Comparison of acoustic localisation techniques for drone position estimation using real-world experimental data

Altena, Anique; Luesutthiviboon, Salil; de Croon, Guido; Snellen, Mirjam; Voskuijl, Mark

Publication date
2023

Document Version
Final published version

Published in
Proceedings of the 29th International Congress on Sound and Vibration, ICSV 2023

Citation (APA)
Altena, A., Luesutthiviboon, S., de Croon, G., Snellen, M., & Voskuijl, M. (2023). Comparison of acoustic localisation techniques for drone position estimation using real-world experimental data. In E. Carletti (Ed.), *Proceedings of the 29th International Congress on Sound and Vibration, ICSV 2023* (Proceedings of the International Congress on Sound and Vibration). Society of Acoustics.

Annual Congress of the International Institute of Acoustics and Vibration (IIAV)

COMPARISON OF ACOUSTIC LOCALISATION TECHNIQUES FOR DRONE POSITION ESTIMATION USING REAL-WORLD EXPERIMENTAL DATA

Anique Altena, Salil Luesutthiviboon, Guido de Croon, Mirjam Snellen
Aircraft Noise and Climate Effects, Delft University of Technology, Delft, The Netherlands
e-mail: A.Altena@tudelft.nl
Mark Voskuijl
Faculty of Military Sciences, Netherlands Defence Academy, Den Helder, The Netherlands

Due to technological advances in the drone industry, security threats induced by unmanned aerial vehicles (UAVs) are becoming more relevant. Fast and accurate localisation systems need to be designed. One approach is localisation of UAVs by their sound using acoustic techniques. So far, a systematic performance assessment of acoustic techniques for drone localisation, based on real-world data, is lacking. This work presents a comparison of selected techniques using real-world measurement data. The achieved performance serves as a baseline for future design of novel localisation methods. Three techniques are chosen. The first technique estimates the time-difference-of-arrival (TDOA) using generalised cross-correlation with phase transform weighting (GCC-PHAT). The second technique is differential evolution, which approaches the localisation task as a global optimisation problem. The third technique is conventional frequency domain beamforming. Real-world data of 5 quadrotor UAVs were used acquired with an acoustic microphone-array. The performance of the techniques is assessed using the absolute error between the estimated source location and the true source location obtained from the onboard GPS tracker of the drones. GCC-PHAT and differential evolution attempt to estimate the drone position in one or few steps. They have a much shorter runtime than beamforming, which is an exhaustive grid search algorithm. However, these techniques result in lower detection ranges and accuracy compared to beamforming.
Keywords: Sound source localisation; Drones; Beamforming; Time-difference-of-arrival; Differential evolution

1. Introduction

Drones, or unmanned aerial vehicles (UAVs), can intentionally or unintentionally pose a threat for critical infrastructure, VIP security or during weaponised conflicts. This urges industry and researchers to investigate fast and accurate UAV detection and mitigation systems. Traditionally, this is done by e.g. radar or vision. Both are hampered by resolution issues. Drones can be small or can appear small at large distances and they can come from any direction [1]. An emerging topic is therefore drone detection using acoustics, as drone sound is very distinct from other sound sources like birds and it can be perceived omni-directional. The use of acoustic imaging techniques for the purpose of drone localisation has caught the attention of researchers over the past years. The suitability of these techniques has been validated many times [2]. UAV localisation techniques that can be used for sound source localisation can roughly be divided into three categories: time-difference-of-arrival (TDOA) techniques, beamforming techniques and bio-inspired techniques.

TDOA methods use correlation to identify the time difference between the arrival of one signal at two different microphones within a microphone array. After estimation of this TDOA, an estimation technique, such as least-squares, is used for estimating the location of the sound source. The second approach, beamforming [3], is an exhaustive grid search approach. This technique applies a delay to signals received at different microphones according to assumed sound source locations within a search area. At the correct location, constructive interference results into a maximum output. However, the microphone array configuration can result into misleading sidelobes. Lastly, bio-inspired techniques contain a wide variety of methods, such as neural networks [4], support vector machines [5] or differential evolution [6]. They can be used as a classification algorithm to detect or classify a drone, or as a regression algorithm to directly estimate the location of the sound source.

