes the behavior of the system at high frequencies. While the W_{eq} equalization filter takes into account the effects of microphone instability and enhances them for low frequencies.

Thus, this work shows the dependence of the efficiency of a differential array of first-order microphones on frequency. It can be supplemented using a model of instability of microphone parameters at low frequencies.

Based on the presented dependence of the directivity on frequency and the instability model of the microphone parameters, a rational operating frequency range for the normal functioning of the microphone array can be determined. The lower limit of this range is limited by the instability of the microphone parameters, and the upper cutoff frequency is determined by the geometry of the array d.

Currently, there are many applications in which acoustic signals are processed. Microelectromechanical microphones (MEMS) are increasingly being used for these purposes. The use of such microphones allows the construction of differential microphone arrays. Microelectromechanical systems (MEMS) are a variety of microdevices of a wide variety of designs and purposes, in the production of which modified microelectronics technological techniques are used. Typically, all elements of such systems are placed on a common silicon base, the size of which is only a couple of millimeters. A MEMS microphone is an electro-acoustic device for converting sound vibrations into electrical waves, which is small enough to be installed in a tightly integrated product, for example: a smartphone, headset, speakerphone, laptop or any other device. There are two fundamentally important elements in such microphones: an integrated circuit (ASIC) and a MEMS sensor. It is the latter that ensures the capture and subsequent transmission of sound. The MEMS sensor itself consists of a flexible membrane and a rigidly fixed cover. Under the influence of air pressure, the membrane moves, changing the capacitance between the plates. This data is recalculated and output as an electrical signal to an integrated circuit. It is this signal that is converted into the sound that we hear.

Thanks to their design, MEMS microphones have the following advantages. Greater resistance to noise, vibration and temperature changes due to the absence of unnecessary connecting elements. Multiple MEMS microphones can be combined together to create a single array. Thanks to capacitive technology, these microphone arrays can capture sound from a precisely defined direction, effectively canceling echoes and background noise. Unlike other small microphones, such as electrets, MEMS microphones include more additional elements, such as preamplifiers, various filters and analog-to-digital converters. This means greater functionality while maintaining microscopic dimensions. Possibility of mounting such devices on the board using soldering.

Despite their many advantages, MEMS microphones are also not without their disadvantages. As we wrote above, MEMS microphones are often used as part of arrays, which increases the sound capture area, but at the same time reduces the service life of the devices. To work correctly, all microphones must work in unison, but the likelihood of one of them breaking is much higher than an individual device. Worse protection from moisture and dust than other microphones.

Microphone arrays include two or more built-in microphones, to which is added a programmable microprocessor designed to continuously determine the primary source of audio input and optimally adjust the output to achieve the best sound quality.

Let us highlight the most significant quality indicators of sound capture systems:

- useful signal/noise ratio, where the useful signal is the sound of the drone engine, and the noise is background noise, the microphone's own noise, and sounds from non-target sources;

- the shape of the radiation pattern and the ability of the system to change it depending on the environment;
- ability to localize the source of a useful signal and measurement accuracy parameters.

The most common way to build a signal capture unit is based on analog microphone arrays. A description of the problems that arise when developing analog microphone arrays, as well as the rationale for reducing the importance of the problem when using digital microphones in audio capture systems, is given in table 1.

Table 1
Problems encountered in the development of analog microphone arrays and justification for the feasibility of using digital microphones.

Problem when designing an analog array microphones	Rationale for using digital mirophones
Sharp increase in cost	Lack of a large number of auxiliary analog components
Reduced yield of suitable products due to the large number of components	Reducing the total number of microcircuits and topological complexity leads to an increase in the percentage of usable products due to the general laws of statistics
Increased development and debugging costs	Implementation of algorithms in code and digital interface blocks, which allows you to attract developers with less qualifications and experience
High sensitivity to electromagnetic radiation and power quality	The use of digital components is less sensitive to static failures and degradation of power supply quality
Increased production cycle	Less topological complexity guarantees the ability to produce a product according to almost any modern technological standards, making the launch process faster and cheaper
Increasing the testing cycle	The digital implementation allows you to write synthetic tests and generate input signals in the same way. Digital generators are more flexible and low cost, and the testing and debugging process is reduced to working with code

As an alternative to existing approaches that have the disadvantages outlined above, the authors of this work proposed to use an architecture built using digital MEMS microphones, which have become widespread recently. The meter for these microphones is located on-chip, so its digital output is minimally affected by the components that surround it. The most simple, inexpensive and perfect solution in terms of signal capture parameters was developed, which consists of using digital MEMS microphones and an Arduino microcomputer.

