

ACOUSTIC LOCALIZATION OF UAVS WITH MICROPHONE ARRAYS. COMPARISON OF DIFFERENT ARRAY GEOMETRIES.

Diana Tejera Berengué, FangFang Zhu Zhou, Manuel Utrilla Manso, Roberto Gil Pita, Manuel Rosa Zurera^{1*}

¹Departamento de Teoría de la Señal y Comunicaciones, Escuela Politécnica Superior.
Campus Universitario, Universidad de Alcalá
28805 Alcalá de Henares, Madrid

RESUMEN

Este trabajo presenta un estudio sobre detección y localización de UAVs utilizando arrays de micrófonos con diferentes geometrías. El estudio se basa en el resultado del análisis espectral de las señales emitidas, que permite identificar las bandas de frecuencia con mayor energía, y características especiales que pueden ser consideradas en el diseño del array de micrófonos. Para probar las diferentes geometrías con las mismas señales, se ha creado un modelo de señal de las señales emitidas por UAVs en movimiento en trayectorias controladas. Este modelo nos permite generar señales sintéticas para probar los algoritmos de localización. En el estudio se han considerado arrays lineales uniformes y arrays circulares, con diferente número de micrófonos y diferentes distancias entre micrófonos. Cada geometría tiene propiedades que pueden dificultar la estimación de la dirección de llegada en diferentes escenarios, y facilitarla en otros. El algoritmo de localización trabaja con tramas, cuya longitud es otro de los parámetros considerados en el estudio. Este estudio se realiza con tres drones cuadricópteros y un helicóptero no tripulado, demostrando que es posible una estimación precisa a distancias de cientos de metros del UAV con drones en movimiento, aunque el uso de micrófonos precisos y algoritmos más potentes augura la ampliación de este rango a distancias mayores.

ABSTRACT

This paper presents a study of UAVs detection and localization using microphone arrays with different geometries. The study is based on the result of the emitted signals spectral analysis, which allows to identify the frequency bands with most of the energy, and special features that can be considered in the microphone array design. To test the different geometries with the same signals, a signal model

of the signals emitted by moving UAVs in controlled trajectories has been created. This model allows us to generate synthetic signals for testing the localization algorithms. Uniform linear arrays and circular arrays, with different number of microphones and different inter-microphone distances, have been considered in the study. Each geometry has properties which can make estimation of the direction of arrival difficult in different scenarios, and easier in others. The localization algorithm works with frames, which length is another parameter considered in the study. This study is performed with three quadcopter drones and one unmanned helicopter, demonstrating that accurate estimation is possible at distances of hundreds of meters from the UAV with moving drones, although the use of precise microphones and more powerful algorithms portends the extension of this range to greater distances.

Palabras Clave— Localization, detection, arrays, microphones, UAV.

1. INTRODUCCIÓN

Los vehículos aéreos no tripulados (“Unmanned Aerial Vehicles”, UAV) representan una de las tecnologías más populares de las últimas décadas, debido a su amplia gama de aplicaciones, así como a su flexibilidad y coste moderado [1]. Se han utilizado para cometer actos ilícitos relacionados con la ciberseguridad, la privacidad y la seguridad pública [2]. La necesidad de detectar UAVs abrió una línea de investigación que se ha desarrollado a lo largo de las dos últimas décadas, proponiéndose diversas técnicas basadas en diferentes tipos de sensores y algoritmos de procesamiento de señal [3].

Los métodos desarrollados adolecen de varios inconvenientes y, hasta la fecha, no existe una solución universal que pueda aplicarse en todos los escenarios posibles. Las técnicas de inspección visual con cámaras,

* **Autor de contacto:** manuel.rosa@uah.es

Copyright: ©2023 First author et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

requieren una línea de visión clara, lo que no es posible en escenarios lluviosos, con niebla, o en presencia de obstáculos que produzcan ocultación. Las cámaras utilizadas pueden ser convencionales o de infrarrojos, pero en este último caso el coste puede ser elevado. La detección basada en radiofrecuencia trata de detectar las señales de comunicación utilizadas por los drones con antenas adecuadas. Puede ser una técnica eficaz, si el UAV se comunica con una estación de control o con otros UAVs, pero inútil con UAVs autónomos. La detección con sensores de radar, activos o pasivos, ha demostrado su eficacia y no se ve afectada por las condiciones de visibilidad, por ejemplo. Tienen una gran fiabilidad y sensibilidad, pero las técnicas de procesamiento que utilizan los hacen ineficaces para detectar drones estáticos o que se mueven a baja velocidad. Además, su coste puede ser muy elevado para muchas aplicaciones. Así, surge la detección acústica, como una alternativa en muchos escenarios, que puede complementar satisfactoriamente a los sistemas anteriores.

