DOI: 10.61782/fa.2025.0253

FORUM ACUSTICUM EURONOISE 2025

CORRECTION OF AZIMUTH AMBIGUITIES OF THE GCC ALGORITHM IN A UNIFORM CIRCULAR ARRAY

Resino Viñas, Alejandro[1] **Corral García, Gonzalo**[1] **Tejera-Berengue, Diana**[1]
Zhu-Zho, FangFang[1] **Rosa Zurera, Manuel**[1]
[1] Department of Signal theory and Communications, University of Alcalá, Spain

ABSTRACT

Array signal processing faces growing challenges in real-world applications where the number of available microphones or computational resources is limited. These applications include industrial device monitoring for anomaly detection and drone localization in surveillance tasks. This paper explores the challenges of implementing direction-of-arrival estimation for audio signals using Uniform Circular Arrays, assessing the feasibility of real-time implementation and addressing potential angular ambiguities in azimuth through different microphone coupling strategies. The generalized cross-correlation algorithm is applied to a uniform circular array of eight microphones, implemented on the MATRIX Creator, an IoT device connected to a Raspberry Pi 3B.

Keywords: *direction of arrival estimation, generalized cross correlation, real time, angular ambiguities, circular array.*

1. INTRODUCTION

A microphone array consists of a set of microphones arranged in a way that allows for the proper estimation of certain signal parameters or characteristics using spatio-temporal and frequency information available at the array's output [1–3]. An important problem, known as Direction of Arrival (DOA) estimation, arises in the localization of sound sources, which is a fundamental problem in many applications, including surveillance, robotics, and sound field analysis [4]. The ability to accurately estimate the DOA of sound sources plays a crucial role in enhancing the performance of systems such as digital hearing aids, and autonomous vehicles. Among various configurations for microphone arrays, circular arrays have gained attention due to their unique geometrical properties, which offer several advantages in source localization tasks.

Circular microphone arrays, typically arranged in a uniform circular pattern, also known as UCA, provide rotational symmetry, making them particularly well-suited for applications where the source can be located in a full 360° range of directions [5]. The inherent symmetry of such arrays simplifies the estimation of direction, reduces ambiguities, and enhances the robustness of localization algorithms. However, this geometry also introduces challenges, particularly in dealing with issues such as noise, reverberation, and the need for precise time delay estimation between microphones.

In this paper, we explore the use of DOA estimation techniques with circular microphone arrays, focusing on algorithms like the Generalized Cross-Correlation (GCC), which have been widely used for time delay estimation between microphone pairs. We also discuss strategies for addressing the challenges of multi-microphone integration, the selection of optimal microphone pairs, and the combination of multiple time delay estimates to achieve accurate localization in real-world environments. The paper aims to provide an in-depth analysis of the methods, their advantages, and their limitations when applied to acoustic source localization using circular arrays.

In the case of a single source, DOA can be determined by calculating the time delay between a pair of microphones. The time delay between two signals is usually obtained through correlation. However, this pro-

**Corresponding author*: *manuel.rosa@uah.es.*

cess is highly sensitive to the non-stationary nature of wideband signals, indoor reverberation effects, and interference from signals with similar spectral characteristics. These problems can be mitigated by applying pre-filtering to the signals, leading to the GCC algorithm [6], a technique used to estimate the time difference of arrival (TDOA) between signals captured by multiple microphones. By analyzing the similarity between the signals over time, GCC identifies the time delay at which they are most aligned. This time delay provides crucial information about the direction from which the sound originated. To improve accuracy, a weighting function—such as the Phase Transform (PHAT)—is often applied to enhance the robustness of the correlation against noise, reverberation, and interfering signals. Once the TDOAs are determined for different microphone pairs, they are used alongside the known microphone positions to estimate the Direction of Arrival (DOA) of the sound source.

Another challenge to consider is selecting which microphones to use for calculating the time delay. The GCC algorithm is based on the calculation of correlation between two microphones, but arrays usually have more elements. When using multiple microphone pairs, it's crucial to determine how to effectively combine the resulting solutions. In this paper, we study three ways of pairing microphones to apply the GCC algorithm to a circular microphone array (consecutive elements, opposite elements, and elements forming a square), presenting results from a case study, in which the array is built with an IoT device, called Matrix-Creator, which is connected to a Raspberry-Pi. The estimated DOA is determined by calculating the mean of the estimated angles in consecutive time frames. An ingenious method is applied to resolve the ambiguities inherent in the GCC algorithm with two microphone arrays.

