

Universidad de Alcalá

Escuela Politécnica Superior

Grado en Ingeniería en Sistemas de Telecomunicación



Trabajo Fin de Grado

Análisis de señales acústicas para la
detección de UAVs

Autor: Diana Tejera Berengué

Tutor: Manuel Rosa Zurera

2021

UNIVERSIDAD DE ALCALÁ

Escuela Politécnica Superior

**GRADO EN INGENIERÍA EN SISTEMAS DE
TELECOMUNICACIÓN**

Trabajo Fin de Grado

Análisis de señales acústicas para la detección de UAVs

Autor: Diana Tejera Berengué

Tutor: Manuel Rosa Zurera

TRIBUNAL:

Presidente: María Pilar Jarabo Amores

Vocal 1ª: Manuel Utrilla Manso

Vocal 2ª: Manuel Rosa Zurera

FECHA: 21 de septiembre de 2021

RESUMEN

En este trabajo se presenta el diseño de un sistema para la localización de pequeños UAVs (drones) usando la señal acústica que emiten los mismos en vuelo. Para ello, se realiza un estudio de la fuente sonora, donde, tras analizar la señal en el dominio del tiempo, frecuencia y tiempo-frecuencia, se calcula la frecuencia fundamental del dron mediante el análisis cepstral. Posteriormente, se caracteriza el canal de propagación, es decir, se calcula la posible atenuación en un ambiente real. Por último, se realiza la detección acústica del dron mediante arrays de micrófonos y algoritmos de estimación de la dirección de llegada (DOA). El desarrollo e implementación de los algoritmos se ha llevado a cabo mediante *MATLAB*.

PALABRAS CLAVE:

UAV, frecuencia fundamental, atenuación, array de micrófonos, DOA

ABSTRACT

This work presents the design of a system for locating small UAVs (drones) using the acoustic signal they emit when flying. To do this, a study of the sound source is carried out, where, after analysing the signal in the time, frequency and time-frequency domains, the fundamental frequency of the drone is calculated using cepstral analysis. Subsequently, the propagation channel is characterized, that is, the possible attenuation in a real environment is calculated. Finally, the acoustic detection of the drone is carried out using microphone arrays and direction of arrival estimation algorithms (DOA). The development and implementation of the algorithms has been carried out using *MATLAB*.

KEYWORDS:

UAV, fundamental frequency, attenuation, microphone array, DOA

RESUMEN EXTENDIDO

Los UAVs de reducido tamaño, los drones, son vehículos no tripulados que gozan de gran popularidad debido a sus múltiples usos. El desarrollo de los drones se debe al requerimiento de artefactos para el uso militar, así surgieron innovaciones como las naves no tripuladas o el control remoto. No obstante, en la actualidad, los drones no solo se usan para fines militares como el salvamento en zonas de difícil acceso, la vigilancia e investigación, conservación de zonas y búsqueda de víctimas, sino también para fines recreativos, factor al que se debe su comercialización. Además, en el futuro, estos dispositivos estarán presentes en el día a día, debido a que pueden utilizarse como puntos de acceso móvil para obtener acceso de banda ancha y pueden realizar gran cantidad de servicios como la entrega de mercancías.

Sin embargo, los UAVs pueden utilizarse para llevar a cabo actos ilícitos, como el reconocimiento no autorizado, la vigilancia de individuos, el contrabando de material ilícito, ataques electrónicos y cinéticos, y han provocado problemas como cierres y demoras de servicios en aeropuertos. Por consiguiente, se hace necesaria la detección de estos dispositivos.

Para llevar a cabo la detección de los drones, se utilizan varias técnicas, aunque todas ellas sufren de ciertos inconvenientes. Por ejemplo, en el caso de utilizar radares convencionales, aunque ofrecen operabilidad en condiciones atmosféricas adversas y no necesitan visión directa entre el dron y el equipo radar, al ser un UAV de tamaño reducido y de valores bajos de RCS, indica que será un objeto de difícil detección. Por otro lado, existen inconvenientes para otras técnicas como: radio frecuencia, ya que no pueden detectar los drones que llevan un GPS integrado, sino solo los que se controlan mediante control remoto; cámaras de infrarrojos, debido a que los motores de los drones tienen bajas emisiones de calor; o, video vigilancia, al no funcionar en condiciones de lluvia, niebla o nieve. No obstante, si se usan diversas técnicas para la detección del UAV, será posible evitar estos inconvenientes.

En este trabajo, se estudia una de estas técnicas, la detección acústica. La detección acústica trata de utilizar la señal sonora producida por el dron para poder obtener su posición espacial. Esta detección es posible gracias a las características de la huella acústica del dron. Para poder realizar esta detección, se van a utilizar los audios de tres drones diferentes: QAV400, Parrot AR y Harop Harpy, obtenidos vía Internet [47]. Además, este estudio se va a realizar en el entorno *MATLAB*, con lo cual, todas las operaciones, pruebas y gráficas serán obtenidas mediante dicho software.

En primer lugar, se estudia la señal sonora en el dominio del tiempo, donde se aprecia que es una señal no estacionaria. Después, se hace necesario el desarrollo de la transformada de Fourier para poder convertir la señal del dominio del tiempo al de la frecuencia. En el dominio frecuencial, se observa que la señal es de banda ancha y rica en armónicos, ya que el sonido que generan los drones se produce por los rotores que originan el movimiento.

Sin embargo, aunque en el análisis frecuencial del UAV se pueda apreciar la estructura de la frecuencia fundamental y los armónicos, se hace necesario el análisis en el dominio tiempo-frecuencial, para observar si el cambio de dirección o velocidad del dron, que resulta en un cambio de movimiento de los rotores del dron, afecta en gran medida al desarrollo espectral. Una vez calculado el espectrograma, se observa que aparecen líneas continuas y periódicas en cualquier instante del audio (excepto en el despegue y aterrizaje), por lo tanto, es posible obtener la frecuencia fundamental del audio, que se corresponderá con la situación de vuelo estable del dron. Para el cálculo de esta frecuencia fundamental o pitch, se compara el uso de la autocorrelación y

el Cepstrum, resultando el cálculo del pitch con la herramienta del análisis cepstral la mejor opción. Así, se obtiene la frecuencia fundamental de cada uno de los drones que se tienen en cuenta en el estudio.

Una vez obtenidos las frecuencias fundamentales y armónicos de los UAVs, se realiza un análisis del entorno de propagación del sonido. Es decir, se calcula la atenuación que va a sufrir la señal sonora del dron en un ambiente real, teniendo en cuenta la atenuación por divergencia geométrica, la absorción atmosférica, el efecto del suelo, las reflexiones de superficies y el apantallamiento por obstáculos. Las ecuaciones y conceptos calculados en esta parte del trabajo se obtienen gracias a la norma ISO 9613.

Finalmente, se realiza un estudio y análisis de la detección sonora, esto es, tanto de los receptores acústicos como de los algoritmos de estimación del ángulo de llegada (DOA). Para llevar a cabo esta parte del estudio, primero, se analizan los arrays de micrófonos de configuración geométrica, y se selecciona el que mejor resultado obtenga para la detección del dron, variando la geometría del array (lineal, circular o rectangular), el número de micrófonos que contendrá dicho array y la distancia entre los sensores correspondientes. Después, se estudian varios algoritmos DOA de banda estrecha correspondientes a la estimación del espectro espacial. Dentro de este ámbito se distinguen dos técnicas: beamforming (formación de haces), donde se estudian los algoritmos del beamformer convencional y MVDR (*Minimum Variance Distorsionless Response*), y técnicas basadas en el subespacio, donde se analiza el algoritmo MUSIC (*MUltiple Signal Classification*). Asimismo, se estudia el uso de un algoritmo de banda ancha, GCC-PHAT (*Generalized Cross Correlation with Phase Transform*).

La señal producida por los rotores de un dron es una señal de banda ancha, por lo que para poder utilizar los algoritmos DOA de alta resolución, se debe realizar una adaptación de la señal, en la que se utiliza las propiedades de la transformada de Fourier, para poder ir seleccionando las subbandas de interés en el rango de frecuencias donde las señales acústicas de los drones tienen el mayor número de componentes frecuenciales.

Por último, se prueban las distintas configuraciones entre los arrays y los algoritmos de forma que se obtiene la solución que obtiene los mejores resultados para la detección de los UAVs.

ÍNDICE GENERAL

Resumen	V
Abstract	VII
Resumen extendido	IX
Índice general	XI
Índice de figuras	XV
Abreviaturas y acrónimos	XIX
1 Introducción	1
1.1 Planteamiento	2
1.2 Objetivos	3
2 Caracterización espectral de los UAVs	5
2.1 Introducción	5
2.2 Análisis en el dominio del tiempo	7
2.3 Autocorrelación.....	8
2.4 Análisis en el dominio de la frecuencia.....	9
2.4.1 Análisis frecuencial de un UAV.....	12
2.5 Espectrograma.....	14
2.5.1 Enventanado	15
A) Longitud de la ventana	15
B) Tipos de ventanas	17
Ventana rectangular.....	17
Ventana de Hanning	18
Ventana Hamming.....	19
Ventana Blackman	19
2.5.2 Espectrograma de un UAV.....	20
2.6 Frecuencia fundamental	23
2.6.1 Comparación entre dos métodos de estimación del pitch (Autocorrelación y Cepstrum).....	23
A) Cepstrum	24
B) Cepstrum de un UAV	26