La detección acústica de vehículos con rotores se lleva a cabo en una serie de pasos, tratando en primer lugar de caracterizar la señal acústica y utilizando los resultados para diseñar el mejor sistema de sensores. Posteriormente debe seleccionarse el algoritmo más adecuado para estimar la dirección de llegada (DOA) de la fuente sonora. Con este fin se han utilizado el conformador de haz de respuesta sin distorsión de varianza mínima (“Minimum Variance Distorsionless Response”, MVDR) [4] o el algoritmo de clasificación de señales múltiples (MUSIC) [5]. Estos algoritmos procesan las señales adquiridas por un array de sensores acústicos (micrófonos).

Los arrays más utilizados son los lineales uniformes (ULA) [3][4], aunque también se han utilizado otras geometrías como un array tetraédrico [6]. Los ULA permiten estimar la dirección de llegada (DoA) con precisión en acimut o elevación. Si el array lineal se coloca en el plano horizontal, existe incertidumbre en la estimación de acimut. Si se coloca verticalmente, puede existir incertidumbre en la estimación de la elevación. La incertidumbre puede desaparecer si se plantean restricciones a las DoA, de forma que algunas no estén permitidas.

Como alternativa, en este trabajo evaluamos el uso de un array circular uniforme (UCA) para estimar la elevación y el azimut simultáneamente. La detección 3D se completaría con la estimación del alcance, para lo que sería necesaria la colaboración de al menos dos arrays de sensores acústicos. Este estudio queda fuera del alcance de este trabajo.

El artículo se organiza de la siguiente forma. En la Sección 2 se caracteriza la señal acústica de varios drones, con el fin de conocer cómo se distribuye la energía en el dominio de la frecuencia y determinar la existencia de una banda predominante que pueda ser utilizada para la estimación de DoA con algoritmos de banda estrecha. Además, se estudian técnicas de procesamiento de array para determinar las características del array de micrófonos a

utilizar. En la Sección 3 se describe el algoritmo, basado en conformación de haz, utilizado para llevar a cabo la estimación de los ángulos de acimut y elevación, considerando fenómenos como la reverberación, la atenuación y el desplazamiento Doppler cuando el dron está en movimiento. En la Sección 4 se presentan los resultados, con el valor del error cuadrático medio (RMSE) del ángulo DoA como métrica. Finalmente, la Sección 5 contiene las conclusiones del estudio.

2. CARACTERIZACIÓN DE LA SEÑAL ACÚSTICA

2.1. Características de la señal acústica emitida por un UAV

El sonido producido por los UAV se origina principalmente en los motores y las palas debido a su rotación. Diferentes motores pueden girar a diferente velocidad angular, lo que daría lugar a un sonido compuesto que no es armónico. Además, si hay aceleración, la velocidad de rotación de las palas cambia. La caracterización de las señales de audio generadas por los UAV es necesaria para el correcto diseño de los sistemas de detección y localización. La señal es generalmente no estacionaria si el UAV está en movimiento, produciendo variaciones de la distribución de energía en el dominio de la frecuencia.

En el dominio de la frecuencia, se observan distintos patrones que pueden ser útiles para la detección de vehículos aéreos no tripulados. Por un lado, la energía se concentra en bandas de frecuencia características, asociadas a las velocidades de rotación medias para la sustentación y el movimiento, residiendo la mayor parte de la energía en baja frecuencia. Por otro lado, la envolvente del espectro presenta picos y valles, resultantes de las señales combinadas de los motores y las palas. Estos picos y valles aparecen aproximadamente equiespaciados y las frecuencias a las que se producen tienen una variación lenta en la práctica. Estas frecuencias y la forma del espectro envolvente están relacionadas con el tipo de UAV, proporcionando información valiosa a efectos de detección, clasificación e identificación.