The paper is organised as follows. After the introduction, Section 2 presents the fundamentals of array processing used in the paper. Section 2.2 introduces the GCC algorithm. Section 3 deals with DOA estimation with a microphone circular array. Section 4 contains the main results of this research. Finally, Section 5 presents the research conclusions.

2. FUNDAMENTALS OF ARRAY SIGNAL PROCESSING

This section presents an introduction to the techniques and algorithms used in this paper for DOA estimation with microphone arrays.

2.1 Uniforma linear arrays

The geometry of the array plays an important role in the formulation of processing algorithms, and thus, in source localisation problems [2, 3]. We start with a Uniform Linear Array (ULA) of microphones, where multiple microphones are arranged in a straight line with equal spacing between them. For far-field applications, the signal acquired by the i-th microphone ($x_i(t)$) can be related to a reference signal $s(t)$, usually the signal impinging on the first microphone of the array.:

$$x_i(t) = \alpha s(t - \delta_{i1}) + n_i(t) \tag{1}$$

where α represents the attenuation and $n_i(t)$ is the noise at the i-th microphone. The signal $x_i(t)$ is a delayed version of $s(t)$, and the delay δ_{ij} between the signals acquired by the i-th and j-th microphones can be calculated using Equation (2), where d is the inter-element spacing and θ is the angle of the impinging wave relative to the normal of the array and c is the speed of sound.

$$\delta_{ij} = \frac{(i-j)d\sin(\theta)}{c} \tag{2}$$

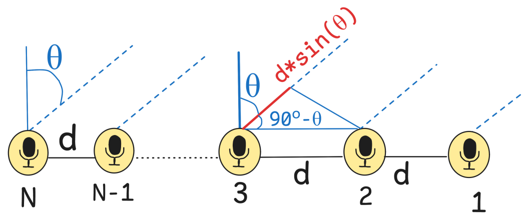

Figure 1: Array ULA de N micrófonos.

In narrow-band applications, the signal acquired by the $i-th$ microphone is expressed as follows, taking the signal at the first microphone as reference:

$$x_i(t) = x_1(t)e^{j\omega_0\tau_{1i}} = x_1(t)e^{j(i-1)\frac{2\pi d sin(\theta_i)}{c}} \tag{3}$$

The received signals can be expressed in matrix form, as shown in Eqn. (4), where $\gamma(\theta)$ represents the steering vector (Eqn. (5)), which contains the phase shift of the sound source received at each of the microphones in the array.

$$\boldsymbol{x}(t) = \boldsymbol{\gamma}(\theta)\boldsymbol{s}(t) + \boldsymbol{n}(t) \tag{4}$$

$$\gamma(\theta_i) = \begin{bmatrix} 1 & e^{j\frac{2\pi d sin(\theta_1)}{c}} & \dots & e^{j\frac{(N-1)2\pi d sin(\theta_i)}{c}} \end{bmatrix}^T \tag{5}$$

In linear microphone arrays, ambiguity in DOA estimation arises because the array can't distinguish between sound sources arriving from symmetrical angles on either side of the array axis. This problem is showed in Fig. 2. This front-back ambiguity is a limitation of the array's geometry and can lead to incorrect angle estimation unless additional spatial cues or array configurations (like circular or planar arrays) are used to resolve it.

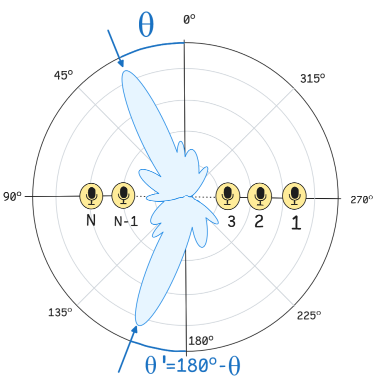

Figure 2: Illustration of ambiguity in DOA estimation for a ULA.