2.6.2	Armónicos de la frecuencia fundamental	29
3	Propagación del sonido en ambientes reales	33
3.1	La atenuación del sonido: ISO 9613	33
3.1.1	Divergencia geométrica <i>Adiv</i>	35
3.1.2	Absorción atmosférica <i>Aatm</i>	36
A)	Frecuencia del sonido	36
B)	Temperatura atmosférica ambiente	36
C)	Humedad relativa del aire	36
D)	Presión atmosférica ambiente	38
3.1.3	Efecto del suelo <i>Agr</i>	39
3.1.4	Apantallamiento por efecto de barreras <i>Abar</i>	41
3.1.5	Otros efectos similares <i>Amisc</i>	42
3.2	Análisis medio de propagación para un UAV	43
4	Receptores acústicos	47
4.1	Micrófono	47
4.1.1	Elementos del receptor acústico	47
4.1.2	Características del receptor acústico	48
A)	Sensibilidad	48
B)	Respuesta en frecuencia	48
C)	Impedancia de salida	48
D)	Ruido eléctrico y distorsión	48
E)	Patrón de directividad	48
	Azimut y elevación	49
4.2	Arrays de micrófonos	50
4.2.1	Geometría de los arrays	50
4.2.2	Características del array	52
5	Detección de ángulos de llegada	55
5.1	Señales de banda ancha y estrecha	55
5.2	Algoritmos de estimación del ángulo de llegada (DOA)	59
5.2.1	Algoritmos de banda estrecha	60

A)	Estimación del espectro espacial.....	60
I)	<i>BEAMFORMING</i>	61
	Bartlett beamforming (Beamsan)	62
	MVDR.....	63
II)	<i>TÉCNICAS BASADAS EN EL SUBESPACIO</i>	64
	MUSIC	64
5.2.2	Algoritmos de banda ancha	65
	<i>GCC-PHAT</i>	65
6	Análisis y resultados.....	67
6.1	Elección del array para cada DOA	67
6.2	Selección configuración mediante barrido de ángulos	70
6.3	Selección array y DOA en un ambiente real	73
7	Conclusión y líneas futuras.....	87
7.1	Conclusiones	87
7.2	Líneas futuras	88
Anexos	91
I.	Pliego de condiciones.....	91
a)	Requisitos de hardware y software.....	92
II.	Presupuesto	95
a)	Presupuesto de ejecución material	95
b)	Presupuesto por contrata	97
c)	Presupuesto total del proyecto.....	97
III.	Manual de usuario	99
a)	Señal acústica UAVs	99
b)	<i>MATLAB</i>	99
	Códigos de <i>MATLAB</i>	100
Bibliografía	113

ÍNDICE DE FIGURAS

Figura 1 - Esquema del análisis de señales acústicas para la detección de UAVs.....	3
Figura 2 – Tipos de UAVs en el estudio: a) QAV400; b) Harop Harpy; c) Parrot AR.	5
Figura 3 – Diagrama de un cuadricóptero con dos hélices por motor [3].	6
Figura 4 – Ejes de movimiento de una aeronave común [4].	6
Figura 5 – Representación de las señales de los UAVs en el dominio del tiempo.....	7
Figura 6 – Representación de la señal producida por los tres tipos de UAVs en el dominio del tiempo, en una ventana temporal de 50ms.	7
Figura 7 - Segmento de la señal producido por el modelo QAV400, sin aparente periodicidad. .	8
Figura 8 – Espectros de: a) La señal original y de la función de muestreo; b) La señal muestreando con $w_s > 2w_M$; c) La señal muestreando con $w_s < 2w_M$ (adaptado de [5]).	10
Figura 9 – Transformación de una señal aperiódica y finita a una señal periódica e infinita [5].	10
Figura 10 – Procedimiento para evitar solapamiento temporal en la DFT [6].	12
Figura 11 – Representación de la señal de los UAVs en el dominio de la frecuencia.	13
Figura 12 – Representación frecuencial de la señal de los UAVs en el rango entre 0 y 3 kHz.	13
Figura 13 – Proceso de la STFT [9].	14
Figura 14 – Señales con solapamiento del 50 % [9].	15
Figura 15 – Espectrograma del QAV400 con una ventana de 64 muestras.	16
Figura 16 – Espectrograma del QAV400 con una ventana de 4096 muestras.	16
Figura 17 – Oscilación de la ventana en el dominio frecuencial: a) Ventana de 16 muestras; b) Ventana de 64 muestras.....	17
Figura 18 – Ventana rectangular de 64 muestras en el dominio temporal y frecuencial.	18
Figura 19 – Ventana de Hanning de 64 muestras en el dominio temporal y frecuencial.	18
Figura 20 – Ventana de Hamming de 64 muestras en el dominio temporal y frecuencial.....	19
Figura 21 – Ventana de Blackman de 64 muestras en el dominio temporal y frecuencial.	19
Figura 22 – Espectrograma de los UAVs con vista de $(-45,65)^\circ$	20
Figura 23 – Espectrograma del QAV400.....	21
Figura 24 – Espectrograma del Parrot AR.	21
Figura 25 – Espectrograma del Harop Harpy.....	22
Figura 26 – Análisis cepstral del QAV400.	27
Figura 27 – Análisis cepstral con liftado del QAV400.	28
Figura 28 – Análisis cepstral con liftado del Parrot AR.	29
Figura 29 – Análisis cepstral con liftado del Harop Harpy.....	29
Figura 30 – Algoritmo de estimación de las frecuencias armónicas [18].	30

Figura 31 – Cálculo de los armónicos mediante la FFT de las señales acústicas de los UAVs.	31
Figura 32 – Localización de las distintas zonas en un diagrama psicrométrico [21].	37
Figura 33 – Diagrama Psicrométrico a 20°C y 40% de humedad [22].	37
Figura 34 – Rango de valores del coeficiente de absorción atmosférica [20].	38
Figura 35 – Regiones de atenuación por efecto del suelo [20].	39
Figura 36 – Cálculo de variables para el cálculo del efecto por suelo para determinadas frecuencias [20].	40
Figura 37 – Efecto por apantallamiento [20]: a) Difracción simple; b) Difracción doble.	42
Figura 38 – Representación de atenuación por follaje [20].	42
Figura 39 – Representación de atenuación por sitios industriales [20].	43
Figura 40 – Atenuación total en función de la frecuencia y distancia de los UAVs.	44
Figura 41 – Ampliación de la atenuación total del QAV400.	45
Figura 42 – Diagrama de radiación de una antena.	49
Figura 43 – Diagrama de azimut y elevación [24].	49
Figura 44 – Arrays de micrófonos con geometría uniforme y : a) Lineal; b) Circular; c) Rectangular.	50
Figura 45 – Patrón de directividad en azimut de un array con geometría: a) ULA; b) UCA; c) URA.	51
Figura 46 – Patrón de directividad en elevación de un array con geometría: a) UCA; b) URA	51
Figura 47 – Patrón de directividad de un array ULA al variar el número de micrófonos, y, con una distancia entre sensores de 0,3 metros.	52
Figura 48 – Patrón de directividad de un array ULA al variar el número de distancia entre micrófonos, y, con un número de 8 sensores	53
Figura 49 – Periodograma señal de: a) Banda estrecha; b) Banda ancha.	55
Figura 50 – Diagrama de convolución para cada valor de k_0 (banda de frecuencia).	56
Figura 51 – Representación del valor real central y las parejas de complejos conjugados.	57
Figura 52 – Valores para formar las subbandas de la señal de entrada.	59
Figura 53 – Comparación entre beamformer convencional y MVDR en la estimación de dos fuentes próximas [32].	64
Figura 54 – Error medio absoluto en azimut del algoritmo Beamscan con un array UCA, donde cada fila corresponde al número de micrófonos y cada columna a la distancia entre ellos.	70
Figura 55 – Selección de configuraciones para cada geometría y algoritmo DOA.	70
Figura 56 – Error medio absoluto en función de un barrido de ángulos azimut para elevaciones de 0, 20, 40, 60 y 80 grados. (Izquierda: Error en azimut; Derecha: Error en elevación).	71
Figura 57 – Representación del error medio absoluto de azimut en función de la distancia entre el dron y el array para distintos ángulos en azimut y a una elevación 20 grados.	75
Figura 58 – Representación del error medio absoluto de elevación en función de la distancia entre el dron y el array para distintos ángulos en azimut y a una elevación 20 grados.	76
Figura 59 – Representación del error medio absoluto de azimut en función de la distancia entre el dron y el array para distintos ángulos en azimut y a una elevación 40 grados.	77

Figura 60 – Representación del error medio absoluto de elevación en función de la distancia entre el dron y el array para distintos ángulos en azimut y a una elevación 40 grados.....	78
Figura 61 – Representación del error medio absoluto de azimut en función de la distancia entre el dron y el array para distintos ángulos en azimut y a una elevación 60 grados.....	79
Figura 62 – Representación del error medio absoluto de elevación en función de la distancia entre el dron y el array para distintos ángulos en azimut y a una elevación 60 grados.....	80
Figura 63 – Representación del error medio absoluto del algoritmo MVDR con UCA en distintas configuraciones en función de la distancia entre el array y el UAV para distintos valores en azimut y a una elevación de 20 grados.	82
Figura 64 – Representación del error medio absoluto del algoritmo MVDR con UCA en distintas configuraciones en función de la distancia entre el array y el UAV para distintos valores en azimut y a una elevación de 40 grados.	84
Figura 65 – Representación del error medio absoluto del algoritmo MVDR con UCA en configuración de 8 micrófonos separados 0,6 metros en función de la distancia entre el array y el UAV para distintos valores en azimut y a una elevación de 60 grados.....	86
Figura 66 – Representación error del MVDR con UCA de 8 micrófonos y distancia de 0,6 metros en función de la distancia entre el array y el UAV para los drones: a) Parrot AR; b) Harop Harpy.....	86
Figura 67 – Tabla de requisitos para la ejecución de MATLAB.	93
Figura 68 – Tabla de costes del inmovilizado material.....	96
Figura 69 – Tabla de costes del inmovilizado intangible.....	96
Figura 70 – Tabla de costes de suministros.....	96
Figura 71 – Coste total del material	97
Figura 72 – Tabla de costes de la mano de obra.	97
Figura 73 – Presupuesto de ejecución material.....	97
Figura 74 – Presupuesto de contrata del proyecto.....	97
Figura 75 – Presupuesto total del proyecto.	98