Variations of the previously mentioned techniques have been tested for the purpose of UAV localisation and reported in literature. Regarding the TDOA methods, the most commonly used technique is generalised cross-correlation with phase transform weighting (GCC-PHAT) [5, 7]. Through the use of historical estimations, localisation ranges of up to 100 m are achieved in these studies. The most commonly used acoustic imaging technique is beamforming and many researchers have tested this method [8–13]. In most of these studies, tests are performed on short distances of up to 10 m. The authors of [12] and [13] have shown promising results by applying a tracking algorithm which leads to localisation ranges up to 200 m. When looking into different bio-inspired techniques, one that stands out are neural networks. These have been trained and tested on raw or processed acoustic drone data several times and some outperform classical acoustic imaging techniques [4, 14–16]. One study that stands out is [16], where a deep neural network has been designed such that it works similar as beamforming. Detection ranges of up to a few 100 m are achieved, although this is done with multiple microphone arrays spread over an area. One downside of neural networks is that they are highly sensitive to the data they are trained on and not all networks can generalise for sound signatures of drones not seen before. One possible bio-inspired technique that does not have to be trained on data is called differential evolution. This is a global optimisation approach which has been used successfully in the past for grid-less localisation of sound sources [6, 17].

The aim of this research is to compare three techniques for acoustic localisation of UAVs. The goal is to set a benchmark performance for localisation techniques to allow for the design of novel localisation methods which can be compared to this benchmark. The techniques that are chosen are GCC-PHAT, conventional beamforming and differential evolution, where the latter has not been used for the purpose of drone localisation yet. Data to test these methods was obtained during a real-world experiment where 5 drones were flown.

The remainder of this contribution consists of the following. Details on the three considered methods are presented in section 2. Details on the experimental set-up are provided in section 3. The results are presented and discussed in section 4 and this paper is concluded in section 5.

2. Sound Source Localisation Methods

The three techniques that are selected for comparison are GCC-PHAT, beamforming and differential evolution. They will be explained briefly. It should be noted that the location of the UAVs will be estimated in Cartesian coordinates, i.e. $[x, y, z]^T$.

2.1 Generalised Cross-Correlation with Phase Transform

The generalised cross-correlation with phase transform between two microphone signals $\boldsymbol{x}_n$ and $\boldsymbol{x}_m$, with Fourier transforms $\boldsymbol{X}_n(f)$ and $\boldsymbol{X}_m(f)$ for the n^{th} and m^{th} microphone respectively, is defined as [18]

$$R_{nm}(\tau) = \int_{-\infty}^{\infty} \boldsymbol{\Psi}_{nm}(f)\boldsymbol{X}_n(f)\boldsymbol{X}_m^*(f)e^{2\pi i f \tau_{nm}}df, \tag{1}$$

where $\boldsymbol{\Psi}_{nm}(f) = \frac{1}{|\boldsymbol{X}_n(f)\boldsymbol{X}_m^*(f)|}$ is the phase transform weighing and * denotes the complex conjugate transpose. If it is assumed that there is one global optimum at the correct time lag, the TDOA can be estimated as

$$\tau_{nm} = \text{argmax } R_{nm}(\tau). \tag{2}$$

After estimating the time lag between each possible microphone pair, the location of the sound source can be estimated using an iterative least-squares estimation. For this, a modelled time lag is used, which is determined using the given microphone pair and the estimated location of the sound source.

2.2 Beamforming

A frequently applied technique for acoustic imaging is conventional frequency domain beamforming. It is robust, fast and intuitive [19]. For the recorded signal of a microphone array with N microphones, consider the Fourier-transformed pressure vector $\boldsymbol{p}(f)$:

$$\boldsymbol{p}(f) = \begin{pmatrix} p_1(f) \\ \vdots \\ p_N(f) \end{pmatrix}, \tag{3}$$

with $p_n(f)$ the Fourier transform of the time domain recorded signal of the n$^{\text{th}}$ microphone. From this, the cross-spectral matrix (CSM) can be obtained according to

$$\boldsymbol{C} = \langle \boldsymbol{p}\boldsymbol{p}^* \rangle, \tag{4}$$

where * is the complex conjugate transpose and $\langle\rangle$ denotes the ensemble average. Furthermore, the elements of steering vector $\boldsymbol{g}_j(f)$ for grid point j and frequency f is defined as