2.2 GCC algorithm

The Generalized Cross-Correlation algorithm is a widely used method for estimating the time difference of arrival between signals captured by pairs of microphones. This time delay information is essential for determining the DOA of a sound source. Unlike basic cross-correlation, GCC enhances the estimation process by applying weighting functions in the frequency domain to improve robustness against noise and reverberation. One popular variant, GCC-PHAT, emphasizes phase information to achieve more accurate results in challenging acoustic environments. Due to its efficiency and effectiveness, GCC is commonly used in real-time audio processing applications such as speaker tracking, acoustic localization, and beamforming [6, 7].

To estimate the delay δ between the two signals, the cross-correlation function, $R_{x_1,x_2}(\tau) = E[x_1(t)x_2(t-\tau)]$ is commonly used, which reaches its maximum at the estimated delay. It can be approximated by Eqn. (6):

$$\hat{R}_{x_1,x_2}(\tau) = \frac{1}{T-\tau}\int_{\tau}^{T} x_1(t)\, x_2(t-\tau)\, dt, \tag{6}$$

The cross-correlation can also be computed as the inverse Fourier transform of the cross-spectrum, which is less computationally expensive than the time domain calculation:

$$R_{x_1,x_2}(\tau) = \mathscr{F}^{-1}\{X_1(w)X_2^*(w)\} = \mathscr{F}^{-1}\{G_{x_1,x_2}(\omega)\} \tag{7}$$

If the signals described in are considered, and the noise is assumed to be uncorrelated with the signals, the following cross-correlation is obtained:

$$\begin{aligned} R_{x_1,x_2}(\tau) = \alpha R_{s_1,s_2}(\tau-\delta) + R_{n_1,n_2}(\tau) = \\ R_{x_1,x_2}(\tau) = \alpha R_{s_1,s_2}(\tau) * \delta(\tau-\delta) \end{aligned} \tag{8}$$

$$G_{x_1,x_2}(\omega) = \alpha G_{s_1,s_2}(\omega)e^{-j\omega\delta} + G_{n_1,n_2}(\omega) \tag{9}$$

The Generalized Cross-Correlation (GCC) is an extension of the standard cross-correlation, where the signals are pre-filtered to improve the accuracy of the delay estimation. GCC is calculated with expression Eqn. (10)

$$\begin{aligned} R_{y_1,y_2}(\tau) = \mathscr{F}^{-1}\{H_1(\omega)H_2(2)X_1(\omega)X_2^*(\omega)\} = \\ \mathscr{F}^{-1}\{H(\omega)X_1(\omega)X_2^*(\omega)\} = \\ \mathscr{F}^{-1}\{H(\omega)G_{x_1,x_2}(\omega)\} = \\ \int_{-\infty}^{\infty} H(\omega)G_{x_1,x_2}(w)e^{j\omega\tau}d\omega \end{aligned} \tag{10}$$

The most commonly used method in the literature is the PHAT filter, $H_P(\omega) = \frac{1}{|G_{x_1,x_2}(\omega)|}$, which gives rise to the following expression for the cross-correlation:

$$\hat{R}_{y_1,y_2}(\tau) = \int_{-\infty}^{\infty} \frac{\hat{G}_{x_1,x_2}(\omega)}{|G_{x_1,x_2}(\omega)|} e^{j\omega\tau} d\omega \tag{11}$$

Assuming $\hat{G}_{x_1,x_2}(\omega) \approx |G_{x_1,x_2}(\omega)|e^{-j\omega\delta}$, the following expression is obtained:

$$\hat{R}_{x_1,x_2}(\tau) \approx \int_{-\infty}^{\infty} e^{-j\omega\delta} e^{j\omega\tau} d\omega \approx \delta(t-\delta) \tag{12}$$

The GCC-PHAT algorithm is particularly useful for reducing the effect of reverberation.