Estableciendo una analogía con el espectro de las señales de voz, debido al comportamiento pseudo-armónico, se ha utilizado el *análisis cepstral* para extraer el valor de la frecuencia más baja con un pico de energía:

$$C(qft) = \mathcal{F}^{-1}\{\log_{10}(\mathcal{F}(s(t)))\} \quad (1)$$

$$C(qft) = \Re \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \log_{10}|S(\Omega)| e^{j\Omega n} d\Omega \right\} \quad (2)$$

donde $\mathcal{F}(s(t))$ representa la transformada de Fourier de la señal $s(t)$, y qft representa la “*quefreny*”, medida en unidades de tiempo. Concretamente, trabajamos con el cepstrum real, que se calcula como la transformada de Fourier inversa del espectro de amplitud. Si la señal es real, el módulo de la transformada de Fourier es real y tiene simetría par, y el

cepstrum real también será real. Sin embargo, en la práctica, el cálculo de las transformadas directa e inversa puede producir un residuo complejo, que debe eliminarse. El *cepstrum* permite identificar la frecuencia con el pico de energía más alto del espectro. En la Figura 1 se presenta un ejemplo.

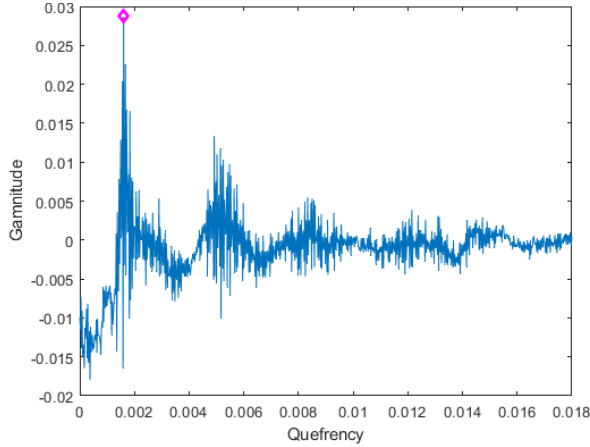


Figura 1. Análisis cepstral de la señal emitida por un cuadricóptero QAV400.

2.2. Arrays de sensores acústicos

La estimación de la dirección de llegada, necesaria para la localización de la fuente sonora, se realiza usando las señales adquiridas con un array de micrófonos con una geometría adecuada. Para mostrar los fundamentos del procesamiento en array, suponemos un array con M micrófonos idénticos. Cuando hay P fuentes sonoras que generan ondas planas en el área bajo observación, el vector $\mathbf{x}^T$ con todas las señales que llegan a los micrófonos se puede expresar con notación matricial de la siguiente forma:

$$\mathbf{x}^T(t) = \mathbf{A}(\theta, \varphi) \mathbf{s}(t) + \mathbf{n}(t), \quad (3)$$

donde $\mathbf{s}^T(t) = [u_1(t), u_2(t), \dots, u_P(t)]$ es el vector con las fuentes sonoras, $\mathbf{n}^T(t) = [n_1(t), n_2(t), \dots, n_M(t)]$ es el vector con las contribuciones de ruido en los micrófonos, la matriz $\mathbf{A}(\theta, \varphi) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_P)]$ es la *matriz de direcciones de llegada* (DOA) con vectores $\mathbf{a}(\theta_i)$ definidos de la siguiente forma:

$$\mathbf{a}^T(\theta_i) = [e^{j\xi \cos(\theta - \theta_1)}, e^{j\xi \cos(\theta - \theta_2)}, \dots, e^{j\xi \cos(\theta - \theta_n)}] \quad (4)$$

siendo $\xi = k r \sin(\varphi)$, $k = \frac{2\pi}{\lambda}$, $\lambda = \frac{c}{f}$, y f la frecuencia. El par (θ, φ) se forma con los valores de acimut y elevación, y r es el radio de la circunferencia si se considera un array circular. Para esta configuración, el factor de array viene dado por la siguiente expresión:

$$AF(\theta, \varphi) = \sum_{n=1}^N \mathbf{a}_n e^{j k r \sin(\varphi) \cos(\theta - \theta_n)} \quad (5)$$