$$g_{n,j}(f) = \frac{e^{-2\pi i f \frac{r_{n,j}}{c}}}{r_{n,j}}, \tag{5}$$

where $r_{n,j}$ is the distance between the n$^{\text{th}}$ microphone and grid point j. c Is the speed of sound. The beamform output can be computed according to

$$B_j(f) = \frac{\boldsymbol{g}_j^*\boldsymbol{C}\boldsymbol{g}_j}{||\boldsymbol{g}_j||^4}. \tag{6}$$

This is done for all points on a predefined grid parallel to the array plane. This grid is considered at a predefined altitude z and can be calculated for a frequency range of interest. The grid in x- and y-direction is now limited to go from -100 m to +100 m with steps of 0.4 m, as it is expected that the localisation range will be lower than 100 m. For a selected time step and frequency range, the grid point where $\boldsymbol{B}_j(f)$ is the highest is considered to be the estimated drone location.

2.3 Differential Evolution

The objective of differential evolution is to minimise an energy function. For this research, the energy function is the Bartlett energy function [6], which is based upon the beamforming expression in Equation 6. This optimisation problem can be formulated as

$$\boldsymbol{g}_j(\boldsymbol{r}_j, f) = \text{argmin -}B_j(f), \tag{7}$$

where the distance vector $\boldsymbol{r}_j$ is calculated from all microphones to a candidate solution $[x_j \;\; y_j \;\; z_j]^T$. A population consisting of candidate solutions is created and are evolved over a predefined amount of generations towards an optimal estimation. The reason for choosing the energy function to be similar to beamforming is to investigate if this technique can achieve similar performance in a grid-free approach.

3. Experimental Setup

In October 2022, flight tests were conducted at the Dutch military base *Luitenant-generaal Bestkazerne*. This former air base provides a large, flat and open grassy area next to a runway. A map of this location is included in Figure 1, including the used reference frame. The y-axis is defined parallel to the runway, whereas the x-axis is defined orthogonal to and pointing away from the runway. During the measurement campaign, an acoustic array with 64 (PUI Audio 665-POM-2735P-R) microphones distributed in an Underbrink spiral configuration [20] was used. This array has been designed and tested by the Delft University of Technology [21]. The configuration has been chosen such that it performs well for a broad range of frequencies considering both the maximum side lobe level and the array resolution [22]. The dimensions of the array are 4×4 m and the microphone signals are sampled with a frequency of 50000 Hz. The microphones are plugged into a metal structure which is covered by foam to prevent ground reflections. Furthermore, the microphones are covered by windshields to lower the noise due to winds.

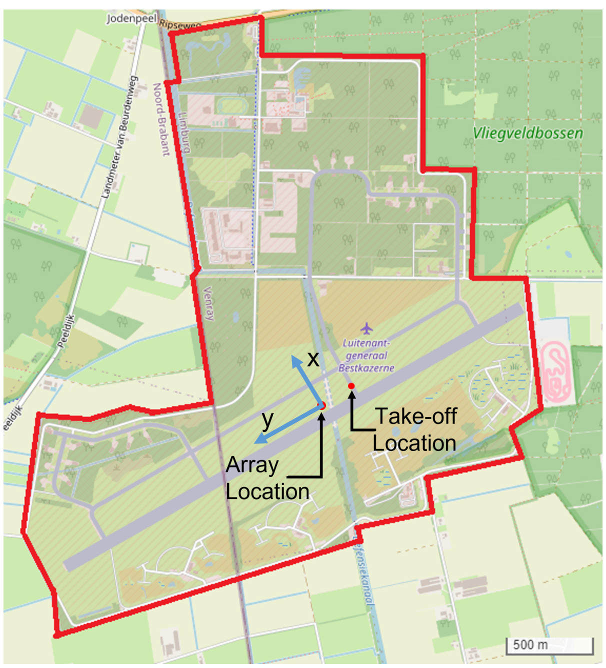

Figure 1: Map of the *Luitenant-generaal Bestkazerne* [23] with array and take-off location

Figure 2: Acoustic array used during the measurement campaign

Five different UAVs were flown, with a total of 26 flights. All drones flew at least 3 times, in a straight line from take-off location as indicated in Figure 1, over the array and back, parallel to the runway. During these flights, the altitude and flight speed were kept as constant as possible. Additionally, some UAVs were flown more than 3 times, where during the remaining flights various flight paths were flown, in combination with a variety of speeds, altitudes and manoeuvres. An overview of the drones that were flown, the number of flights per drone and their take-off weight is given in Table 1.