3. DOA CALCULATION WITH A UCA BASED ON GCC

The GCC algorithm can be used to calculate the delay between two microphones $\delta_{i,j}$, obtaining the angles of arrival $\theta_{i,j}$ by solving Equation Eqn. (3). In a circular array with N microphones, the number of unique microphone pairs you can form is given by the number of combinations of N elements taken 2 at a time:

$$\text{N}^{\text{o}}\text{ microphone pairs} = \binom{N}{2} = \frac{N(N-1)}{2} \tag{13}$$

Each pair of microphones produces two possible solutions to the DOA estimation problem, due to ambiguity with respect to the array axis. Another issue to consider is that each microphone pair calculates the angle with respect to the normal of the line connecting the two microphones, so the resulting solutions are not referenced to a common reference, as shown in Fig. 3.

Corrections must be applied to refer all DOA estimates to a common reference axis, using Equation (14), where α is the angular correction to be implemented, l is the distance between the two considered elements, c is the sound propagation velocity, and i is the index of the reference microphone in each pair. The angular correction in the azimuth plane depends on the geometric position of the pair or microphones, (Mic_i, Mic_j), from which δ_{ij} has been computed. After the correction, all solutions are referenced to the axis normal to the first microphone pair used to compute the initial delay (see Fig. 3).

$$\beta_k = \sin^{-1}(\frac{c}{l} \cdot \delta_{ij}) - (i-1) \cdot \alpha \tag{14}$$

Angles estimated in this way are referenced to the normal of the axis connecting the first pair of microphones and can be corrected to refer to that axis, resulting in the following values: $\gamma_k = \pm(\frac{\pi}{2} - \beta_k)$, one of which arises from the inherent ambiguity of linear arrays. An additional correction angle, Ω, can be applied to reference all angles to a chosen symmetry axis of the circular array. From this point on, consider γ_k to represent the angles referenced to the symmetry axis of the microphone array, which is taken as a reference.

Once the DOA estimates from each pair of microphones have been referenced to the chosen symmetry axis, a higher concentration of solutions is expected around the true direction, while the angles estimated due to the linear array ambiguity tend to appear scattered. Thus, the angles are assigned to one of four quadrants $(q(j), j = 1, \cdots, 4)$, and the correct quadrant is selected as the one that accumulates the most estimates. The final DOA estimate is then calculated as the average of the number of angles within the selected quadrant, $n(q(j))$.

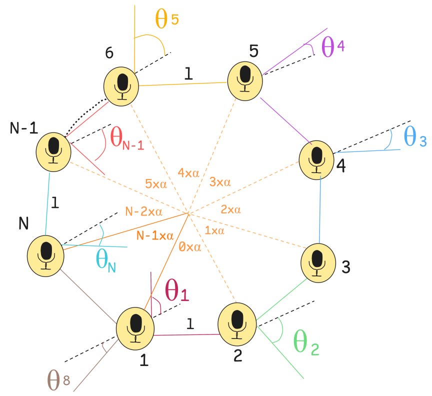

Figure 3: Angular correction $(i-1)\alpha$ on θ_i for each microphone binomial of the UCA array in GCC-1

$$D\hat{O}A = \frac{1}{n(q_j)} \sum_{k=0}^{n(q_j)} \gamma_k \tag{15}$$

3.1 First approach: GCC with consecutive microphones (GCC-1)

In the first approach, microphone pairs are formed by adjacent ones $(\{Mic_1, Mic_2\}, ..., \{Mic_N, Mic_1\})$. N pairs can be formed, therefore, N delays are obtained, δij, $i \in \{1, ..., N\}$, and $j = (i+1) \; mod \; N$. These delays are used to calculate the DOA angles $\theta_i = \sin^{-1}\left(\frac{\delta_{i,i+1} c}{l}\right)$, where $l = 2R\sin\left(\frac{\pi}{N}\right)$ is the inter-element spacing, and from them, the angles referenced to the microphone array axis, $\gamma_i, i \in \{1, ..., N\}$, are calculated. Each microphone pair produces two results, due to the inherent ambiguity of linear arrays.

The relation between θ_i and the corresponding γ_i is obtained with $\alpha = 2\pi/N$. This case is represented in Fig. 3. As the axis formed by $\{Mic_1, Mic_2\}$ is parallel to the cartesian coordinates axis in that figure, an additional correction is unnecessary to reference the DOA angles with the abscissa axis.