$|AF(\theta, \varphi)|$ define la directividad. El ancho de haz de un array uniforme con N elementos, separados una distancia d , a la frecuencia f (Hz), se calcula con la siguiente expresión:

$$BW = \frac{2c}{Ndf} \quad (6)$$

Para evitar la aparición de réplicas del lóbulo principal, la distancia d entre micrófonos debe ser menor que la mitad de la longitud de onda, a la frecuencia considerada. Se ha seleccionado una distancia entre micrófonos $d=40\text{cm}$, que cumple esta condición para $f < 425\text{Hz}$, que es suficiente, teniendo en cuenta que la potencia de la señal se concentra en baja frecuencia. Para los resultados que se presentan, el número de micrófonos se ha fijado a $N=8$, que da lugar a un ancho de haz $BW=0.5$ rad, para $f=425\text{Hz}$.

3. ESTIMACIÓN DEL ÁNGULO DE LLEGADA

Se han considerado arrays lineales y uniformes. La distancia entre sensores adyacentes es $d=40\text{cm}$, para evitar la aparición de réplicas del lóbulo principal, como se ha mencionado previamente. El array de sensores se posiciona un metro y medio sobre el suelo, de forma que se considerará un efecto de reverberación del 30%.

El algoritmo de localización proporciona una estimación del DOA en acimut y elevación cada medio segundo, lo que supone procesar 4000 muestras, si la frecuencia de muestreo es de 8kHz (se ha seleccionado esta frecuencia de muestreo porque en el dominio de la frecuencia hay información útil hasta 4 kHz como máximo).

En este trabajo utilizamos algoritmos de banda estrecha para estimar la dirección de llegada, aplicados a bandas centradas en múltiplos de la frecuencia estimada con el procedimiento descrito en el subapartado II.A. El algoritmo para estimar el DOA es el siguiente:

- a) Se estima la frecuencia con un máximo de energía (f_{fund}) aplicando el análisis cepstral descrito en el apartado II.A. El algoritmo de estimación del ángulo de llegada (DOA) de banda estrecha se aplica a las bandas centradas en f_{fund} , y los primeros seis múltiplos de la misma, $n \cdot f_{fund}$, $n = 2, \dots, 7$, porque por encima de la banda que abarca a estas frecuencias la energía es relativamente baja. Alrededor de cada múltiplo de f_{fund} , se busca el pico del espectro de amplitud en una ventana de 50 muestras en frecuencia de la transformada discreta de Fourier (DFT), considerando que es posible que el máximo en el espectro de amplitud en esa banda no coincida con el múltiplo de f_{fund} .
- b) Una vez que se han obtenido esos valores, se eligen cinco componentes frecuenciales alrededor de las frecuencias para las que hemos obtenido los picos en el espectro de amplitud, dando lugar a un total de 35 componentes frecuenciales. Este procedimiento evita el barrido de frecuencias, que requeriría un mayor tiempo de procesamiento, y minimiza el impacto de otras fuentes sonoras presentes en la escena acústica, que pueden tener

una elevada potencia en bandas de frecuencia no utilizadas.

- c) A efectos de evaluación, se ha construido una malla de ángulos de acimut y elevación, con una separación de 2° en ambas dimensiones.
- d) Para estimar la posición del UAV, se calcula la respuesta del formador de haz MVDR en las frecuencias de interés para cada punto de la rejilla. El ángulo estimado corresponde al formado entre el punto de la rejilla con máxima respuesta y el centro del array, con el eje correspondiente, promediando seis estimaciones consecutivas.

4. EXPERIMENTOS Y RESULTADOS

Esta sección describe los experimentos realizados y los principales resultados obtenidos, mediante simulaciones, pero en un escenario realista. El experimento realizado para generar las señales se describe en los siguientes pasos:

- a) El array de micrófonos se considera centrado en el origen de coordenadas. El estudio trata de determinar la posibilidad de detectar UAVs en movimiento y estimar la dirección de llegada en acimut y elevación
- b) Se utilizan grabaciones de señales de audio reales, obtenidas con diferentes tipos de UAVs, todos ellos en reposo [2]. Para los resultados presentados en este trabajo, las señales generadas por los UAVs se han escalado para simular un nivel de presión sonora (SPL) medio de 80dB.
- c) El fenómeno de atenuación se ha caracterizado con las normas ISO 9613-1 [8] y 9613-2 [9], de forma que la atenuación de la señal recibida varía en función de la frecuencia y de la distancia fuente-receptor (Figura 2). Para calcular esta atenuación, se ha elegido un paso de frecuencia de aproximadamente 3,9 Hz. La atenuación resultante se ha aplicado en el dominio de la frecuencia a tramas de 1000 muestras.