Table 1: Details on drone numbering, number of flights per drone and take-off weight

Nomenclature	Drone Type	Flights	Weight
Drone 1	DJI Mavic 3	5	895 g
Drone 2	Autel EVO II	4	1191 g
Drone 3	DJI Mini 2	3	242 g
Drone 4	DJI Phantom 3	3	1216 g
Drone 5	DJI Phantom 4	4	1380 g

For each drone, the first flight was chosen for testing the different localisation techniques. During this flight, the drone flew a straight line over the array and back, during which it flew over the array twice. For analysis, a segment with one fly-over was chosen. The flight number and foam and windshield conditions for the chosen flights are indicated in Table 2. It should be noted that due to the set-up of the experiment, the location of the drone varied from a few hundreds of meters in y-direction, to only a few meters in x-direction. For performance analysis, the estimation in y-direction will therefore be compared.

4. Results and Discussion

To perform a fair comparison between different techniques, but especially to compare among the different drones, first drone characteristics will be discussed. Two characteristics that are relevant for the performance of localisation techniques are the signal-to-noise ratio and the selected range of frequencies. After this has been established for the different drones, the techniques can be assessed and compared.

4.1 Drone Characteristics

For each drone, the signal-to-noise ratio will be determined according to [5]

$$R_{SNR} = 10 \log_{10} \frac{|X_{drone}(f)|^2}{|X_{noise}(f)|^2}, \tag{8}$$

where $X_{drone}(f)$ is the Fourier transformed signal of one fly-over of the drone. To obtain the Fourier transformed signal of the background noise, $X_{noise}(f)$, several measurements were taken without drones flying. The obtained ratios are summarised in Table 2 in the second column. It is clear that both drone 4 and 5 are loudest and it is therefore expected that for these flights the localisation ranges will be highest.

Table 2: Details on the SNR and frequency band used for analysis

Nomenclature	SNR [dB]	Frequencies [Hz]
Drone 1	2.393	250-2000
Drone 2	1.593	300-2000
Drone 3	0.187	250-2000
Drone 4	3.195	300-2000
Drone 5	3.043	250-2000

Another aspect that will influence this range is the frequency range that is selected for analysis. For all drones, the range with most pronounced harmonics of the rotors is between 100 and 2000 Hz. However, due to wind noise at lower frequencies, the lower bound of this region of interest has been raised to 250 of 300 Hz for all drones, based upon visual inspection of the spectrograms. The chosen ranges can be found in Table 2 in the last column.

4.2 Localisation Performance

To assess the localisation performance and compare results of different drones and techniques, the same evaluation method as described in [12] is chosen. The authors propose to calculate the absolute error of the estimated location and true location taken from the onboard GPS tracker and present these errors in box plots for set distance ranges. For the current study, the distance range from 0 to 100 m is subdivided into five 20-meter-wide distance bins. The error is determined as $\epsilon = |y_{true} - y_{est}|$.

For conciseness, a summary of the median values for all flights and methods within the 0 to 20 m bin is provided in Figure 3. From this figure, a clear ranking among the three methods is observed. Beamforming localises the drones best, followed by differential evolution and GCC-PHAT. These results are reasonable, as beamforming is an exhaustive grid search technique and therefore the most accurate. On the other hand, differential evolution is a grid-free approach which requires some iterations before getting to a good estimation, therefore being faster. In this study, it is less accurate as it also estimates the z-position of the drone. GCC-PHAT requires noisy correlation results of the time-difference-of-arrival to estimate the location using an iterative least-squares estimation, making it the least accurate method. Another observation is that the localisation performance is worst for drone 3 and best for drone 4. This is a logical result when the signal-to-noise ratios in Table 2 are compared. Drone 3 shows the lowest SNR and drone 4 the highest. A last observation is that within this bin of 0 to 20 meter, beamforming and differential evolution show median absolute errors of maximum 1.65 m and 1.87 m respectively. This is a low error of only 10%, which means that the estimation can be regarded as accurate.