When choosing a digital microphone for use in a linear differential microphone array, it is important to consider the following factors:

- Sensitivity: The sensitivity of the microphone is a measure of how well it can convert sound waves into electrical signals. A higher sensitivity microphone will be able to pick up quieter sounds, but it may also be more susceptible to noise.
- Signal-to-Noise Ratio (SNR): The SNR of the microphone is a measure of the ratio of the desired signal (sound) to the undesired signal (noise). A higher SNR microphone will have less noise, resulting in cleaner recordings.
- Dynamic range: The dynamic range of the microphone is the range of sound pressure levels that it can accurately measure. A wider dynamic range microphone will be able to capture both very loud and very quiet sounds without distortion.
- Linearity: The linearity of the microphone is a measure of how accurately it can reproduce the input signal. A more linear microphone will produce recordings that are more faithful to the original sound. In addition to the above factors, it is also important to consider the cost and availability of the microphone when making a selection.

Each digital MEMS microphone can be simplified into the model shown in figure 9. Input sound vibrations are converted through a MEMS membrane into a weak electrical signal, which is then fed to the input of amplifier A. The pre-amplified signal then passes through an analog low-pass filter (LPF), which is necessary to protect against aliasing. The final element of signal processing in the microphone is a 4th order $\Sigma - \Delta$ modulator, which converts the input analog signal into a one-bit digital stream. The frequency of data bits from the output of the $\Sigma - \Delta$ modulator is equal to the frequency of the input timing signal CLK and, as a rule, lies in the range from 1 to 4 MHz.

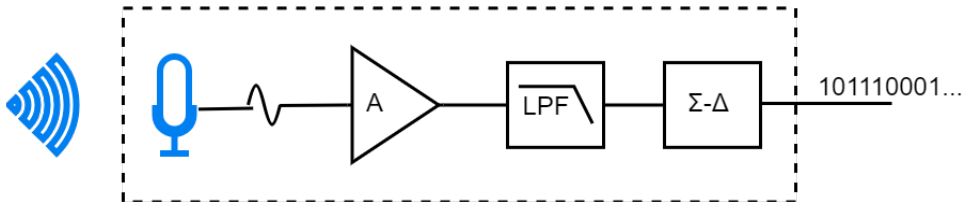

Figure 9: A simple model of a digital MEMS microphone.

In the time domain, the output of a $\Sigma - \Delta$ modulator is a jumbled collection of ones and zeros. However, if we assign a value of 1.0 to each high logical level of the microphone output, and a value of –1.0 to each low level and then perform a Fourier transform, we will obtain a spectrogram of the output data from the microphone.

Let's look at the pins of a digital microphone. VDD – microphone power supply, GND – Ground, CLK – input clock signal, synchronously with which the DATA line switches its DATA states. During one half of the CLK cycle this pin is in a high impedance state, and during the second half it serves as a pin for reading data from the $\Sigma - \Delta$ output of the microphone modulator. L/R_{Sel} – this pin is used to control switching of the DATA line. If L/R_{Sel} is connected to VDD, then after some time after detecting the rising edge of the CLK signal, the DATA pin goes into a high impedance state, and after the arrival of the falling edge of the CLK signal, the DATA pin is connected to the $\Sigma - \Delta$ output of the microphone modulator. If L/R_{Sel} is connected to GND, the edges of the CLK signal, along which the DATA line switches, are reversed (figure 10).

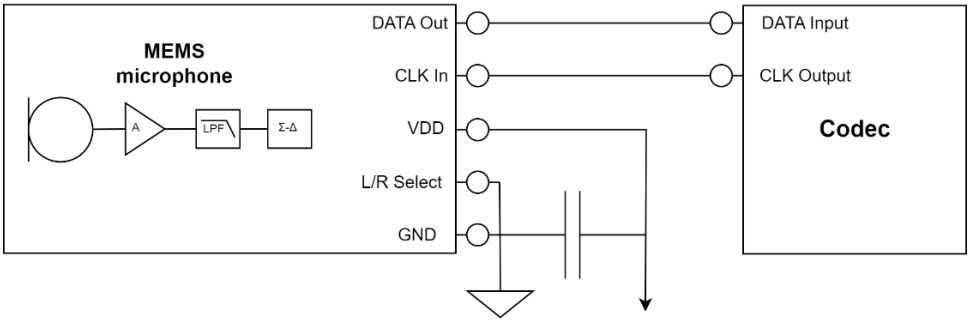

Figure 10: The pins of a digital microphone.