ABREVIATURAS Y ACRÓNIMOS

<i>2D</i>	Dos dimensiones
<i>3D</i>	Tres dimensiones
<i>A</i>	Atenuación
<i>AWGN</i>	Ruido Blanco Gaussiano Aditivo (<i>Additive White Gaussian Noise</i>)
<i>B</i>	Ancho de banda (<i>Beamwidth</i>)
<i>C</i>	Centígrado
<i>c</i>	Velocidad del sonido (343 metros por segundo)
<i>d</i>	Distancia
<i>dB</i>	Decibelios
<i>DFT</i>	Transformada Discreta de Fourier (<i>Discrete Fourier Transform</i>)
<i>DIT</i>	Diezmado en tiempo (<i>Decimation-In-Time</i>)
<i>DOA</i>	Detección de Llegada (<i>Direction of Arrival</i>)
<i>DTFT</i>	Transformada de Fourier de Tiempo Discreto (<i>Discrete-time Fourier transform</i>)
<i>f</i>	Frecuencia
<i>FFT</i>	Transformada Rápida de Fourier (<i>Fast Fourier Transform</i>)
<i>FNBW</i>	Ancho de haz entre nulos (<i>First Null Beam Width</i>)
<i>g</i>	Gramos
<i>GCC-PHAT</i>	Correlación Cruzada Generalizada con transformada de fase (<i>Generalized Cross Correlation with Phase Transform</i>)
<i>GPS</i>	Sistema de Posicionamiento Global (<i>Global Positioning System</i>)
<i>h</i>	Altura
<i>HERM</i>	Modulación del rotor del helicóptero (<i>Helicopter Rotor Modulation</i>)
<i>HPBW</i>	Ancho de haz a 3 dB (<i>Half-Power Beam Width</i>)
<i>Hz</i>	Hercios (Hertz)
<i>IEEE</i>	Instituto de Ingenieros Eléctricos y Electrónicos (<i>Institute of Electrical and Electronic Engineers</i>)
<i>ISO</i>	Organización Internacional de Normalización (<i>International Organization for Standardization</i>)
<i>K</i>	Kelvin

<i>k</i>	Kilo (10^3)
<i>m</i>	Metros
<i>ML</i>	Máxima verosimilitud (<i>Maximum Likelihood</i>)
<i>MUSIC</i>	Clasificación de señal múltiple (<i>MUltiple SIgnal Classification</i>)
<i>MVDR</i>	Conformador de haz de respuesta de mínima varianza sin distorsión (<i>Minimum Variance Distorsionless Response</i>)
<i>n</i>	Tiempo discreto
<i>Pa</i>	Pascales
<i>pW</i>	Pico watio (10^{-12} watios)
<i>qft</i>	Cuefrecia (<i>quefreny</i>)
<i>RCS</i>	Sección equivalente de radar (<i>Radar Cross-Section</i>)
<i>s</i>	Segundos
<i>SNR</i>	Relación señal a ruido (<i>Signal to Noise Ratio</i>)
<i>STFT</i>	Transformada de Fourier de Tiempo Reducido (<i>Short-term Fourier Transform</i>)
<i>t</i>	Tiempo
<i>TDOA</i>	Diferencia en tiempo de llegada (<i>Time Difference of Arrival</i>)
<i>TF</i>	Transformada de Fourier
<i>UAV</i>	Vehículo aéreo no tripulado (<i>Unmanned Aerial Vehicle</i>)
<i>UCA</i>	Array circular uniforme (<i>Uniform Circular Array</i>)
<i>ULA</i>	Array lineal uniforme (<i>Uniform Lineal Array</i>)
<i>URA</i>	Array rectangular uniforme (<i>Uniform Rectangular Array</i>)
<i>W</i>	Vatios (Watt)
<i>XLR</i>	Línea de Retorno Externa (<i>eXternal Lineal Return</i>)

CAPÍTULO 1

1 INTRODUCCIÓN

Los UAVs de reducido tamaño tienen gran popularidad hoy en día, ofreciendo servicios como la inspección, investigación, vigilancia, además de su gran uso recreativo. Sin embargo, los drones también pueden ser utilizados para cometer actos ilícitos, por lo que se hace necesaria su detección, para poder implementar contramedidas que neutralicen sus actos.

Este trabajo pretende proporcionar un método basado en la detección acústica de los UAVs para tener conocimiento de su localización y evitar así su uso ilícito. La detección acústica podría combinarse con otro tipo de técnica de detección, como el radar, para poder obtener los mejores resultados posibles y minimizar los posibles inconvenientes.

La investigación de la detección acústica se realiza debido a que la señal producida por un dron está formada por una gran cantidad de componentes en el dominio frecuencial. Estos componentes se generan por las hélices en movimiento del motor del dron. Los rotores se mueven a una u otra velocidad dependiendo del movimiento o dirección que se quiera conseguir. Sin embargo, aunque se requiera que el dron este en reposo en el aire, las hélices siguen girando debido a que el mínimo soplo de viento variaría la posición del dron. Cualquier cambio en el movimiento de los rotores conlleva una variación de la energía sonora emitida, que podría caracterizarse a partir del análisis en el dominio temporal, mediante el cálculo de ciertas características.

Después, se hace necesario pasar al dominio de la frecuencia y calcular las componentes frecuenciales fundamentales. Un sonido es una señal compleja que se puede separar en un sumatorio de señales simples. El oído humano, por medio del caracol, descompone las señales sonoras que le llegan en sus componentes, y ésta es la información básica que se transmite por el nervio auditivo al cerebro. Por tanto, se puede afirmar que el proceso de audición se fundamenta en la descomposición en frecuencias de la señal sonora [1]. Los receptores acústicos tienen este mismo funcionamiento.

Sin embargo, como el movimiento de las hélices del dron va cambiando, dependiendo del viento o del movimiento, la señal en el dominio del tiempo no es estacionaria, con lo cual, se debe transformar a un dominio tiempo-frecuencia donde se puedan observar las características espectrales en función del tiempo. En este nuevo dominio aparecen líneas paralelas correspondientes a los armónicos de la señal sonora que son en su mayoría constantes con lo cual es posible obtener la frecuencia fundamental que caracteriza a cada uno de los drones.

Tras realizar el análisis espectral de la señal emitida por los UAVs, se realizará una caracterización del medio por el cual se propaga el sonido, estimando la atenuación debido a los diversos efectos del entorno en el que está presente el dron.

Por último, se realiza un estudio de los receptores acústicos y los algoritmos de detección que estimarán la posición en la que se encuentra el UAV. Por un lado, los receptores acústicos que se van a seleccionar serán conformaciones de arrays de micrófonos, para así poder estimar la posición del dron en 3D, es decir, dirección, altura y distancia entre la fuente sonora y el receptor.

Por otro lado, los algoritmos de estimación que se van a usar son los caracterizados por la estimación del espectro espacial, y dentro de este ámbito se analizan tanto los algoritmos de beamforming como los basados en el subespacio. Además, se estudia un algoritmo que capta el sonido directamente como una señal de banda ancha y está basado en la correlación cruzada.

Una vez realizado el estudio del escenario (fuente sonora, medio de propagación y recepción acústica), se realiza un análisis para seleccionar el receptor acústico y el algoritmo de estimación del ángulo de llegada que más se ajuste a las características obtenidas de la señal sonora del dron.

Este estudio se ha realizado gracias al software *MATLAB* y a los audios de tres drones obtenidos de la investigación auditiva ‘*A Study into 21st Century Drone Acoustics*’ realizada por el compositor Gonçalo F. Cardoso y el diseñador Ruben Pater obtenida vía Internet [47].

1.1 PLANTEAMIENTO

Este trabajo consiste en el estudio de la señal acústica producida por un UAV, debida a la rotación de las aspas de sus motores, en las características atenuantes que van a estar presentes en el medio en el que se propaga dicha señal sonora, y en el diseño de un sistema que contenga arrays de micrófonos y algoritmos de detección de la dirección de llegada para tal fin.

Con el objetivo de abordar estas características se sugiere el siguiente planteamiento.

- CAPÍTULO 1. Se realiza una breve introducción al trabajo desarrollado, formulando los objetivos a conseguir.
- CAPÍTULO 2. Se realiza la caracterización espectral de los UAVs. Para ello, en primer lugar, se introduce la señal sonora y se aborda un análisis en el dominio del tiempo, en el de la frecuencia y en tiempo tiempo-frecuencia. El capítulo termina calculando la frecuencia fundamental de los drones con el método cepstral, así como los armónicos de la señal.
- CAPÍTULO 3. Se estudia el medio de propagación del sonido como un ambiente real. Para ello, se calculan las atenuaciones y los efectos adversos en función de la frecuencia y la distancia entre la fuente sonora y el receptor acústico.
- CAPÍTULO 4. Se analizan los elementos y características de los micrófonos y los arrays de micrófonos. Asimismo, se abordan estructuras geométricas de estos arrays.
- CAPÍTULO 5. Este capítulo trata de explicar los algoritmos utilizados para la estimación del ángulo de llegada. Muchos algoritmos clásicos han sido propuestos para la detección de fuentes de radiofrecuencia o blancos radar, donde las señales son de banda estrecha en relación con la frecuencia central. Para poder usar estos algoritmos, se necesita convertir la señal de banda ancha en un conjunto de señales de banda estrecha, gracias a las propiedades de la transformada de Fourier y la convolución.
- CAPÍTULO 6. Se realiza un diseño con los resultados obtenidos en los capítulos anteriores y se prueba la mejor configuración del array de micrófonos con el algoritmo de detección para poder realizar la estimación de la posición del UAV.
- CAPÍTULO 7. Trata las conclusiones del estudio y las posibles líneas futuras que se puedan llevar a cabo.

- **CAPÍTULO 8.** Corresponde a distintos anexos correspondientes a este trabajo como el pliego de condiciones, el presupuesto y el manual de usuario.

El planteamiento que se lleva a cabo en el trabajo se muestra en el siguiente esquema, donde los resultados finales de cada una de las ramas son los que se utilizan para llevar a cabo el diseño del sistema.