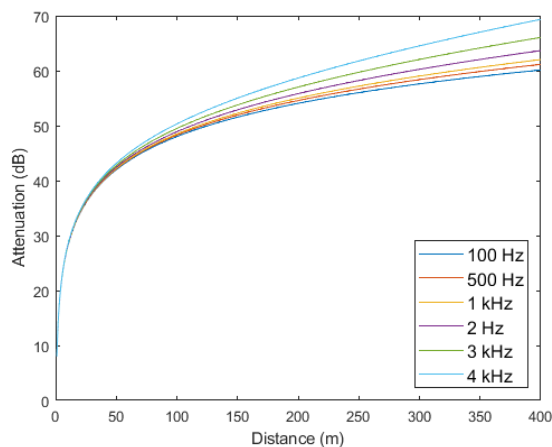


Figura 2. Atenuación dependiente de la frecuencia, en función de la distancia.

- d) El movimiento del UAV se modela generando desplazamientos Doppler, según la dirección del movimiento. El movimiento descrito por el UAV es lineal y paralelo al suelo, en una maniobra que consiste en acercarse y luego alejarse, de forma que la elevación es máxima cuando el ángulo acimutal es mínimo. La posición inicial del UAV viene definida por el vector $\mathbf{r}_0 = [20, y_0, 10]^T$ (m), donde y_0 es variable, dependiendo del experimento a realizar. El vector velocidad viene definido por $\mathbf{v} = [0, v_y, 0]^T$ (m/s). El vector posición en cada instante de tiempo viene determinado por la ecuación del movimiento rectilíneo uniforme: $\mathbf{r} = \mathbf{r}_0 + \mathbf{v} \cdot \mathbf{t}$. La representación de las variaciones de los ángulos de acimut, elevación y distancia, para la maniobra de movimiento descrita, se muestra en la Figura 3.

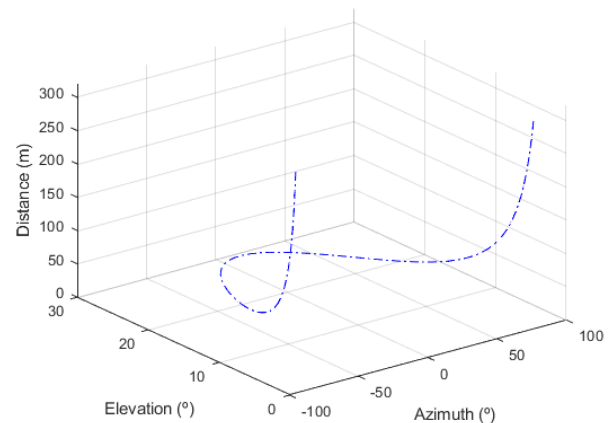


Figura 3. Movimiento simulado del UAV: Acimut, elevación y distancia al centro del conjunto de micrófonos.

- e) En el entorno se encuentra una fuente ruidosa que genera ruido de banda ancha (ruido blanco) con un nivel de 40 dB SPL en la posición de los micrófonos. La señal de ruido generada interfiere con la señal de interés, y la relación señal/ruido (SNR) es inferior a 0dB cuando la distancia de la fuente a los sensores es superior a 50m, aproximadamente.

A modo de ejemplo, se muestran los valores del error cuadrático medio (RMSE) en la estimación de los ángulos de acimut y elevación, para el cuadricóptero AirRobot AR-100, cuando se mueve a una velocidad de 50 m/s (los resultados obtenidos con otros UAVs son similares). En la Figura 4 se observa que los errores obtenidos en la estimación de DOA en acimut y elevación son mayores a medida que aumenta la distancia, debido principalmente al efecto de la atenuación sufrida por la señal. No obstante, se observa que la precisión en la estimación apenas varía hasta que la distancia entre el UAV y el array de micrófonos es superior a 300m, distancia a la cual la SNR es negativa.