3.2 Second approach: GCC between opposite microphones (GCC-2)

In this case, the pairs are formed by microphones located at diametrically opposite positions in the circular array. The number of microphones in the array must be even, and the pairs are formed as follows: $\{Mic_1, Mic_{N/2+1}\}, \cdots, \{Mic_{N/2-1}, Mic_N\}$. The number of pairs is $N/2$. The smaller number of possible DOA estimates is compensated by the larger aperture of the linear array formed by each pair, which results in a narrower main lobe and a more accurate DOA estimate from each microphone pair.

The angles θ_i would be obtained in this case using the following expression, where $l = 2R$:

$$\theta_i = \sin^{-1}(\frac{c}{l} \cdot \delta_{i,i+\frac{N}{2}}), i \in \{1, \cdots, \frac{N}{2}\} \tag{16}$$

The relation between θ_i and the corresponding γ_i is obtained again with $\alpha = 2\pi/N$. This case is represented in Fig. 4. If the line connecting the pair $\{Mic_1, Mic_{N/2+1}\}$ is not aligned with the system's reference axis, the angle between both lines or axes must be corrected. If the first microphone is located at position (x, y) in the reference system centered at the middle of the circular array, that angle can be calculated as:

$$\Omega = \sin^{-1}(\frac{|y|}{R}) \tag{17}$$

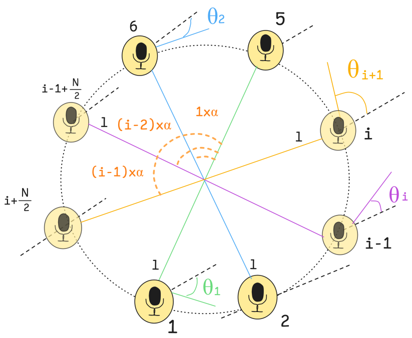

Figure 4: Circular array with pairs of microphones located at diametrically opposite positions.

3.3 Third approach: GCC applied to square-like microphone arrays (GCC-3)

GCC-1 generates N estimations of DOA, corresponding to N arrays of two microphones, but the distance between microphones is low. In contrast, GCC-2 generates only N/2 DOA estimates, but the distance between the microphones is longer, resulting in a narrower main lobe. One way to combine the positive aspects of both approaches is the use of arrays consisting of alternating microphones (one yes and one no). If the number of microphones is even, the number of pairs that can be formed matches the number of microphones N, and the length of each pair is longer than in the GCC-1 case. This organization is represented in Fig. 5.

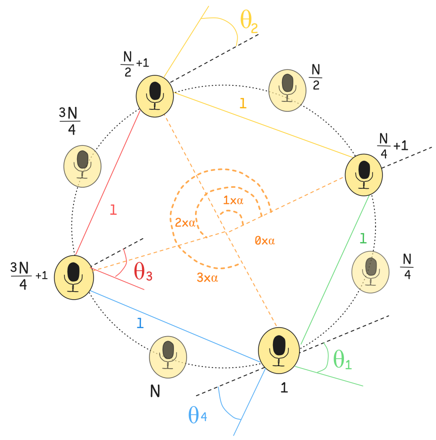

Figure 5: Circular array microphones with square-like shape. Microphones located at diametrically opposite positions.

This type of arrangement can be extended to arrays of $N = 4n$ microphones by pairing each microphone with a separate microphone at $N/4$ positions. N arrays of two microphones are thus formed, and the length of each array is equal to the length of the side of the square inscribed in the circle of radius R containing the microphones. With this reasoning, the angles θ_k are estimated with the following expression:

$$\begin{gathered} \theta_k = sin^{-1}(\tfrac{\delta_{ij} c}{l}), \\ i = 1, \cdots, N; \\ j = (i + \tfrac{N}{4}) \ mod \ N, \\ l = 2Rcos(\pi/4) \end{gathered} \tag{18}$$

Corrections to obtain angles γ_{ij} are implemented in a similar way to GCC-1 and GCC-2 algorithms.

4. RESULTS

In this section, we present the results of the evaluation of the three approaches previously described. The evaluation was carried out using two implementations: one in MATLAB and another in C++, running on a Raspberry Pi equipped with a Matrix Creator, an IoT device that includes a circular array of eight microphones, among other sensors [8], to evaluate the possibility of the real-time implementation of the algorithms running on a RaspberryPi.