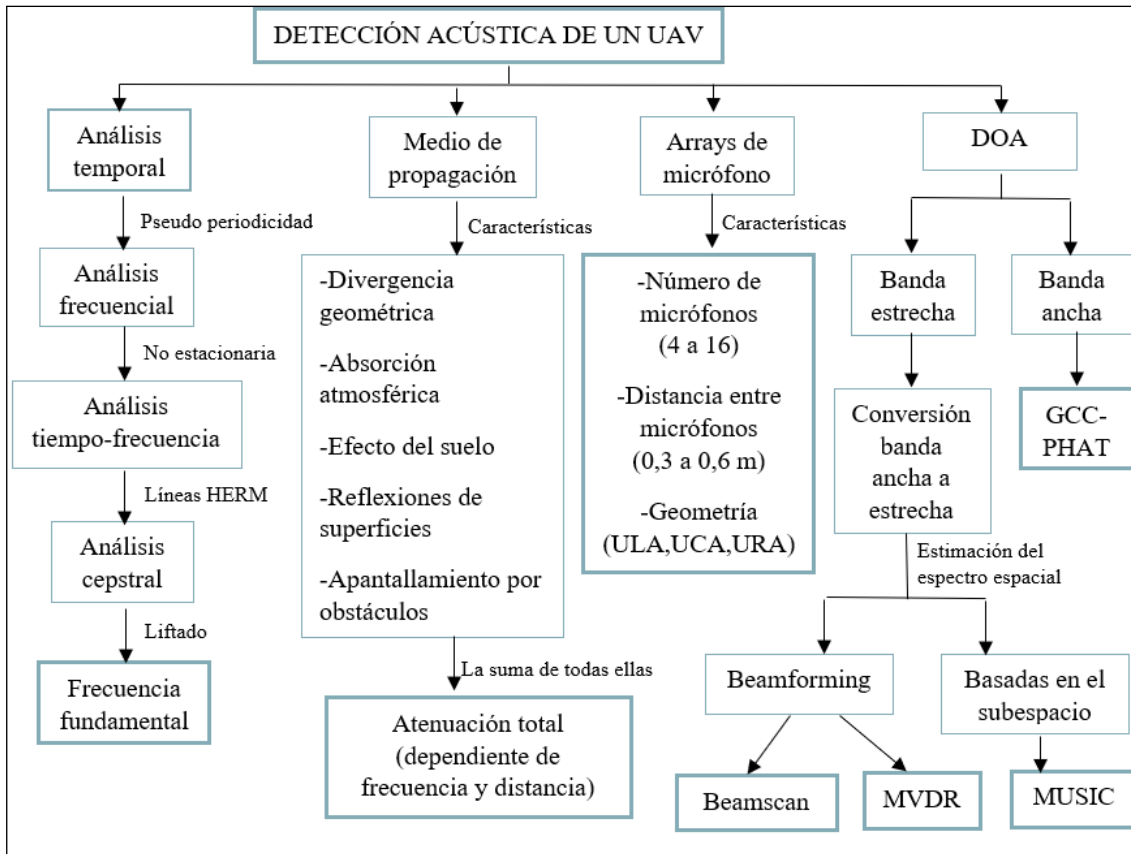


Figura 1 - Esquema del análisis de señales acústicas para la detección de UAVs.

1.2 OBJETIVOS

En este trabajo se propone el estudio de la posibilidad de detectar UAVs usando el sonido que generan debido a la rotación de las aspas, y el diseño de un sistema basado en arrays de micrófonos para tal fin. Se construye, además, una prueba de concepto, que permite validar las principales conclusiones del estudio.

Para ello, se establecen los siguientes objetivos:

- 1) Caracterización espectral de los sonidos generados por drones.
- 2) Estudio de la propagación del sonido generado por drones en ambientes reales y selección de bandas de frecuencia más favorables para la detección.
- 3) Diseño de un array de micrófonos para la localización del dron a partir de la señal en la banda seleccionada.
- 4) Implementación de un pequeño demostrador que muestre la prueba de concepto.

CAPÍTULO 2

2 CARACTERIZACIÓN ESPECTRAL DE LOS UAVS

2.1 INTRODUCCIÓN

Un dron según el diccionario de Oxford es un pequeño vehículo aéreo no tripulado, utilizado en el ámbito militar (para reconocimiento táctico desde gran altura, vigilancia del campo de batalla o guerra electrónica) y civil (vigilancia de manifestaciones, control de la contaminación y de incendios forestales, etc.) [2].

Para poder llevar a cabo la caracterización espectral del sonido producido por los drones, en primer lugar, se describen ciertas características comunes a cualquier sonido.

El sonido se genera por variaciones en la presión del medio por el que se propaga, en este caso, el aire. Estas variaciones de presión se deben a la diferencia entre la presión atmosférica y la presión instantánea generada por la señal sonora. Las variaciones de presión se convierten en ondas mecánicas que producen un estímulo físico, la sensación auditiva. Para producir dicho estímulo, las ondas mecánicas pasan a ser ondas eléctricas en el receptor acústico, ya sea el oído o un micrófono. Las ondas que genera una fuente sonora son esféricas y se propagan longitudinalmente, es decir, se comportan de forma que varían estados de mayor o menor compresión.

En la ecuación (1), se calcula el nivel de presión sonora, donde p es la presión instantánea y p_{ref} es la presión de referencia o presión ambiente media.

$$SPL = 20 \log_{10} \left(\frac{p}{p_{ref}} \right) \text{ dB} \quad (1)$$

Por otro lado, al recibir la señal sonora, dependiendo de sus cualidades, se percibirá en el receptor con distinta intensidad (volumen), duración (tiempo), altura (agudo o grave) y timbre (número de armónicos y su intensidad).

En cuanto al sonido generado por un UAV, existen distintos tipos de drones. En este trabajo, en concreto, se analizan tres modelos: QAV400, Harop Harpy y Parrot AR, obtenidos en la investigación auditiva ‘A Study into 21st Century Drone Acoustics’ [47].

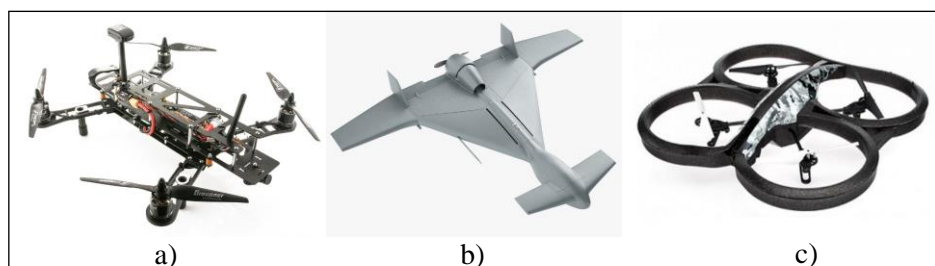


Figura 2 – Tipos de UAVs en el estudio: a) QAV400; b) Harop Harpy; c) Parrot AR.

El sonido generado por un UAV está compuesto por una serie de frecuencias armónicas, debido a que es generado por las hélices de los motores del dron. Si se estudia el principio de operación

de, por ejemplo, el caso de un dron cuadricóptero como QAV400 y Parrot AR, existen dos tipos de configuraciones distintas para controlar sus motores: configuración cruzada, donde los cuatro motores del cuadricóptero controlan la rotación de cabeceo y alabeo; y, configuración plus, donde dos motores controlan la rotación de cabeceo y otros dos la de alabeo.

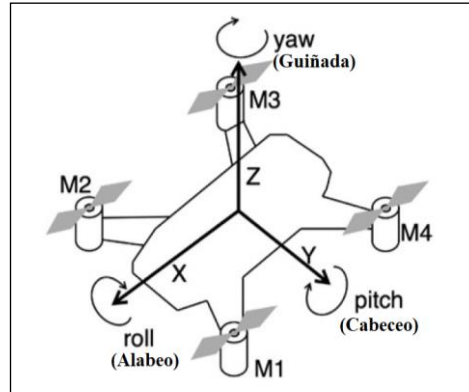


Figura 3 – Diagrama de un cuadricóptero con dos hélices por motor [3].

En las Figuras 3 y 4 se observan los distintos tipos de movimiento de un UAV, estos son:

- **Alabeo.** Es el movimiento respecto al eje longitudinal x del dron.
- **Cabeceo.** Es el giro respecto al eje transversal y del UAV.
- **Guiñada.** Rotación respecto al eje vertical z del dron.

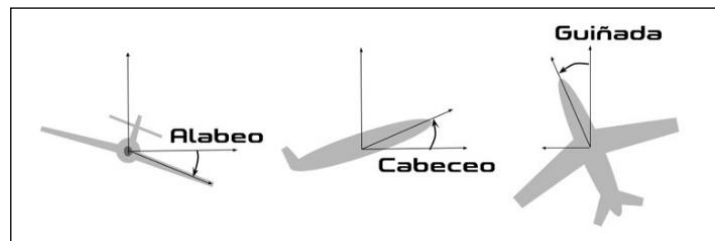


Figura 4 – Ejes de movimiento de una aeronave común [4].

Al analizar el movimiento de un cuadricóptero, se observa que actúan varias fuerzas sobre éste, además de la resistencia del aire. Las fuerzas dominantes son el empuje y la gravitación [3].

Para poder volar, un cuadricóptero tiene que estar constantemente monitorizando y corrigiendo la velocidad de rotación de los motores, ya que incluso una pequeña brisa de viento hará que el dron necesite una corrección. Esto pasa tanto en movimiento hacia delante como hacia detrás y lateral. Incluso hacia arriba, aunque se necesite de una menor corrección. La diferencia de estos movimientos determina la velocidad respecto al suelo. Entonces, si el cuadricóptero se está moviendo, existirán diferencias en las velocidades de rotación de los cuatro motores [3]. Por ejemplo, si se tiene un cuadricóptero con configuración cruzada y se busca el movimiento hacia la izquierda, las velocidades de rotación de los motores M1 y M4 en la Figura 3 serán mayores que las de los motores M2 y M3.

