Bearing Estimation of Specific Audio Events using Microphone Arrays

Macarena Varela, Wulf-Dieter Wirth und Miriam Häge

Fraunhofer FKIE- SDF, Fraunhoferstr. 20, 53343 Wachtberg, Germany,
E-Mail: macarena.varela@fkie.fraunhofer.de

Abstract

Automatic detection, bearing estimation and source localization of specific audio events, such as drone detection and impulsive sound localization, play an important role in security applications. Using nearly omni-directional microphones arranged in groups and exploiting the advantages of array processing techniques, such as beamforming, is an approach to solve those tasks. The spatial distribution and number of microphones in arrays has a crucial influence on the bearing estimation of audio events, therefore in this paper three microphone array configurations are described, and their performance in different open field experiments is presented.

Keywords: microphone array, sound detection, bearing estimation, sound source localization

Introduction

In surveillance or security applications, the detection of specific events, such as sounds of vehicles or explosions, is essential. To fulfill these purposes, various microphone array systems were created, having as aim the detection of multiple sources of audio events depending on the type of scenario, and to improve their bearing estimation (see Figure 1).

All systems are composed of a microphone array, a sound card and a computer as indicated in Figure 2. The linear arrangement was the first array to be constructed and consists of 8 equidistant omni-directional microphones placed linearly. Due to its linearity, this arrangement only provides an azimuth estimation of the sound, but no elevation information. The goal of the second and third arrangements was to obtain a two-dimensional estimation of the bearing (azimuth and elevation). The double-layer array was the second array to be built and is composed of 12 microphones equidistantly distributed in 2 circles over 2 layers, where one of the circles is twice the size of the second one. The last arrangement to be created was the "crow's nest" array. In this arrangement, 10 microphones are randomly distributed over a sphere. The output signals of the microphones are digitized with a sampling rate up to 192 KHz, and the data is processed online and in real-time by using MATLAB. Subsequently, the data is processed with the specific procedure shown in Figure 1.
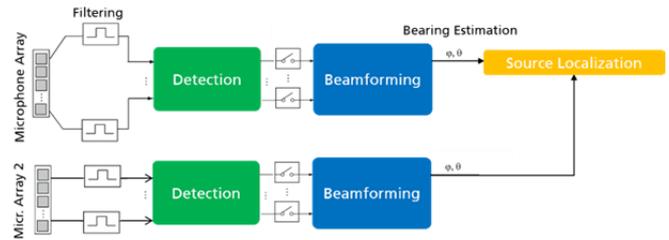

Figure 1: Flow chart of the Experimental System

First, a bandpass filter is applied to all the signals to remove background noise, such as wind noise or noise emanating from a moving platform. Then, beamforming is performed only on detected audio events, where the direction of arrival (DoA) corresponds to the position of the main beam direction of the beam pattern. Once the bearing is estimated, source localization can be achieved as explained in the two following cases: multiple arrays are simultaneously used and detect the same event, and in the particular case of a source producing 2 sounds. For example, gunshot sound events produce two types of sound waves: the shock wave and the muzzle blast. In that specific type of event, additionally to the shooter localization, the bullet trajectory could also be estimated.

Experimental System

The automatic online and real-time experimental system is continuously recording in outdoor environments. This data is then processed as shown in Figure 1. The signals are windowed into 1 second segments and filtered with a bandpass filter. For every detected event, the system is activated, and produces an alarm containing relevant information, such as the localization of the sound source and the time it occurred. Figure 2 shows one of the technical realizations of a microphone array system.
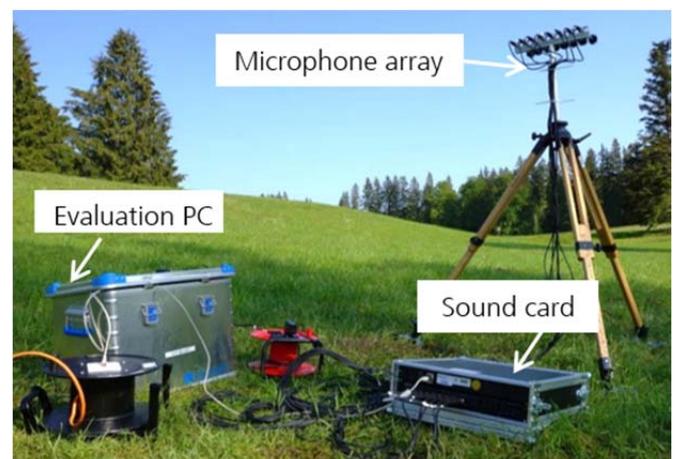

Figure 2: Sensor system using a linear microphone array

The detection method used varies depending on the type of sound to be detected, and holds a crucial role in the performance of the system, since only detections of true and false positive events are desired. Consequently, the detection block is explained for one type of event: impulsive sound events. Then, the beamforming method is shortly explained.