To isolate the audio frequency band signal, the data from the microphone must be filtered and resampled at a lower frequency (usually 50–128 times lower than the sampling frequency of the $\Sigma - \Delta$ modulator). A digital low-pass filter filters out external noise and the microphone's own noise outside the operating band ($f > F_{CLK}/2M$) to protect against aliasing, and also makes it possible to reduce the data repetition rate. In figure 11 presents one of the possible options for processing a one-bit data stream from a microphone, implemented in software on a DSP or in hardware in audio codecs. Shown in figure 11, the sampling frequency compression circuit (compressor) lowers the sampling frequency due to the fact that from every M samples of the filtered signal $w(mM)$, $M\breve{}1$ sample is discarded. The

input and output of the converter shown in figure 8 are related by the following expression:

$$y(m) = w(mM) = \sum_{k=-\infty}^{\infty} h(k)x(mM-k). \tag{15}$$

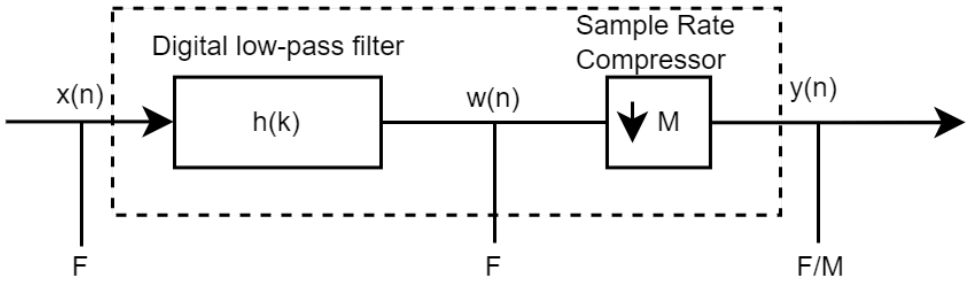

Figure 11: Signal conversion by $\Sigma - \Delta$ modulator.

MEMS microphones have PDM (Pulse-Density Modulation, PDM) outputs. Pulse density modulation is a method of transmitting the relative change in signal per sample, which can be mathematically described by the formula

$$x[n] = -A(-1^{a[n]}), \tag{16}$$

where $x[n]$ contains in each term the relative change in the signal in the form of 1 bit with a sign, which is specified by the transition. A negative increment is a transition from 1 to 0, a positive increment is from 0 to 1. Repeating ones increases the overall amplitude of the signal, and repeating zeros decreases (figure 12).

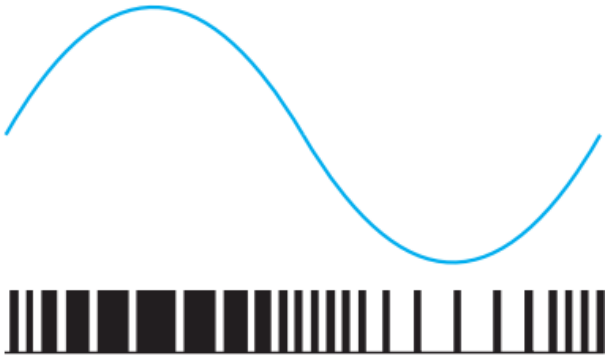

Figure 12: Period of a sine wave per 100 samples.

A mathematical model for pulse density modulation can be obtained using a delta-sigma modulator model. In the discrete frequency domain, the operation of a delta-sigma modulator can be described by the formula

$$O(z) = I(z) + E(z)(1 - z^{-1}), \tag{17}$$

where $O(z)$, $I(z)$ are the signal spectra at the input and output of the modulator; $E(z)$ is the sampling error of the delta-sigma modulator; $1 - z^{-1}$ is high-pass filter. As a result of transforming the formula, we get

$$O(z) = E(z)[I(z) - O(z)z^{-1})]\frac{1}{1 - z^{-1}}. \tag{18}$$

According to this formula, the error $E(z)$ reduces the value of the signal at the output $O(z)$ in the low-frequency region and increases it in the high-frequency region, as a result of which the quantization noise spectrum shifts predominantly to the high-frequency region.

Table 2
Array characteristics (ULA with 2 microphones).

Array characteristic	Value
Array directivity	2.97 dBi at 0 Az; 0 El
Array span	x=0 m y=17 mm z=0 m
Number of elements	2
HPBW	60.50° Az / 360.00° El
FNBW	180.00° Az / -° El
SLL	- dB Az / - dB El
Element polarization	None

Let $i[n]$ be a sample of the signal at the input of the modulator in the time domain, and $o[n]$ be a sample of the output signal, then, using the inverse z-transform, we can proceed to the expression

$$o[n] = i[n] + e[n] - e[n-1], \tag{19}$$

where

$$o[n] = \begin{cases} 1 & \text{if} x[n] \geq e[n-1]; \\ -1 & \text{if} x[n] < e[n-1]; \end{cases} \tag{20}$$

$$e[n] = o[n] - i[n] + e[n-1]. \tag{21}$$

The signal from the output signal sample $o[n]$ is represented as 1 bit and takes values ± 1, and is implemented so that the value of the current quantization error $e[n]$ is minimal. In this case, the quantization error $e[n]$ of each sample appears at the device input during the subsequent sample.