Con el objetivo de detectar un UAV, es importante tener conocimiento del sonido que producen estos dispositivos. El espectro en frecuencia del sonido generado por un dron corresponde con las velocidades de rotación de los rotores. Será un espectro ancho, con las frecuencias fundamentales de cada rotor y sus correspondientes armónicos. Con lo cual, se deberían extraer directamente las características del espectro en frecuencia ya que es un espectro ancho y rico en armónicos.

2.2 ANÁLISIS EN EL DOMINIO DEL TIEMPO

Gracias a los archivos de audio obtenidos de Internet, se puede caracterizar la huella acústica de distintos tipos comerciales de UAVs. En primer lugar, se representan las señales en el dominio del tiempo, es decir, la amplitud que tiene la señal en función del tiempo. Con esta representación se pueden caracterizar parámetros importantes de la señal como la amplitud máxima, la duración y la frecuencia fundamental si fuera una señal periódica.

El sonido producido por un dron es una señal no estacionaria, ya que va cambiando con el tiempo. Sin embargo, se pueden encontrar características de periodicidad en ciertos momentos cuando la velocidad de los rotores del dron se mantiene constante. Esto es debido a que los rotores estarán girando a velocidad constante. Si se considera, por ejemplo, el sonido producido por un UAV QAV400, Harop Harpy y Parrot AR, se obtiene en el dominio del tiempo las siguientes señales:

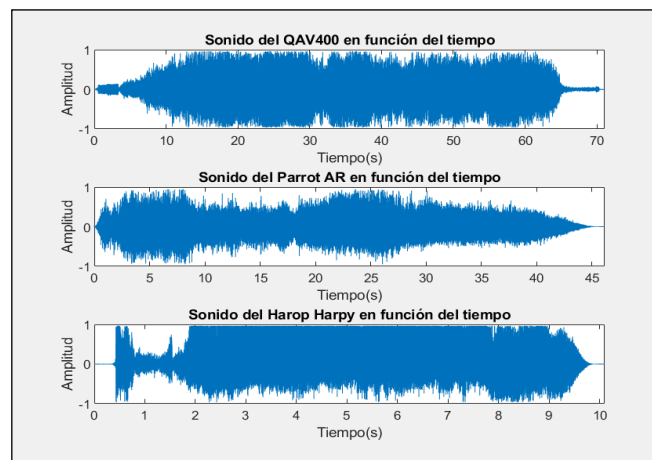


Figura 5 – Representación de las señales de los UAVs en el dominio del tiempo.

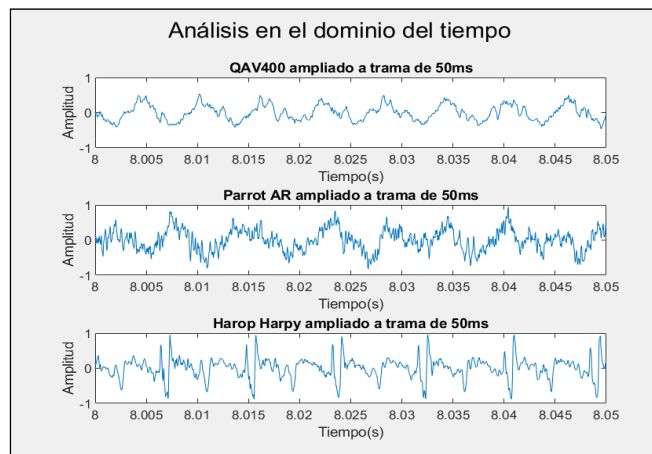


Figura 6 – Representación de la señal producida por los tres tipos de UAVs en el dominio del tiempo, en una ventana temporal de 50ms.

En las imágenes anteriores, se puede ver como cuando se hace el análisis de la señal completa, la señal sonora queda lejos de ser periódica, mientras que, si se analizan franjas de tiempo con una ventana de, por ejemplo, 50 ms, se aprecian indicios de pseudo-periodicidad. Es decir, en lugar de tener repeticiones iguales de la señal a lo largo del tiempo, al analizar un instante en el que la velocidad de rotación de las hélices es constante, se obtiene una réplica similar de un segmento de la señal, pero desplazada un cierto tiempo.

Sin embargo, si se representa un segmento de la señal de cualquiera de los sonidos de los drones anteriores en instantes en los que el dron esté cambiando de velocidad, ya sea por movimiento o por un despegue o aterrizaje, se observará una señal donde para la componente debida a un rotor el periodo tendrá una variación según algún tipo de ley (lineal, cuadrática, o de otro tipo). Además, este tipo de movimientos se produce haciendo girar a los rotores a distintas velocidades, siendo el sonido generado la superposición de dos señales periódicas, cuyo resultado no tiene porqué ser periódico si no existe relación armónica entre las frecuencias de las señales que se superponen. Un ejemplo de este tipo de señal es el representado en la Figura 7, que corresponde al modelo QAV400, donde se puede comprobar la pérdida de cualquier tipo de pseudo-periodicidad:

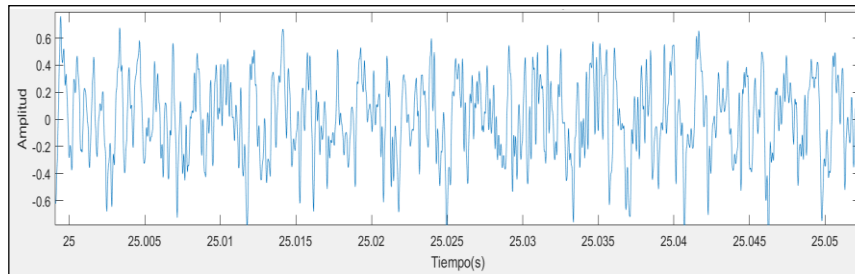


Figura 7- Segmento de la señal producido por el modelo QAV400, sin aparente periodicidad.

Debido a que la señal cambia a lo largo del tiempo (es una señal no estacionaria), no se podrá identificar un pitch o frecuencia fundamental en dicho dominio, por lo que se necesita pasar al dominio de la frecuencia para poder detectar componentes frecuenciales dominantes, aunque no estén relacionadas armónicamente, al ser una señal compleja.

2.3 AUTOCORRELACIÓN

La autocorrelación, en el procesado de señales, se puede usar para encontrar patrones de periodicidad en una señal. La autocorrelación trata de calcular el grado de correlación o semejanza entre una señal y ella misma desplazada un cierto intervalo de tiempo.

$$R_x(\tau) = \int_{-\infty}^{\infty} x^*(t) x(t + \tau) dt \quad (2)$$

En el dominio del tiempo, se ha visto que dependiendo de en qué instante de tiempo se localice la ventana de análisis, se obtiene un resultado pseudo-periódico o aperiódico. Es por ello por lo que se usa la autocorrelación de la señal completa, ya que ésta informa del comportamiento periódico de la señal, de forma que tiene en menor consideración el ruido de fondo. Esto se debe a que el ruido blanco gaussiano tiene una distribución aleatoria con una densidad espectral de potencia constante, por lo que su autocorrelación es una delta localizada en el origen si la media es nula, y no influirá en modo alguno en la autocorrelación de la señal acústica de interés, salvo en el origen.

En la señal de voz, existen dos tipos de señales dependiendo de si vibran o no las cuerdas vocales, estas son las señales sonoras y sordas, respectivamente. La vibración de las cuerdas vocales produce un tren de deltas que al ajustarse la tensión hace que la señal vibre oscilatoriamente, produciendo en términos de procesado de señal, periodicidad. Sin embargo, las señales sordas se generan interrumpiendo el flujo de aire en el tracto vocal y forzándolo para formar turbulencias, pero no hacen vibrar a las cuerdas vocales, por lo que la señal sorda no es periódica, sino que será una fuente parecida al ruido blanco gaussiano.

La autocorrelación de una señal sonora, es decir, periódica, alcanza un valor máximo en el origen y, después, aparecen réplicas menores a este valor máximo con una separación equivalente al periodo. Por el contrario, la autocorrelación de una señal sorda (aperiódica) tiene una delta muy pronunciada en el origen y valores próximos a cero en el resto, careciendo de cualquier tipo de estructura, tal como pasa en el caso del ruido blanco gaussiano.

2.4 ANÁLISIS EN EL DOMINIO DE LA FRECUENCIA

Anteriormente se ha evaluado que analizando la huella acústica de un dron en el dominio temporal no es suficiente, por lo que se necesita pasar al dominio frecuencial, para así poder obtener información acerca de la contribución de cada frecuencia a la formación de la señal. Además, en el dominio frecuencial se pueden obtener parámetros importantes como pueden ser la frecuencia fundamental, el timbre (la cantidad de armónicos y su amplitud distinguen entre dos sonidos del mismo tono) o la altura (sonido grave o agudo).

Asimismo, en el dominio del tiempo también se observó que la señal acústica de un dron no es estacionaria, es decir, sus características varían a lo largo del tiempo.

La densidad espectral de potencia de señales estacionarias puede estimarse con técnicas paramétricas y no paramétricas. Las técnicas paramétricas se basan en la asunción de un modelo de producción. Cuando no se conoce dicho modelo, pueden aplicarse técnicas no paramétricas, como el periodograma. Para calcular el periodograma de la señal es necesario aplicar la transformada de Fourier. La transformada de Fourier consiste en transformar la señal en función del tiempo a una señal en función de la frecuencia. Una propiedad importante de esta transformada es que es reversible, es decir, se puede volver a obtener la señal original con la transformada inversa.

En primer lugar, se va a introducir la transformada de Fourier continua. Esta transformación ha de tener como argumento de entrada una señal analógica, por ejemplo, una señal de audio $s(t)$.

$$\begin{cases} S(\omega) = \int_{-\infty}^{\infty} s(t) e^{-j\omega t} dt \rightarrow \text{Ecuación de análisis} \\ s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{j\omega t} d\omega \rightarrow \text{Ecuación de síntesis} \end{cases} \quad (3)$$

El resultado obtenido con la transformada de Fourier es una función de la frecuencia con información de las amplitudes de las distintas frecuencias de la señal de entrada, en otras palabras, cómo contribuye cada frecuencia a la formación del sonido. Sin embargo, una señal continua no es computable por lo que se necesita muestrear la señal temporal para poder operar con valores discretos.