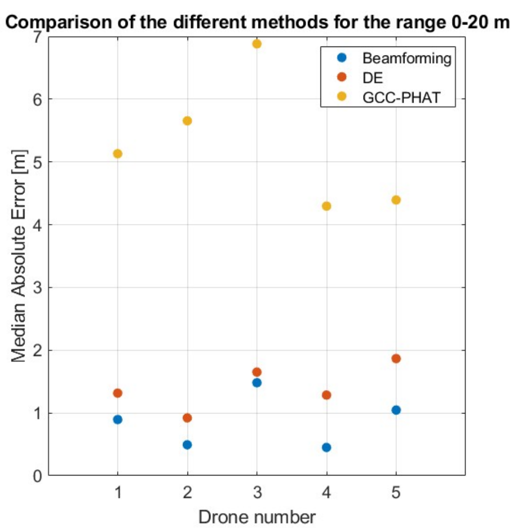

Figure 3: Medians for the range 0 - 20 m

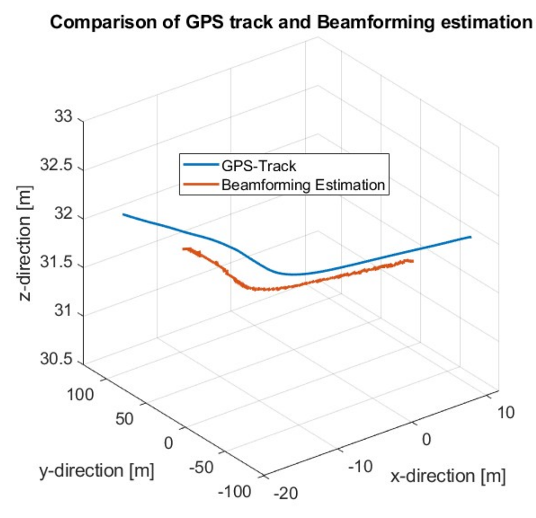

Figure 4: Beamforming estimation for drone 4

In Figure 4, the GPS track of drone 4 is plotted together with the estimation of the drone location by beamforming. This drone is chosen, as it is the most accurate estimation based upon previous analysis. The results for beamforming are included for the ranges in which the estimation is stable. In the y-direction, the localisation performance is good. Opposed to this, in the x-direction a constant offset can be observed. For other techniques and drones, similar offsets are noticed. This raises concerns regarding the accuracy of the calculated GPS tracks. It is likely that this discrepancy is a result of offsets between