When implementing frequency converters in software, a finite impulse response (FIR) filter or an Infinite impulse response (IIR) filter can be used as a digital LPF. Developers should be very careful when choosing the type of filter, its length and bit depth, since the performance of the entire system as a whole directly depends on this. A correctly calculated and implemented decimator (frequency converter) in some cases will significantly reduce the cost of products and increase its technical characteristics.

As a second option, audio codecs adapted for this can be used to convert data from the output of a digital microphone, which will significantly reduce product development time. For example, Analog Devices offers the ADAU1361 and ADAU1761 codecs, which are suitable for the ADMP521 microphones. In our work we used a microphone ADMP521. However, the process of creating digital audio devices becomes simple in terms of hardware implementation and complex in terms of writing programs for the microcontrollers used.

Next, we conducted a simulation and computational experiment of a uniform linear array of 2 omnidirectional microphones using Matlab Sensor Array Analyzer.

The following model parameters were used. The distance between the microphones is 20 mm. The board has two MEMS microphones spaced 20 mm apart. This spacing is ideal for detecting acoustic events. Additionally, the 20 mm spacing is equivalent to $8 \cdot 2.54$ mm, which makes it suitable for DIP (Dual In-line Package) – a type of housing for microcircuits, electronic modules and some other electronic components. Experimentally, a distance of 0.017 m was determined for the formation of a bi-directional pattern.

Next, we conducted a simulation and computational experiment of a uniform linear array of 2 omnidirectional microphones using Matlab Sensor Array Analyzer. The following model parameters were used. The distance between the microphones is 17 mm. The speed of sound is 343 m/s, the signal frequency is 10 kHz. As a result, we obtained the parameters listed in table 2.

The Matlab script is listed below:

```
% Create a Uniform Linear Array Object
Array = phased.ULA('NumElements',2, 'ArrayAxis','y');
Array.ElementSpacing = 0.017;
```

```
Array.Taper = ones(1,2).';
% Create an omnidirectional microphone element
Elem = phased.OmnidirectionalMicrophoneElement;
Elem.FrequencyRange = [0 10000];
Array.Element = Elem;
% Assign Frequencies and Propagation Speed
Frequency = 10000;
PropagationSpeed = 343;
% Create Figure
% Plot Array Geometry figure;
viewArray(Array,'ShowNormal',true,
'ShowTaper',true,'ShowIndex','All',
'ShowLocalCoordinates',true,'ShowAnnotation',true,
'Orientation',[0;0;0]);
% Find the weights
w = ones(getNumElements(Array), length(Frequency));
% Plot 2d azimuth graph
format = 'polar';
cutAngle = 0;
plotType = 'Directivity';
plotStyle = 'Overlay';
figure;
pattern(Array, Frequency, -180:180, cutAngle, 'PropagationSpeed',
PropagationSpeed, 'CoordinateSystem', format ,'weights', w,
'Type', plotType, 'PlotStyle', plotStyle);
% Find the weights
w = ones(getNumElements(Array), length(Frequency));
% Plot 2d elevation graph
format = 'polar';
cutAngle = 0;
plotType = 'Directivity';
plotStyle = 'Overlay';
figure;
pattern(Array, Frequency, cutAngle, -90:90, 'PropagationSpeed',
PropagationSpeed, 'CoordinateSystem', format ,'weights', w,
'Type', plotType, 'PlotStyle', plotStyle);
% Find the weights
w = ones(getNumElements(Array), length(Frequency));
% Plot U Pattern
format = 'uv';
plotType = 'Directivity';
plotStyle = 'Overlay';
figure;
pattern(Array, Frequency, -1:0.01:1, 0, 'PropagationSpeed',
PropagationSpeed, 'CoordinateSystem', format,'weights', w,
'Type', plotType, 'PlotStyle', plotStyle);
```

The resulting pattern has the shape of a bi-directional (figure 13). As you can see, the design of the grille allows you to create a grille with the main lobes directed at -90 and 90 degrees. To form a cardiode radiation pattern, as mentioned above, it is necessary to use delay-and-sum and filter-and-sum algorithms. The meaning of these algorithms is that microphone signals are added with different delays (different phase shifts), aligning the phases of signals coming from the selected direction (source

localization) for each frequency. In this case, the beamforming algorithm makes it possible to amplify the signals generated by sound coming from the selected direction, i.e. performs a kind of focusing of sounds.