Con el fin de muestrear la señal, es decir, de convertir una señal analógica en una señal digital, hay que tener en cuenta varios factores, siendo el primero de ellos el Teorema de Muestreo. Dicho teorema dice:

Siendo $x(t)$ una señal limitada en banda con $X(\omega) = 0$ para $|\omega| > \omega_M$, entonces $x(t)$ está determinada únicamente por sus muestras $x(nT)$, $n = 0, \pm 1, \pm 2, \dots$ si $\omega_s > 2\omega_M$ donde $\omega_s = 2\pi/T$. Dadas estas muestras, se puede reconstruir $x(t)$ generando un tren periódico de impulsos en el que los impulsos sucesivos tengan amplitudes de valores de muestras sucesivas. Este tren de impulsos es después procesado por un filtro paso bajo con ganancia T y

frecuencia de corte mayor a la w_M y menor a $(w_s - w_M)$. El resultado será una señal igual a $x(t)$. [48]

Es decir, si una señal está limitada en banda y sus muestras están lo suficientemente cerca unas de otras en relación con la frecuencia más alta presente en la señal, se podrá reconstruir la señal perfectamente. Por otro lado, para que no haya pérdida de información, si la señal tiene una banda de frecuencia limitada, dicha señal tiene que ser muestreada con una frecuencia mayor a dos veces la frecuencia de banda. Si esto no sucediera, existiría aliasing (la señal se solaparía de una banda a otra y se perdería parte de la información).

Otro aspecto importante es la interpolación, siendo necesario convolucionar la señal digital con una función de interpolación para recuperar la señal original.

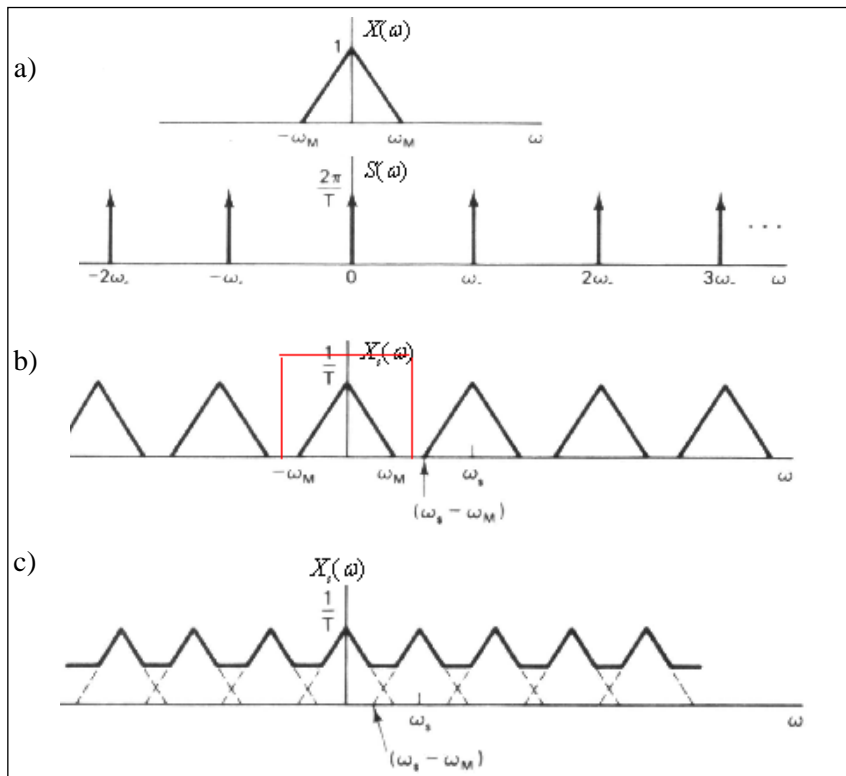


Figura 8 – Espectros de: a) La señal original y de la función de muestreo; b) La señal muestreado con $w_s > 2w_M$; c) La señal muestreado con $w_s < 2w_M$ (adaptado de [5]).

Tras muestrear la señal continua, se obtiene una señal de entrada de valores discretos.

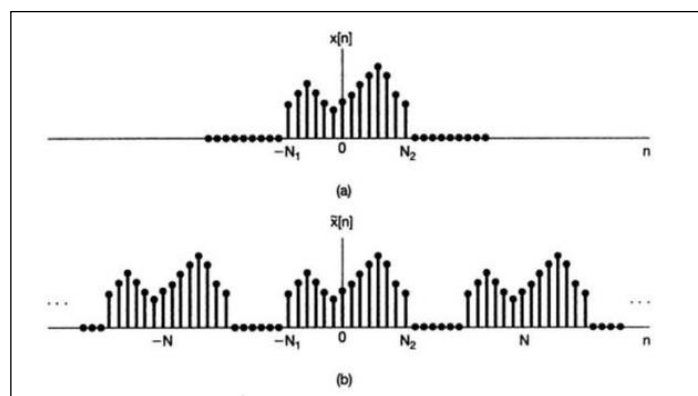


Figura 9 – Transformación de una señal aperiódica y finita a una señal periódica e infinita [5].

Capítulo 2 – Caracterización espectral de los UAVs

Existe una relación entre la transformada de Fourier de una señal aperiódica y los coeficientes del desarrollo en series de Fourier de la señal que se obtiene mediante la extensión periódica de la señal original. La extensión periódica de la señal $x[n]$ se representa en la Figura 9. La relación entre la transformada $X(\Omega)$ y los coeficientes a_k de la señal periódica de frecuencia fundamental Ω_0 se expresa en (4):

$$X(\Omega) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\Omega n} \quad a_k = \frac{1}{N}X(k\Omega_0) \quad (4)$$

Si se combinan ambas expresiones, y teniendo en cuenta que Ω_0 se usa para referirse al espacio entre muestras $2\pi/N$, se obtiene lo siguiente:

$$\tilde{x}[n] = \frac{1}{2\pi} \sum_{k=\langle N \rangle} X(k\Omega_0)e^{jk\Omega_0 n} \quad (5)$$

Si $N \rightarrow \infty$, $\Omega_0 \rightarrow 0$, esta ecuación se convierte en una integral de límite 2π . Las siguientes expresiones definen la transformada de Fourier de señales de tiempo discreto:

$$\begin{cases} X(\Omega) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\Omega n} \rightarrow \text{Ecuación de análisis} \\ x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\Omega)e^{j\Omega n} d\Omega \rightarrow \text{Ecuación de síntesis} \end{cases} \quad (6)$$

De este modo, la transformada de Fourier de tiempo discreto, es una función continua en la pulsación de tiempo discreto, por lo que tampoco va a poder ser computable en un ordenador. Entonces, de la misma forma que se ha muestreado la señal de tiempo continuo para formar $x[n]$ en el dominio del tiempo, ahora se ha de muestrear el espectro de la señal $x[n]$ después de ser aplicada la DTFT. Esta operación define la DFT (Transformada Discreta de Fourier).

$$X[k] = X(\Omega)|_{\Omega=\frac{2\pi}{M}k}, k = 0, 1, \dots, M-1 \quad (7)$$

Al muestrear la función $X(\Omega)$ se obtienen muestras en frecuencia que son proporcionales a los coeficientes del desarrollo en series de Fourier de la extensión periódica $\tilde{x}[n]$.

$$X[k] = \begin{cases} X\left(e^{j\left(\frac{2\pi k}{M}\right)}\right), & 0 \leq k \leq M-1 \\ 0, & \text{resto} \end{cases} \quad (8)$$

$$x_p[n] = \begin{cases} \sum_{r=-\infty}^{\infty} x[n-rM], & 0 \leq n \leq M-1 \\ 0, & \text{resto} \end{cases} \quad (9)$$

Para que $x_p[n] = x[n]$ hay que evitar el solapamiento temporal. Esto se consigue con más muestras en frecuencia que las que tiene la señal finita $x[n]$ en el tiempo, por lo tanto, $M > N$.

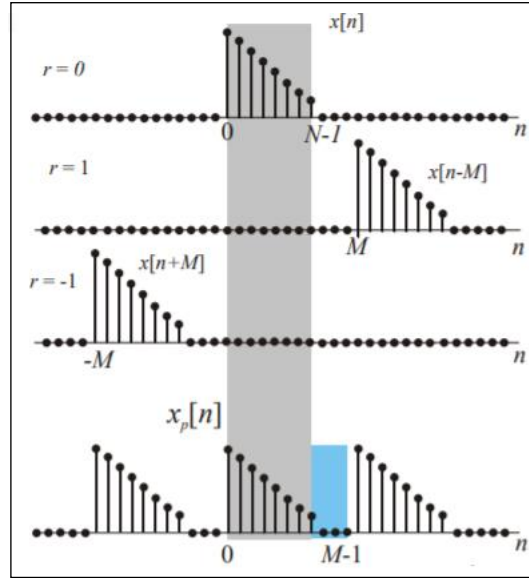


Figura 10 – Procedimiento para evitar solapamiento temporal en la DFT [6].

Finalmente, obtenemos las siguientes ecuaciones para la transformada discreta de Fourier (DFT):

$$\begin{cases} X[k] = \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi}{N}kn}, & k = 0, 1, \dots, N-1 \rightarrow \text{Ecuación de análisis} \\ x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{j\frac{2\pi}{N}kn}, & n = 0, 1, \dots, N-1 \rightarrow \text{Ecuación de síntesis} \end{cases} \quad (10)$$

Para utilizar la transformada discreta de Fourier (DFT) de forma óptima, se aplicará en *MATLAB* el algoritmo FFT (Transformada Rápida de Fourier) que es un algoritmo DIT (Decimation-In-Time algorithm). La FFT es una herramienta que calcula la DFT con el menor número de cálculos posible, ya que elimina gran parte de los cálculos repetidos, siendo mucho más rápida que la DFT. La FFT de raíz-2 divide la señal en dos de duración mitad, de forma iterativa, hasta poder realizar transformadas de dos muestras. Después, se van agrupando dichas transformadas y se resuelven de nuevo hasta llegar al nivel de muestras de entrada. Para poder realizar la FFT de la forma más eficiente, debe tener como argumento de entrada un número de puntos de la señal correspondiente a una potencia de 2, definiendo así el algoritmo de raíz-2.