A 5-second audio signal was played from an external source located one meter away from the array, with the DOA falling between microphones 5 and 6 (corresponding to an angle between 67.5º and 112.5º). The audio signal corresponds to a segment of Tchaikovsky's Nutcracker. The audio was recorded by the Matrix Creator at a sampling rate of 48 kHz. The signals were processed in frames of 512 samples each, and the GCC-1, GCC-2, and GCC-3 algorithms were applied.

In figs. 6 to 8, the results obtained using the MATLAB implementation of the three algorithms (GCC1, GCC2, and GCC3) are presented. The results from GCC1 are acceptable, showing a distribution centered around the true DOA, located between Mic_5 and Mic_6. The variance of the DOA estimates is high, as expected, given the small distance between the microphones. The results from GCC2 exhibit a multimodal distribution, which may be attributed to reverberation effects in the environment. However, the spread around each mode is low, due to the greater distance between microphones. Finally, the results from GCC3 show a distribution with relatively low variance and no multimodal behavior, effectively combining the best features of the other two methods.

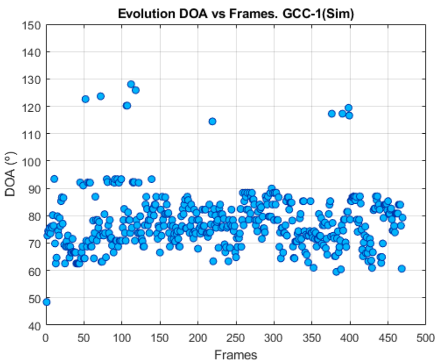

Figure 6: Evolution of DOA estimation using GCC-1 implemented with Matlab

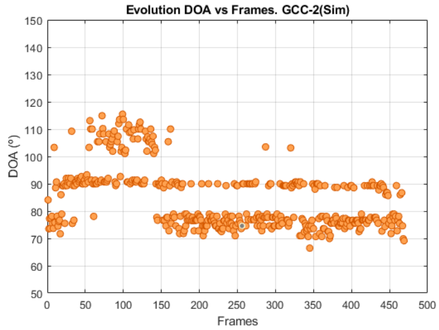

Figure 7: Evolution of DOA estimation using GCC-2 implemented with Matlab

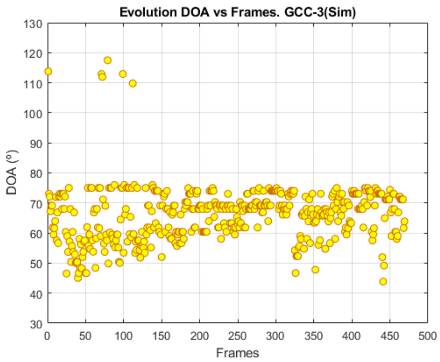

Figure 8: Evolution of DOA estimation using GCC-3 implemented with Matlab

Table 1 summarizes the execution times of the algorithms on the Raspberry Pi. GCC3 has been simplified, and only half of the possible linear arrays are implemented, forming a square shape array. The execution times indicate that the three configurations require a similar amount of time to process each frame, approximately 21 ms, and the total execution time to process the 5-second audio segment ranges between 9.95 s and 10 s. Thus, by computing the ratio $\frac{Sound\ Duration}{Total\ Time}$ and assuming a total time of around 10s and a sound duration of 5 seconds in all cases, the system operates in a pseudo real-time scenario, effectively processing one out of every two incoming frames. This result is acceptable to implement a DOA estimation system, where supplying the results of the estimation every 42ms allows the localization of the sound sources.

In terms of processing time (**TT**), as shown in Table 1, GCC-3 proves to be the least computationally expensive, with a total execution time of 0.98 seconds. In contrast, both GCC-1 and GCC-2 exceed one second due to the absence of the simplifications used in GCC-3. It is important to note that the processing time is mainly determined by the number of Fast Fourier Transforms (FFTs) that must be computed, rather than the number of correlations. All algorithms require computing the FFT of the signals acquired by each microphone considered.