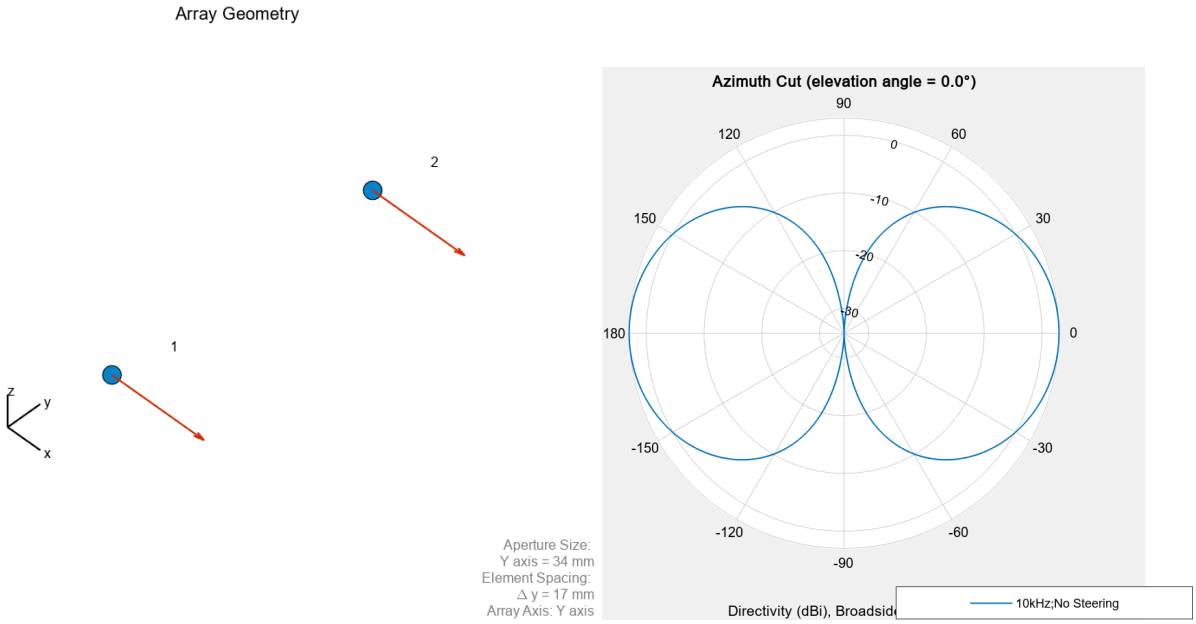

Figure 13: The geometry of the array and the directional diagram (linear array of 2 microphones).

ADMP521 microphones were connected to the ADAU1761 codec in accordance with the technical specifications of both products (figure 14).

A model was also created based on an array of four omnidirectional microphones located at a distance of 20 mm (figure 15). The following model parameters were used. The distance between the microphones is 17 mm. The speed of sound is 343 m/s, the signal frequency is 10 kHz. As a result, we obtained the parameters listed in table 3.

Table 3
Array characteristics (ULA with 4 microphones).

Array characteristic	Value
Array directivity	5.98 dBi at 0 Az; 0 El
Array span	x=0 m y=51 mm z=0 m
Number of elements	4
HPBW	26.52° Az / 360.00° El
FNBW	60.58° Az / -° El
SLL	11.30 dB Az / - dB El
Element Polarization	None

The Matlab script has the following form:

```
% Create a Uniform Linear Array Object
Array = phased.ULA('NumElements',4, 'ArrayAxis','y');
Array.ElementSpacing = 0.017;
Array.Taper = ones(1,4).';
% Create an omnidirectional microphone element
Elem = phased.OmnidirectionalMicrophoneElement;
Elem.FrequencyRange = [0 10000];
Array.Element = Elem;
```