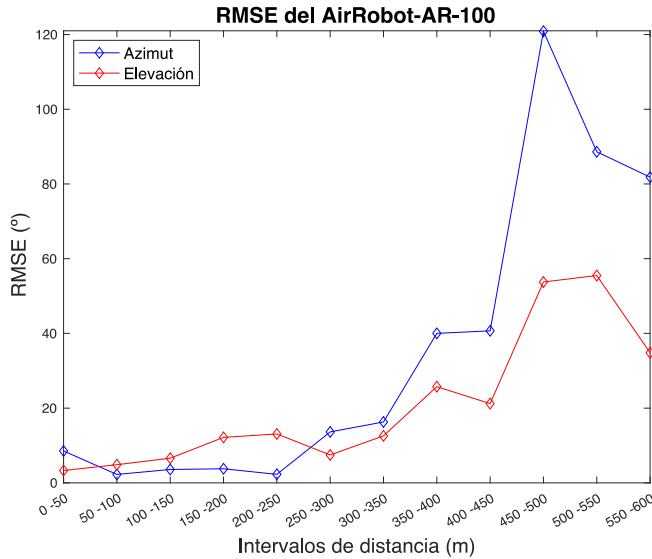


Figura 4. Valor cuadrático medio del error en la estimación de los ángulos de acimut y elevación, para un drone AiRobot-AR-100, en un entorno con ruido de 40dB SPL, con una array circular.

También se presentan los resultados obtenidos con un array lineal (Figura 5). Se observa que los resultados son comparables a los obtenidos con arrays circulares. Debe considerarse que este entorno es incluso desfavorable, puesto que aún con la presencia de ruido de fondo, la conformación de haz necesaria para el apuntamiento del array en las distintas direcciones permitirá atenuar la potencia de ruido mediante filtrado espacial, haciendo la localización más fiable.

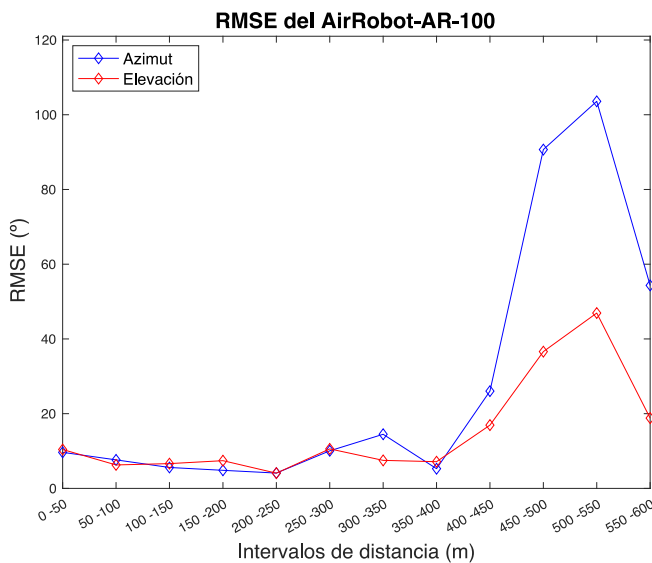


Figura 5. Valor cuadrático medio del error en la estimación de los ángulos de acimut y elevación, para un drone AiRobot-AR-100, en un entorno con ruido de 40dB SPL, con una array lineal.

5. CONCLUSIONES

Este trabajo se ha presentado un estudio para evaluar la posibilidad de localizar la posición de UAVs usando arrays de micrófonos. Se han estudiado arrays lineales y circulares, tratando de evaluar la precisión en la estimación de la dirección de llegada de la onda sonora. Para el estudio, se ha estudiado la señal emitida por distintos tipos de UAVs, demostrándose que la potencia se emite fundamentalmente en baja frecuencia, y que si bien no tiene características pseudo-periódicas por estar formada por la suma de señales generadas por rotores que pueden girar a frecuencias diferentes, sí que observamos una estructura que es comparable al espectro de la señal de voz, con picos y valles. Por esta razón, se ha adoptado el análisis homomórfico (cepstrum), para estimar la frecuencia a la que obtenemos un pico de energía en la frecuencia. Esta frecuencia se ha tomado referencia para el diseño del array.