Detection of Impulsive Sound Events

For Impulsive Sound Events, a two-step detection is applied. All windowed array signals are first summed-up incoherently and compared with a selected threshold to detect peak signals. Then, the final detection consists in comparing the signals before and after the peak [1]. There are different surrounding sounds before and after such peaks, and therefore an appropriate comparison of those surrounding sounds, can lead to a successful decision toward an event detection. This innovative detection method demonstrated robustness and accuracy in the detection of impulsive sounds, such as gunshots, present in noisy open field experiments, and is based on the preliminary work of Damarla [1].

Beamforming and Direction of Arrival

The direction of arrival (DoA) estimation, also called bearing angle, was determined by coherent broadband beamforming [2]. In this method, the acoustic wideband array signals are windowed, Fourier transformed and concatenated to form extended space-frequency vectors. Those vectors are then weighted considering the linear array geometry and added to produce a beam pattern. The bearing angle of the gunshot corresponds to the position of the main beam direction of the beam pattern.

Microphone Arrays

The following three microphone array arrangements shown in Figure 3 were used in different open field experiments and their performance is explained successively.
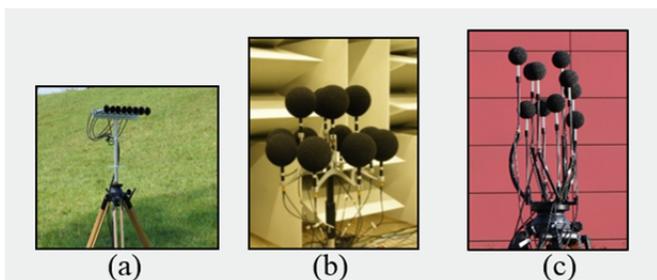

Figure 3: Types of microphone arrays: Uniform Linear (a), Double-Layer (b), and Crow's Nest (c)

Uniform Linear Array (ULA)

The ULA [2] is shown in Figure 3a; it was the first array of microphones to be realized and its initial purpose was findings the DoA of impulsive sounds present at the ground level, providing no elevation information about the event. Due to the arrangement of the microphones (see Figure 4a), the azimuth pattern presents a strong back lobe that may lead to ambiguity problems with respect to the DoA, especially if the scenario is not previously known. On the other hand, the beamwidth is narrow and the main lobe extremely accurate.
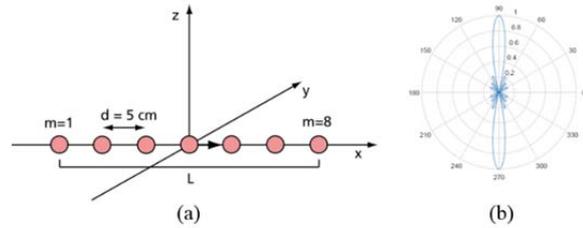

Figure 4: Uniform Linear Array: array geometry (a), and azimuth pattern (b)

Behring microphones are used in this array, having a frequency response up to 20 kHz. The processing time of a detected event is less than 1 second. This array can be used ground-based as shown in Figure 2 or mounted on a moving platform such as the one presented on Figure 5.

Figure 5: ULA mounted on a moving platform

Double-Layer Array (DLA)

The DLA [2] consists of two uniform circular arrays composed of 8 microphones in the lower circular array and 4 in the upper circular array (see Figure 3b). The inter-microphone distance is of 0.08m. It was the second array to be realized with the purpose of extending the azimuth to 360° and removing the back lobe encountered in the ULA arrangement. Moreover, the two layers of this array enable an elevation estimation. However, due to the limited microphone layers, the elevation pattern is not as accurate as the azimuth pattern.
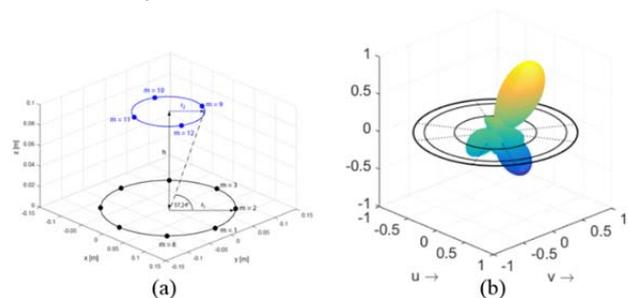

Figure 6: Double-Layer Array: array geometry (a), and directivity pattern (b)

Brüel & Kjaer microphones are used in this array, covering frequencies up to 20 kHz. The processing time of a detected event is of about 1.5 seconds. This arrangement provides a two-dimensional estimation of the bearing (azimuth and elevation). This array can be used ground-based or mounted on a moving platform.

Crow's Nest Array

Figure 3c shows the CNA [2] realized at Fraunhofer-FKIE which is the only configuration among the others having a full hemispherical coverage. The microphones are randomly distributed as shown in Figure 7a, and their inter-microphone distance is of at least 0.08m. This is the last arrangement realized and its aim is to have high accuracy in azimuth and in elevation (Figure 7b).