2.4.1 Análisis frecuencial de un UAV

Una vez explicado el marco teórico de la transformada de Fourier, se va a aplicar a la señal acústica del dron dicha transformada para poder conocer como contribuye cada frecuencia a la formación del sonido.

Para ello, se va a aplicar la FFT, donde se ha seleccionado un número de puntos correspondiente a la siguiente potencia de base 2 del número de muestras de la señal acústica. Esto rellenará la señal de audio con ceros hasta obtener dicha potencia para poder mejorar el rendimiento de la FFT. Posteriormente se calcula el módulo de la FFT partido del número de muestras seleccionado, para así poder obtener el espectro de la señal.

Al calcular el espectro de las huellas acústicas de los drones se obtiene la siguiente figura.

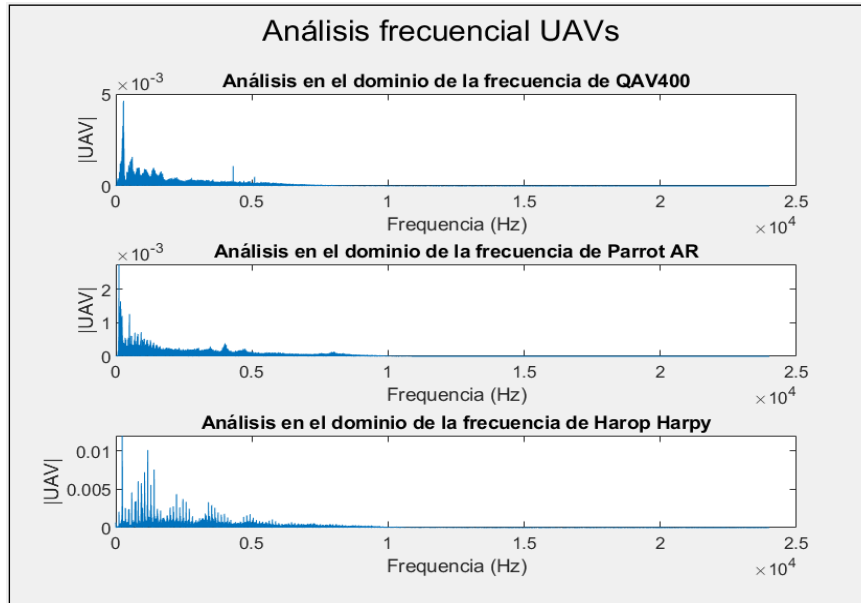


Figura 11 – Representación de la señal de los UAVs en el dominio de la frecuencia.

Las representaciones de la Figura 11 muestran que la mayoría de la información de la señal la albergan las frecuencias menores de 3000 hercios. Si se representa un detalle de los espectros de las señales de forma que solo muestren la información obtenida en el rango entre 0 y 3 kHz, se puede apreciar más información de los espectros.

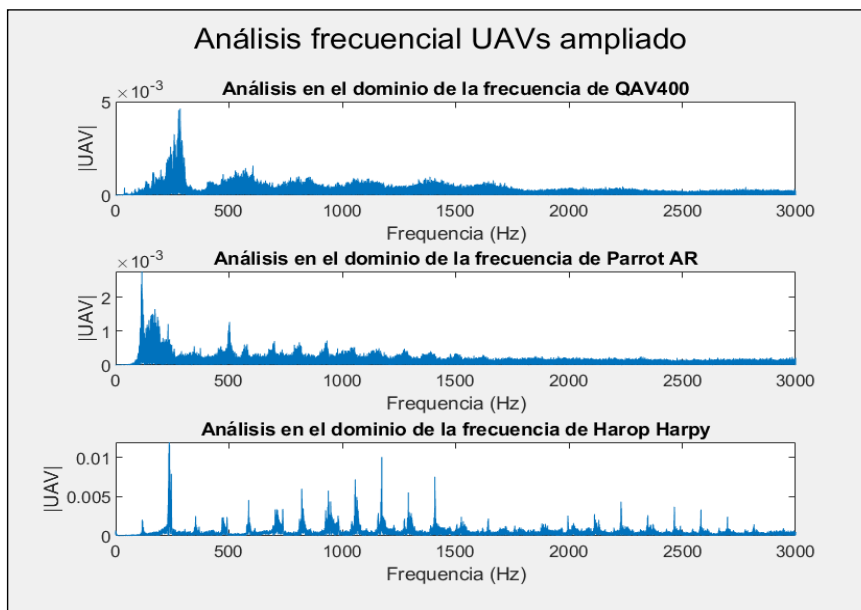


Figura 12 – Representación frecuencial de la señal de los UAVs en el rango entre 0 y 3 kHz.

Como se puede observar, el sonido producido por los rotores de un dron es una señal en la que se distinguen una serie de picos a frecuencias que son aproximadamente múltiplos de una frecuencia fundamental, por lo que se adivina cierta estructura periódica, que estará relacionada con la velocidad de giro de los rotores. Es una señal de banda ancha rica en armónicos. Resulta importante conocer la frecuencia fundamental y, la banda de frecuencia en la que se concentra la energía, que puede no coincidir con la banda donde se localice la frecuencia fundamental.

Además, debido a que la huella acústica de un UAV no es estacionaria, se debe aplicar un espectrograma, para poder ver cómo evoluciona la potencia de la señal en función tanto del tiempo como de la frecuencia, ya que la transformada de Fourier no ofrece información sobre el momento en el que la señal presenta las características frecuenciales exhibidas y enmascara en la información espectral las variaciones locales bruscas que aportan poca energía [7].

2.5 ESPECTROGRAMA

La gran debilidad de la transformada de Fourier es que no puede hacer un análisis localmente, es decir, no se puede hacer un análisis de las frecuencias en algún intervalo de tiempo. Debido a esto, se desarrolló un método que permite realizar un análisis tanto en tiempo como en frecuencia [8], para obtener información no sólo de las componentes frecuenciales de la señal, sino de los instantes en los que aparecen.

Para poder calcular el espectrograma de la señal, se tiene que aplicar una STFT (Transformada de Fourier de Tiempo Reducido). La STFT consiste en dividir la señal en segmentos cortos de forma que las características de la señal de entrada en ese tramo de tiempo sean estacionarias y se calcula la transformada de Fourier a cada segmento. De esta forma, se obtiene una función en la que se puede saber qué componente frecuencial es la que tiene mayor amplitud en cada instante de tiempo, tal y como se muestra en la Figura 13. La STFT se implementa mediante la siguiente ecuación:

$$STFT\{x[n]\} = X(m, \Omega) = \sum_{n=-\infty}^{\infty} x[n] h[n - mR] e^{-j\Omega n} \quad (11)$$

En la fórmula anterior, $x[n]$ es la señal de entrada, $h[n - mR]$ es la ventana que selecciona el segmento de la señal donde se va a aplicar la transformada de Fourier, siendo mR el centro de la ventana, y, R es la diferencia entre la longitud de la ventana y el factor de solapamiento entre segmentos.

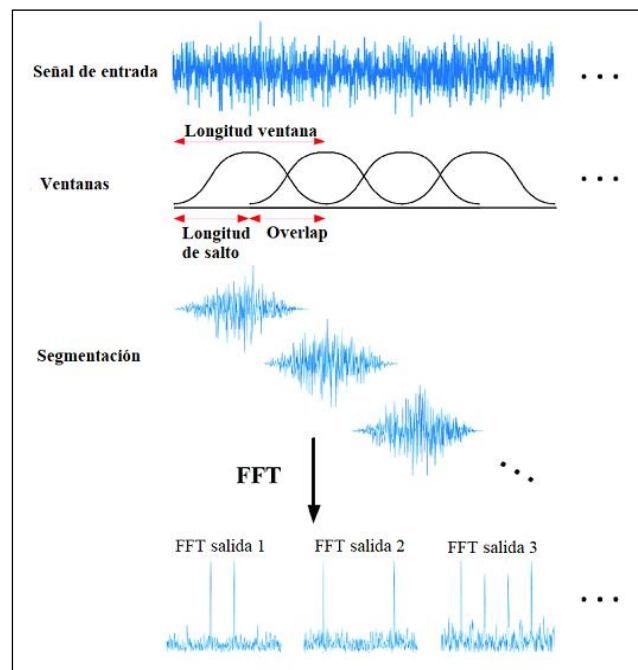


Figura 13 – Proceso de la STFT [9].

Esta transformada, es muy útil para señales no estacionarias, ya que, al cambiar las propiedades de la señal a lo largo del tiempo, se hace necesario evaluar un instante de tiempo muy breve para poder tener las mismas características frecuenciales. Por lo tanto, se divide la señal en pequeños segmentos de igual longitud y se aplica una ventana a cada uno de los segmentos donde se va aplicando la FFT.

Para poder escoger una longitud de segmento en el que se pueda confirmar que las características sean estacionarias, se va a explicar brevemente como debe aplicarse la STFT a una señal de voz. Una señal de voz es una señal no estacionaria, debido a que en el habla se expresan fonemas cuya representación tanto en el tiempo como en la frecuencia varían. En términos de la señal de voz, se puede expresar de 4 a 20 fonemas cada segundo, cuya duración mínima son 50 ms. Por lo tanto, para poder escoger una longitud de segmento en el que se pueda confirmar que las características sean estacionarias, se eligen 20 ms. Es por esto mismo, por lo que elegiremos ventanas de tiempo de esta duración para el caso que se está abordando.

Por otro lado, para conseguir que las tramas consecutivas estén correladas y no se pierda parte de la información de la señal, ya que el eje temporal podría sufrir discontinuidades, se va a introducir un solapamiento de 10 ms, es decir, del 50 % de cada segmento. El enventanado con solapamiento del 50% se representa en la Figura 14.