Para probar las diferentes geometrías con las mismas señales, se ha creado un modelo de señal de las señales emitidas por UAVs en movimiento, en trayectorias controladas. El modelo de señal usado también considera la atenuación del sonido con la distancia y en función de la frecuencia.

Se han comparado arrays lineales y circulares. Se ha comprobado que ambas geometrías, en principio, ofrecen resultados similares, demostrándose que la localización es posible a distancias de hasta 400 metros entre el UAV y el array de micrófonos. Esta distancia no es muy grande, pero permitirá la aplicación de estos sensores como complemento de otras tecnologías, para evitar los inconvenientes de las mismas.

Los resultados obtenidos abren nuevas líneas de investigación. Por un lado, debe considerarse el diseño de nuevas geometrías de arrays de micrófonos, como arrays coprimos. El estudio, además, debe ser validado en un ambiente real, usando las señales reales emitidas por UAVs. Finalmente, el estudio debe completarse con un estudio de la capacidad de filtrar espacialmente los sonidos interferentes y de la posibilidad de usar micrófonos directivos y de alta sensibilidad.

En cualquier caso, el análisis presentado permite aventurar la viabilidad de esta tecnología, siempre como solución complementaria para evitar problemas de otros sensores, como la oclusión por obstáculos, o la imposibilidad de detección en reposo.

10. AGRADECIMIENTOS

Este trabajo ha sido financiado por el Ministerio de Ciencia e Innovación, a través del proyecto PID2021-129043OB-I00 (MSI/FEDER), por la Unión Europea con el proyecto NextGenerationEU/PTR, y por la Junta de Comunidades de Castilla La Mancha, a través del proyecto SBPLY/19/180501/000350.

12. REFERENCIAS

- [1] F. Nex, C. Armenakis, M. Cramer, D.A. Cucci, M. Gerke, E. Honkavaara, A. Kukko, C. Persello, J. Skaloud, "UAV in the advent of the twenties: Where we stand and what is next", *ISPRS journal of photogrammetry and remote sensing*, Vol. 184, pp. 215-242, 2022.
- [2] E. Vattapparamban, I. Güvenç, A.I. Yurekli, K. Akkaya, and S. Uluagaç, "Drones for Smart Cities: Issues in Cybersecurity, Privacy, and Public Safety;" in 2016 International Wireless Communications and Mobile Computing Conference (IWCMC), Paphos, 2016, pp. 216-221.
- [3] M.A. Khan, H. Menouar, A. Eldeeb, A. Abu-Dayya, F.D. Salim, "On the detection of unauthorized drones-techniques and future perspectives: A review". *IEEE Sensors Journal*, Vol. 22, No. 12, pp. 11439-11455, 2022.
- [4] D. Salvati, C. Drioli and G. L. Foresti, "A weighted MVDR beamformer based on SVM learning for sound source localization," *Pattern Recognition Letters*, vol. 84, pp. 15-21, December 2016.
- [5] C. Yang, Z. Wu, X. Chang, X. Shi, J. Wo and Z. Shi, "DOA Estimation Using Amateur Drones Harmonic Acoustic Signals," 2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM), Sheffield, UK, 2018, pp. 587-591.
- [6] X. Chang, C. Yang, X. Shi, P. Li, Z. Shi and J. Chen, "Feature Extracted DOA Estimation Algorithm Using Acoustic Array for Drone Surveillance," 2018 IEEE 87th Vehicular Technology Conference (VTC Spring), Porto, Portugal, 2018, pp. 1-5.
- [7] G. F Cardoso and R. Pater, "A Study into 21st Century Drone Acoustics," [Online], Available in: <http://www.droneacoustics.org/>
- [8] "Acoustics — Attenuation of sound during propagation outdoors — Part 1: Calculation of the absorption of sound by the atmosphere," ISO 9613-1:1993.
- [9] "Acoustics — Attenuation of sound during propagation outdoors — Part 2: General method of calculation," ISO 9613-2:1996.