The microphones used in this array have a frequency range from 5 Hz to 96 kHz. The processing time of a detected event is of about 1.5 seconds. This array is will be used ground-based and mounted on a moving platform.
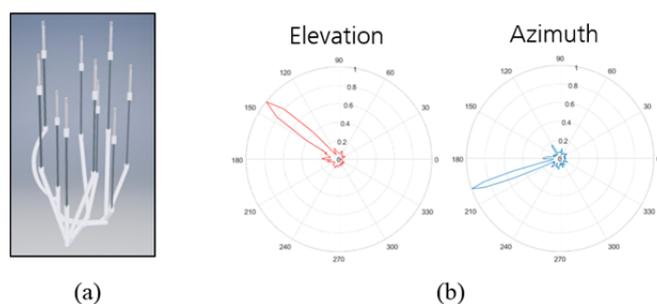

Figure 7: Crow's Nest Array: array geometry (a), and directivity pattern (b)

Experimental Results

Within the framework of the project "Process Chain to Support Automated Reconnaissance" [2], gunshots were detected and their direction of arrival estimated using one ULA as shown in Figure 2. In the field exercise, gunshots audio recording were carried out using a firearm of type G3 with blank ammunitions, namely, only producing a muzzle blast and no shock wave. The locations of the shooters were not fixed nor previously known, but their area of movement was restricted as shown in Figure 8. The ground-based microphone array was consequently placed toward that area in order to avoid any ambiguities created by the strong back lobe of that array (Figure 4b). The array position and orientation was measured with high precision using global positioning system (GPS).
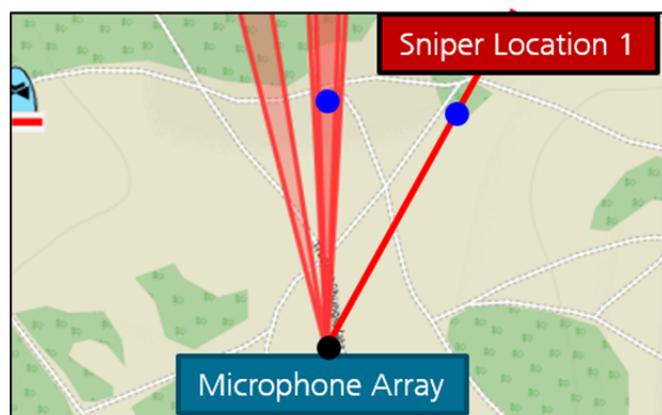

Figure 8: Microphone array system (black dot), bearing angles (red sectors and lines) and the sniper locations ground truths (blue dots); Source: openstreetmap-contributors

A UGV, such as the one shown in Figure 9 , was moving near the microphone array, thus creating disturbing noises to be removed by the experimental system filter (see Figure 1).

The microphones used for data collection had a frequency response which was enough to capture emitted muzzle blast waves. Shock waves have a much higher frequency content, therefore if shock waves would have been present in the data, microphones having a higher frequency response would have been necessary, such as the ones used for the CNA. During the field exercise (see Figure 8), gunshots were effectively detected in real time and their corresponding bearing angles were estimated to an accuracy of 1°.

Figure 8 shows the estimated positions for three shooter positions during the conducted field exercise to prove the accuracy of the system. The firearms were fired from three distinct locations, where two of the positions could be measured using GPS, and then be used as ground truth information.

Figure 9: Unmanned Ground Vehicle

Conclusion and Future Work

The method used by the experimental system (see Figure 1) presented excellent DoA results. In fact, it provides an automatic detection, bearing estimation and in some particular cases source localization of specific audio events, such as impulsive sounds. The spatial distribution has a strong influence on the bearing estimation of audio events, and the CNA proved to perform best, since it has a narrow beamwidth with an extremely accurate main lobe, is the only arrangement that provides a full hemispherical coverage.

Within the framework of the successor project called "AutoAuge", similar field exercises will be conducted using multiple CNAs. In the exercises involving blank ammunitions, the shooters will be localized by the fusion of the bearings delivered by the multiple arrays, whereas for live ammunitions, the shooters will be localized by one single array. In fact, if live ammunition is used, microphones with higher frequency response, such as the ones used in the CNA, will accurately not only capture the signals from the muzzle blast, but also the rapid rising and falling times of the projectile shock wave produced by the bullet. The characteristics of the shock wave abrupt "N" shape provide information about the caliber, the Mach number, and consequently about the initial speed of the bullet. The sniper

localization can then be estimated by using the time difference of arrival (TDOA) of both waves and the bullet's initial velocity as explained in [1]. Therefore, for live ammunitions a single sensor will provide detection, bearing information, classification and localization results.

References

[1] Damarla, T.: Battlefield Acoustics, Springer (2015)

[2] Wirth, W.-D.: Radar Techniques Using array Antennas, 2nd Edition, IET-Radar, Sonar and Navigation Series, 2013

[3] Brüggemann, B.: Manned-Unmanned Teaming to Support a Dismounted Infantry Platoon, Specialists meeting on Intelligence & Autonomy (NATO IST-127), 2016