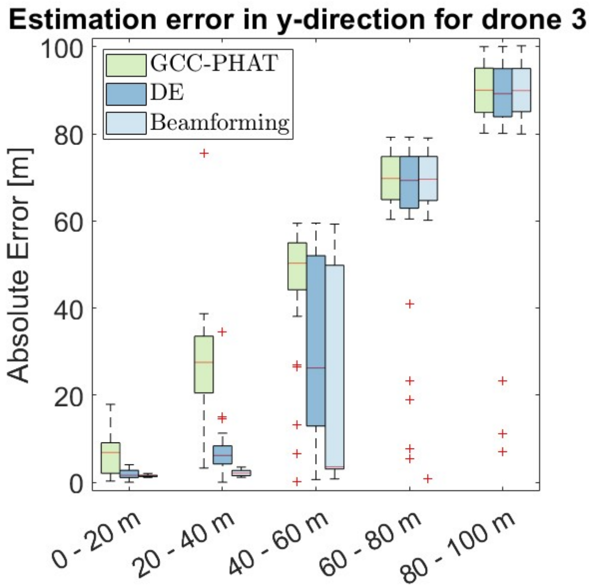

Figure 5: Results for drone 3

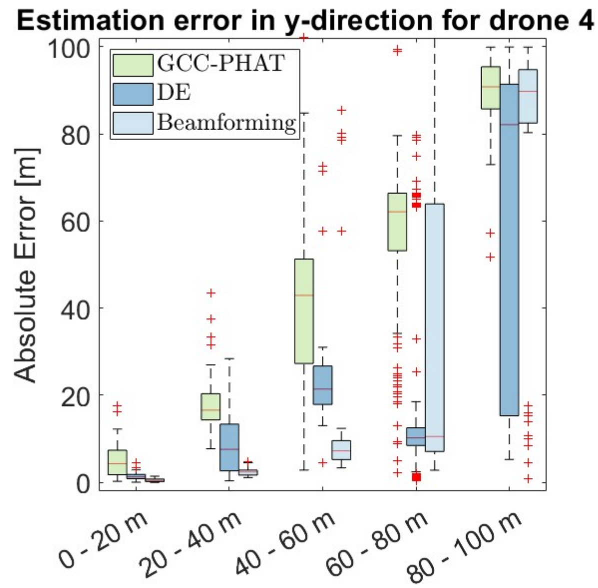

Figure 6: Results for drone 4

the onboard GPS track of the drone and the measured GPS location and orientation of the array. It is still fair to compare the different drones and methods, as the bias is systematic for all methods and drones.

Figure 5 and Figure 6 show the box plots of drone 3 and 4 respectively. These figures once more confirm the superiority of beamforming compared to the other two localisation techniques. Above this, the impact of the signal-to-noise ratio on the drone localisation using beamforming can be seen in these figures. When comparing the localisation ranges, it can be concluded that for drone 3, which can be described as the most silent drone, the localisation fails beyond 40 m. For drone 4, this only happens beyond 60 m, which indicates an increase in localisation range of approximately 150%.

Even though localisation is possible for all drones and localisation ranges of close to 80 m can be achieved, more emphasis can be put on tones emerging from drone rotors through data enhancement to improve the localisation range. Next to this, estimating the position of the drone in Cartesian coordinates is less accurate when drones are flying in the far-field. It is therefore useful to employ a spherical coordinate system and estimate the azimuth and elevation angle. By doing so, the different methods only need to estimate two variables.

5. Conclusion

In this study, GCC-PHAT, beamforming and differential evolution are tested and compared on five drone flights. The most accurate technique is beamforming, with localisation ranges varying between 40 to close to 80 m, depending on the signal-to-noise ratio of the drone. These results are achieved without enhancing any signals, which leaves room for improvement. Another way to improve this range is estimating the azimuth and elevation angle instead of the Cartesian coordinates. This reduces the number of variables that must be estimated and can therefore lead to faster and more accurate estimations.

References

1. I. Güvenc, F. Koohifar, S. Singh, M. L. Sichitiu, and D. Matolak, "Detection, Tracking, and Interdiction for Amateur Drones," *IEEE Communications Magazine*, vol. 56, pp. 75–81, 2018.

2. L. P. Lamotte, V. Baron, and S. Bouley, "UAV detection from acoustic signature: requirements and state of the art," pp. 1–13, 2020.

3. B. D. V. Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, pp. 4–24, 1988.

4. H. Pujol, É. Bavu, and A. Garcia, "BeamLearning: An end-to-end deep learning approach for the angular localization of sound sources using raw multichannel acoustic pressure data," *The Journal of the Acoustical Society of America*, vol. 149, p. 4248, 2021.

5. Z. Shi, X. Chang, C. Yang, Z. Wu, and J. Wu, "An Acoustic-Based Surveillance System for Amateur Drones Detection and Localization," *IEEE Transactions on Vehicular Technology*, vol. 69, pp. 2731–2739, 2020.

6. A. M. N. Malgoezar, M. Snellen, and R. Merino-Martinez, "On the use of global optimization methods for acoustic source mapping," *The Journal of the Acoustical Society of America*, vol. 141, pp. 453–465, 2017.

7. X. Chang, C. Yang, J. Wu, X. Shi, and Z. Shi, "A Surveillance System for Drone Localization and Tracking Using Acoustic Arrays," in *2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop*, pp. 573–577, 2018.

8. B. Harvey and S. O'Young, "A harmonic spectral beamformer for the enhanced localization of propeller-driven aircraft," *Journal of Unmanned Vehicle Systems*, vol. 7, pp. 156–174, 2019.