Table 1: Comparison of GCC algorithm computation times in Matlab simulation and Raspberry Pi implementation

Method	Matlab		Raspberry Pi	
	FPT [1]	TT [1]	FPT [1]	TT [1]
GCC-1	3.1 ms	1.5 s	21.33 ms	10 s
GCC-2	4.1 ms	1.916 s	21.2 ms	9.94 s
GCC-3	2 ms	0.98 s	21.2 ms	9.947 s

5. CONCLUSIONS

In this paper, three algorithms for estimating the Direction of Arrival (DOA) of a wideband sound signal through continuous frame-by-frame processing have been presented and evaluated. These methods combine the results of the Generalized Cross-Correlation (GCC) algorithm using a circular microphone array. The three approaches differ in how the microphones are paired to form two-element linear arrays, which are then used to apply the GCC algorithm. The effectiveness of each method is analyzed, and the results in terms of DOA estimation accuracy and computation time on a Raspberry Pi device are presented. The signals were acquired using the MATRIX Creator IoT device.

The proposed algorithms are capable of producing reliable localization results while operating in soft real-time, as demonstrated in Table 1, on a programmable device such as the Raspberry Pi, in combination with the MATRIX Creator microphone array.

The first approach pairs consecutive microphones. Although the distance between microphones is small, the number of linear arrays equals the number of microphones. The DOA estimation results are acceptable, with a mean value close to the true DOA, though the variance is relatively high.

[1] FPT (Frame Processing Time), TT (Total Time)

The second approach pairs diametrically opposed microphones. This increases the distance between elements, but the number of linear arrays is halved. The resulting estimates are more accurate, though the distribution is multimodal—likely due to reverberation effects in the environment.

The third approach pairs microphones in a way that the resulting linear arrays form a square-like shape. This method appears to combine the strengths of the previous two approaches, yielding accurate and stable DOA estimates.

Regarding execution time, it is important to note that this parameter is influenced not only by the number of linear arrays considered but also by the number of Fast Fourier Transforms (FFTs) required. In both the first and second approaches, FFTs must be computed for all acquired signals, resulting in similar processing times. A more in-depth analysis of execution time is needed, but the current evaluation suggests that DOA estimation using all three methods is feasible in pseudo real-time with the selected hardware. This makes the proposed algorithms suitable for scenarios where IoT devices are applicable, such as in industrial environments.

6. ACKNOWLEDGMENTS

This work was funded by the Spanish Ministry of Science and Innovation under Project PID2021-129043OB-I00 (funded by MCIN/AEI/10.13039/501100011033/FEDER, EU), and by the Community of Madrid and University of Alcalá under project EPU-INV/2020/003.

7. REFERENCES

[1] J. Benesty, J., C., and Y. Huang, *Microphone Array Signal Processing*. Springer, Berlin-Heidelberg: Springer Topics in Signal Processing, 2008.

[2] H. Krim and M. Viberg, “Two decades of array signal processing research: the parametric approach,” *IEEE Signal Processing Magazine*, vol. 13, no. (4), pp. 67–94, 1996.

[3] M. Pesavento, M. Trinh-Hoang, and M. Viberg, “Three more decades in array signal processing research: An optimization and structure exploitation perspective,” *IEEE Signal Processing Magazine*, vol. 40, pp. 92–106, June 2023.

[4] D. Desai1 and N. Mehendale, *A Review on Sound Source Localization Systems*, vol. 29 of

Archives of Computational Methods in Engineering, p. 4631–4642. Springer Nature, 2022.

[5] M. Gaber, T. Elnady, and A. Elsabbagh, "Sound source localization in 360 degrees using a circular microphone array," in *Euronoise 2018 - Conference Proceedings*, (Crete, Greece), pp. 2613–2620, European Acoustics Association, June 2018.

[6] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 4, no. 4, pp. 320–327, August 1976.

[7] B. Kwon, Y. Park, and Y. s. Park, "Analysis of the gcc-phat technique for multiple sources," *ICCAS*, pp. 2070–2073, 2010.

[8] "Matrix creator overview." https://matrix-io.github.io/matrix-documentation/matrix-creator/overview/, 2019. Accessed on April 7th, 2025.