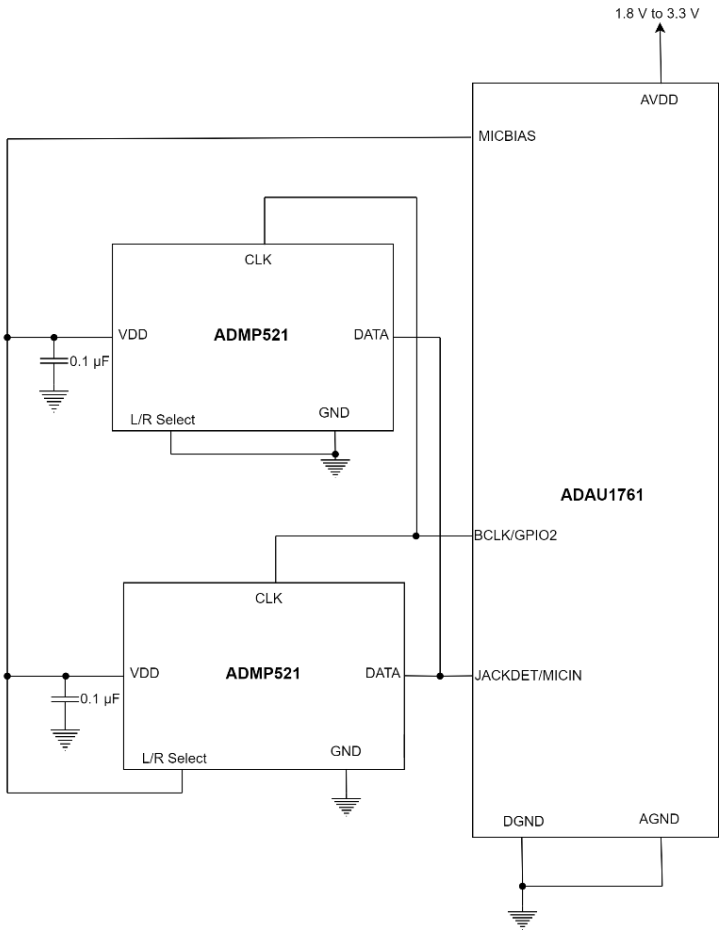

Figure 14: Connection diagram of ADMP521 microphones to ADAU1761 codec.

```
% Assign Frequencies and Propagation Speed
Frequency = 10000;
PropagationSpeed = 343;
% Create Figure
% Plot Array Geometry
figure;
viewArray(Array,'ShowNormal',true,'ShowTaper',true,'ShowIndex','All',
'ShowLocalCoordinates',true,'ShowAnnotation',true,'Orientation',[0;0;0]);
% Calculate Steering Weights
Freq3D = 10000;
% Find the weights
w = ones(getNumElements(Array), length(Frequency));
% Plot 3d graph
format = 'polar';
plotType = 'Directivity';
figure;
pattern(Array, Freq3D , 'PropagationSpeed', PropagationSpeed,
'CoordinateSystem', format,'weights', w(:,1),
'ShowArray',false,'ShowLocalCoordinates',true,
'ShowColorbar',true,'Orientation',[0;0;0],'Type', plotType);
% Find the weights
w = ones(getNumElements(Array), length(Frequency));
```