9. V. Baron, S. Bouley, M. Muschinowski, J. Mars, and B. Nicolas, "Drone localization and identification using an acoustic array and supervised learning," in *Proc. SPIE 11169, Artificial Intelligence and Machine Learning in Defense Applications*, p. 19 September 2019, 2019.

10. T. Blanchard, J.-H. Thomas, and K. Raoof, "Acoustic localization and tracking of a multi-rotor unmanned aerial vehicle using an array with few microphones," *The Journal of the Acoustical Society of America*, vol. 148, pp. 1456–1467, 2020.

11. G. Herold, A. Kujawski, C. Strümpfel, S. Huschbeck, M. U. de Haag, and E. Sarradj, "Flight path tracking and acoustic signature separation of swarm quadcopter drones using microphone array measurements," in *Quiet Drones 2020 - International e-Symposium on UAV / UAS Noise*, pp. 1–19, 2020.

12. M. Blass and F. Graf, "A Real-Time System for Joint Acoustic Detection and Localization of UAVs," in *Quiet Drones 2020 - International e-Symposium on UAV / UAS Noise*, pp. 1–18, 2020.

13. M. Blass, A. Maly, F. Graf, P. Wellig, and B. Ott, "Towards mobile microphone array based UAV tracking," in *Quiet Drones 2022 - International e-Symposium on UAV / UAS Noise*, pp. 1–18, 2022.

14. A. Kujawski, G. Herold, and E. Sarradj, "A deep learning method for grid-free localization and quantification of sound sources," *The Journal of the Acoustical Society of America*, vol. 146, pp. EL225–EL231, 9 2019.

15. P. Castellini, N. Giulietti, N. Falcionelli, A. F. Dragoni, and P. Chiariotti, "A neural network based microphone array approach to grid-less noise source localization," *Applied Acoustics*, vol. 177, p. 107947, 6 2021.

16. É. Bavu, H. Pujol, A. Garcia, C. Langrenne, S. Hengy, O. Rassy, N. Thome, Y. Karmim, S. Schertzer, and A. Matwyschuk, "Deeplomatics: A deep-learning based multimodal approach for aerial drone detection and localization," in *Quiet Drones 2022 - International e-Symposium on UAV / UAS Noise*, pp. 1–13, 2022.

17. B. von den Hoff, R. Merino-Martínez, D. G. Simons, and M. Snellen, "Using global optimization methods for three-dimensional localization and quantification of incoherent acoustic sources," *JASA Express Letters*, vol. 2, 2022.

18. J. H. DiBiase, H. F. Silverman, and M. S. Brandstein, "Robust Localization in Reverberant Rooms," in *Microphone Arrays — Signal Processing Techniques and Applications*, pp. 157–180, Berlin, Heidelberg: Springer-Verlag, 2001.

19. R. Merino-Martínez, P. Sijtsma, M. Snellen, T. Ahlefeldt, J. Antoni, C. Bahr, D. Blacodon, D. Ernst, A. Finez, S. Funke, T. Geyer, S. Haxter, G. Herold, X. Huang, W. Humphreys, Q. Leclère, A. Malgoezar, U. Michel, T. Padois, A. Pereira, C. Picard, E. Sarradj, H. Siller, D. Simons, and C. Spehr, "A review of acoustic imaging methods using phased microphone arrays," *CEAS Aeronautical Journal*, vol. 10, pp. 197–230, 2019.

20. James R. Underbrink, "Circularly symmetric, zero redundancy, planar array having broad frequency range applications," 2001. U.S. Patent Number 6,205,224 B1.

21. A. Vieira, M. Snellen, and D. G. Simons, "Experimental Assessment of Sound Quality Metrics for Takeoff and Landing Aircraft," *AIAA Journal*, vol. 59, no. 1, pp. 240–249, 2021.

22. Z. Prime and C. Doolan, "A comparison of popular beamforming arrays," in *Annual Conference of the Australian Acoustical Society 2013, Acoustics 2013: Science, Technology and Amenity*, pp. 151–157, 01 2013.

23. OpenStreetMap. `www.openstreetmap.org/#map=14/51.5217/5.8461`, 2022. (Accessed on 16-12-2022).