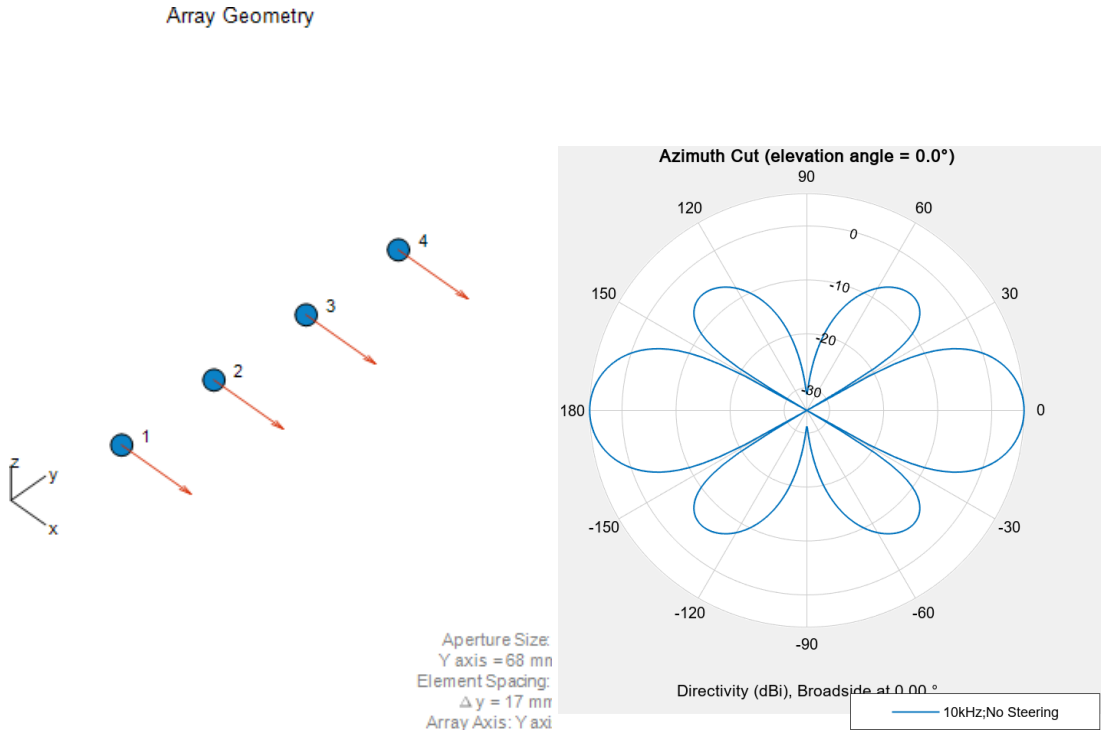

Figure 15: The geometry of the array and the directional diagram (linear array of 4 microphones).

```
% Plot 2d azimuth graph
format = 'polar';
cutAngle = 0;
plotType = 'Directivity';
plotStyle = 'Overlay';
figure;
pattern(Array, Frequency, -180:180, cutAngle, 'PropagationSpeed',
PropagationSpeed,'CoordinateSystem', format ,'weights', w,
'Type', plotType, 'PlotStyle', plotStyle);
% Find the weights
w = ones(getNumElements(Array), length(Frequency));
% Plot 2d elevation graph
format = 'polar';
cutAngle = 0;
plotType = 'Directivity';
plotStyle = 'Overlay';
figure;
pattern(Array, Frequency, cutAngle, -90:90, 'PropagationSpeed',
PropagationSpeed,'CoordinateSystem', format ,'weights', w,
'Type', plotType, 'PlotStyle', plotStyle);
```

So, during the computational experiment, we built 2 linear microphone arrays with bi-directionality. The directionality of these arrays can be easily converted to unidirectional (cardioid) using known algorithms or hardware (codecs). The tuning of the circuit to create cardioid directivity will be considered in further studies.

5. Discussion

The following questions require additional discussion and clarification: the dependence of the azimuthal pattern of the proposed linear microphone arrays on the source frequency; the choice of analog or digital MEMS microphones and labor-intensiveness in the development of a microphone array; the use of directional microphones instead of omnidirectional; peculiarities of localization of a moving sound source (Doppler effect, reflection from obstacles, etc.); higher-order differential beam array formers; signal processing algorithms of microphone arrays.

6. Conclusions

The study looks at setting up a microphone array to determine the position of a UAV (unmanned aerial vehicle) based solely on the sound of its engines. The location of the microphones plays a crucial role for accurate localization. A mathematical model of pulse density modulation of a digital MEMS microphone is also considered. This work shows the dependence of the efficiency of a differential array of first-order microphones on frequency. Based on the presented dependence of the directivity on frequency and the instability model of the microphone parameters, a rational operating frequency range for the normal functioning of the microphone array can be determined.

A model of a linear microphone array based on MEMS omnidirectional microphones is proposed, which with a certain geometrical arrangement give a bi-directional pattern, which, in principle, can be easily transformed into a unidirectional one with the use of special algorithms or hardware (for example, ADAU1761 codecs). Refinement of the circuit to achieve cardioid directivity will be addressed in forthcoming research.

7. Author contributions

Conceptualization, methodology – Andrii V. Riabko, Oksana V. Zaika; setting tasks, conceptual analysis – Tetiana A. Vakaliuk, Oksana V. Zaika; development of the model – Andrii V. Riabko, Valerii V. Kontsedailo; software development, verification – Andrii V. Riabko, Roman P. Kukharchuk; analysis of results, visualization – Roman P. Kukharchuk, Tetiana A. Vakaliuk; drafting of the manuscript – Valerii V. Kontsedailo, reviewing and editing – Tetiana A. Vakaliuk.

All authors have read and approved the published version of this manuscript.

References

[1] M. Cobos, F. Antonacci, A. Alexandridis, A. Mouchtaris, B. Lee, A Survey of Sound Source Localization Methods in Wireless Acoustic Sensor Networks, Wireless Communications and Mobile Computing 2017 (2017) 3956282. doi:10.1155/2017/3956282.

[2] A. R. Petrosian, R. V. Petrosyan, I. A. Pilkevych, M. S. Graf, Efficient model of PID controller of unmanned aerial vehicle, Journal of Edge Computing 2 (2023) 104–124. doi:10.55056/jec.593.

[3] M. Risoud, J.-N. Hanson, F. Gauvrit, C. Renard, P.-E. Lemesre, N.-X. Bonne, C. Vincent, Sound source localization, European Annals of Otorhinolaryngology, Head and Neck Diseases 135 (2018) 259–264. doi:10.1016/j.anorl.2018.04.009.

[4] W. A. Yost, M. T. Pastore, Y. Zhou, Sound Source Localization Is a Multisystem Process, Springer International Publishing, Cham, 2021, pp. 47–79. doi:10.1007/978-3-030-57100-9_3.

[5] E. King, A. Tatoglu, D. Iglesias, A. Matriss, Audio-visual based non-line-of-sight sound source localization: A feasibility study, Applied Acoustics 171 (2021) 107674. doi:10.1016/j.apacoust.2020.107674.

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

[7] T. Yamada, K. Itoyama, K. Nishida, K. Nakadai, Sound Source Tracking by Drones with Microphone Arrays, in: 2020 IEEE/SICE International Symposium on System Integration (SII), 2020, pp. 796–801. doi:10.1109/SII46433.2020.9026185.

[8] A. D. Firoozabadi, P. Irarrazaval, P. Adasme, D. Zabala-Blanco, P. Palacios-Játiva, H. Durney, M. Sanhueza, C. Azurdia-Meza, Three-dimensional sound source localization by distributed microphone arrays, in: 2021 29th European Signal Processing Conference (EUSIPCO), 2021, pp. 196–200. doi:10.23919/EUSIPCO54536.2021.9616326.

[9] Y. Sasaki, R. Tanabe, H. Takemura, Probabilistic 3D sound source mapping using moving microphone array, in: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2016, pp. 1293–1298. doi:10.1109/IROS.2016.7759214.

[10] M. C. Catalbas, M. Yildirim, A. Gulten, H. Kurum, S. Dobrišek, Estimation of Trajectory and Location for Mobile Sound Source, International Journal of Advanced Computer Science and Applications 7 (2016). doi:10.14569/IJACSA.2016.070934.

[11] K. Hoshiba, K. Washizaki, M. Wakabayashi, T. Ishiki, M. Kumon, Y. Bando, D. Gabriel, K. Nakadai, H. G. Okuno, Design of UAV-Embedded Microphone Array System for Sound Source Localization in Outdoor Environments, Sensors 17 (2017) 2535. doi:10.3390/s17112535.

[12] T. Tachikawa, K. Yatabe, Y. Oikawa, 3D sound source localization based on coherence-adjusted monopole dictionary and modified convex clustering, Applied Acoustics 139 (2018) 267–281. doi:10.1016/j.apacoust.2018.04.033.

[13] M. Wakabayashi, H. G. Okuno, M. Kumon, Drone audition listening from the sky estimates multiple sound source positions by integrating sound source localization and data association, Advanced Robotics 34 (2020) 744–755. doi:10.1080/01691864.2020.1757506.

[14] L. Carneiro, A. Berry, Three-dimensional sound source diagnostic using a spherical microphone array from multiple capture positions, Mechanical Systems and Signal Processing 199 (2023) 110455. doi:10.1016/j.ymssp.2023.110455.

[15] S. Kita, Y. Kajikawa, Fundamental study on sound source localization inside a structure using a deep neural network and computer-aided engineering, Journal of Sound and Vibration 513 (2021) 116400. doi:10.1016/j.jsv.2021.116400.

[16] F. R. do Amaral, J. Rico, M. A. F. de Medeiros, Design of microphone phased arrays for acoustic beamforming, Journal of the Brazilian Society of Mechanical Sciences and Engineering 40 (2018) 354. doi:10.1007/s40430-018-1275-5.

[17] F. Grondin, F. Michaud, Lightweight and optimized sound source localization and tracking methods for open and closed microphone array configurations, Robotics and Autonomous Systems 113 (2019) 63–80. doi:10.1016/j.robot.2019.01.002.

[18] C. Rascon, I. Meza, Localization of sound sources in robotics: A review, Robotics and Autonomous Systems 96 (2017) 184–210. doi:10.1016/j.robot.2017.07.011.

[19] M. J. Bianco, P. Gerstoft, J. Traer, E. Ozanich, M. A. Roch, S. Gannot, C.-A. Deledalle, Machine learning in acoustics: Theory and applications, The Journal of the Acoustical Society of America 146 (2019) 3590–3628. doi:10.1121/1.5133944.

[20] H. Niu, Z. Gong, E. Ozanich, P. Gerstoft, H. Wang, Z. Li, Deep-learning source localization using multi-frequency magnitude-only data, The Journal of the Acoustical Society of America 146 (2019) 211–222. doi:10.1121/1.5116016.

[21] W. He, P. Motlicek, J.-M. Odobez, Deep Neural Networks for Multiple Speaker Detection and Localization, in: 2018 IEEE International Conference on Robotics and Automation (ICRA), 2018, pp. 74–79. doi:10.1109/ICRA.2018.8461267.

[22] A. Ebrahimkhanlou, S. Salamone, Single-Sensor Acoustic Emission Source Localization in Plate-Like Structures Using Deep Learning, Aerospace 5 (2018) 50. doi:10.3390/aerospace5020050.

[23] S. Adavanne, A. Politis, J. Nikunen, T. Virtanen, Sound Event Localization and Detection of Overlapping Sources Using Convolutional Recurrent Neural Networks, IEEE Journal of Selected Topics in Signal Processing 13 (2019) 34–48. doi:10.1109/JSTSP.2018.2885636.

[24] G. Chardon, Theoretical analysis of beamforming steering vector formulations for acoustic source localization, Journal of Sound and Vibration 517 (2022) 116544. doi:10.1016/j.jsv.2021.

116544.

[25] B. da Silva, A. Braeken, K. Steenhaut, A. Touhafi, Design Considerations When Accelerating an FPGA-Based Digital Microphone Array for Sound-Source Localization, J. Sensors 2017 (2017) 6782176:1–6782176:20. doi:10.1155/2017/6782176.