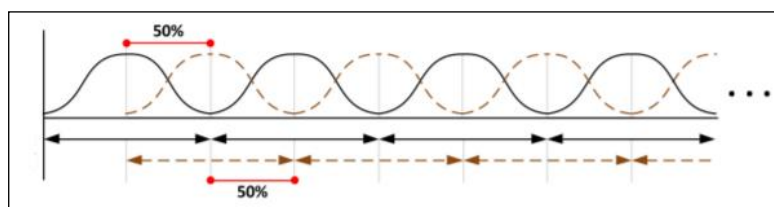


Figura 14 – Señales con solapamiento del 50 % [9].

En cuanto al enventanado, una ventana, en el contexto del procesado de señales, se utiliza para seleccionar un segmento de datos de una señal mediante una multiplicación de la señal por dicha ventana. De esta forma es posible obtener solo el segmento deseado de la señal, además de variar sus características.

2.5.1 Enventanado

El enventanado consiste en la multiplicación de una señal por una función, llamada ventana. En el caso que se está abordando, dicha función es una ventana deslizante que permite seleccionar partes de la señal para efectuar un análisis localizado. De esta forma, la ventana se va desplazando por toda la señal, para poder obtener la respuesta deseada en cada segmento por separado.

A) Longitud de la ventana

Uno de los conceptos de mayor importancia en la estimación del espectrograma es la longitud de la ventana. A mayor longitud de la ventana existirá peor resolución temporal mientras que si la longitud de la ventana es menor, la resolución temporal aumentará mientras que la frecuencial disminuirá. Esto se debe a que la selección de la longitud de la ventana está fuertemente ligada al principio de incertidumbre de Heisenberg. Este principio dice que no se puede conocer con exactitud qué valor de frecuencia existe en un momento determinado, sólo se puede conocer qué componentes frecuenciales existen dentro de un intervalo de tiempo. Por ejemplo, si se obtuviera el espectrograma de la señal del QAV400 aplicando una ventana de 64 muestras se obtendría la siguiente figura.

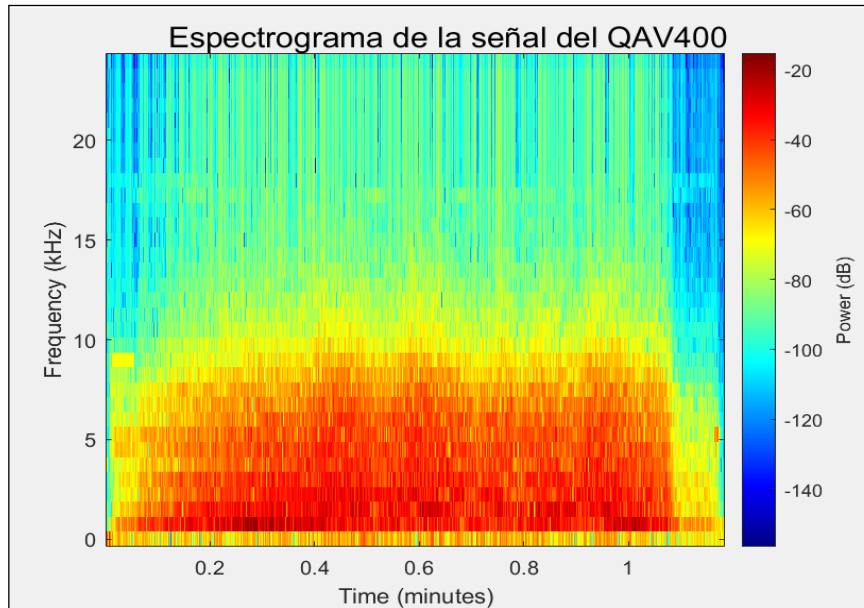


Figura 15 – Espectrograma del QAV400 con una ventana de 64 muestras.

Como se puede observar, la resolución temporal es notablemente superior a la frecuencial, ya que en el eje del tiempo se pueden ver líneas verticales correspondientes a cada instante de tiempo, pero la frecuencia queda difuminada, ya que sólo se podrán observar las componentes de frecuencias con una resolución igual y mayores al inverso del tamaño de la ventana [10].

Por el contrario, si se aplica una ventana de 4096 muestras se obtiene:

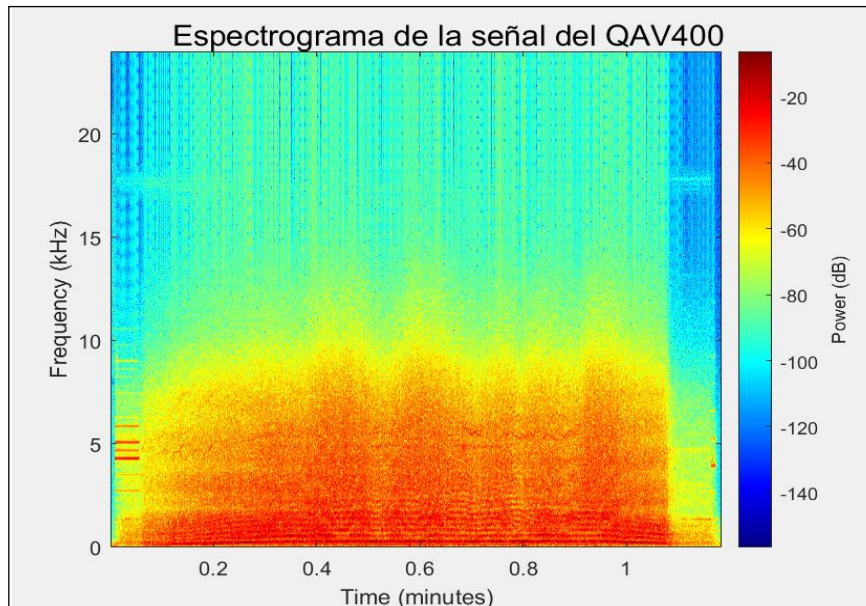


Figura 16 – Espectrograma del QAV400 con una ventana de 4096 muestras.

En esta figura, se ha obtenido el caso contrario, donde la información en el dominio de la frecuencia es muy superior a la del tiempo. Sin embargo, como lo que se está buscando son las frecuencias en las que la señal acústica del dron tiene mayor energía, y que estas frecuencias sean mayormente constantes a lo largo del audio independientemente de la velocidad de los rotores del dron, se ha seleccionado una ventana de 2048 muestras, de forma que tenga buena resolución en frecuencia sin que se pierda la información temporal.

La selección de 2048 muestras no solo indicará la longitud de la ventana, sino los pasos por cero de la función. Es decir, cuanto mayor sea el valor de M muestras, mayor será la oscilación de la ventana en el dominio de la frecuencia, y, por tanto, existirán más nulos en la función. Este efecto se puede ver en la figura siguiente.

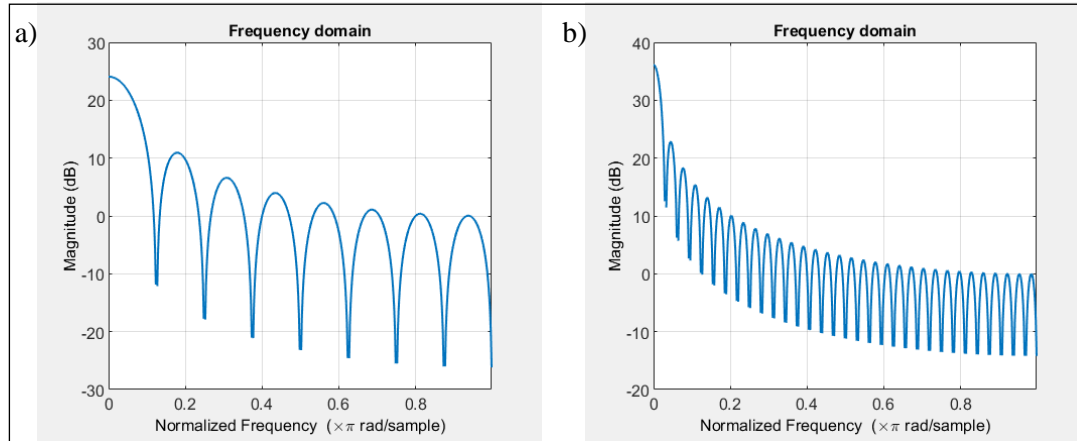


Figura 17 – Oscilación de la ventana en el dominio frecuencial: a) Ventana de 16 muestras; b) Ventana de 64 muestras.

En la Figura 17 se puede observar también como cuánto mayor es el orden de la ventana (número de muestras), menor es la anchura del lóbulo principal, que es el lóbulo de mayor amplitud. Sin embargo, aunque se aumente el orden, los lóbulos secundarios siguen manteniendo la misma relación entre las amplitudes con el principal, ya que los lóbulos siempre abarcan la misma área.

B) Tipos de ventanas

Con el fin de reducir la distorsión espectral causada por el enventanado, es deseable que la función ventana satisfaga, fundamentalmente, dos características: que posea una alta resolución frecuencial, esencialmente un lóbulo principal estrecho y agudo; y, que posea una gran atenuación de los lóbulos laterales con respecto al principal, para que los primeros no afecten al espectro de la señal [11]. Sin embargo, estos dos requisitos son incompatibles entre sí, con lo cual, se debe alcanzar un compromiso para así poder seleccionar la ventana más apropiada para el escenario que se está abordando. Para resolver el problema específico de la voz, en este documento se van a estudiar las ventanas de mayor uso en las aplicaciones de procesamiento de señal, éstas son las ventanas: rectangular, Hanning, Hamming y Blackman.

Ventana rectangular

La ventana rectangular viene dada por las siguientes ecuaciones en el dominio del tiempo y de la frecuencia.

$$w_{rect}[n] = \begin{cases} 1, & 0 \leq n \leq M \\ 0, & \text{resto} \end{cases} \xrightarrow{\text{Transf. Fourier}} W_{rect}(\Omega) = \frac{\text{sen}\left(\Omega\left(\frac{M+1}{2}\right)\right)}{\text{sen}\left(\frac{\Omega}{2}\right)} e^{j\frac{M}{2}\Omega} \quad (12)$$

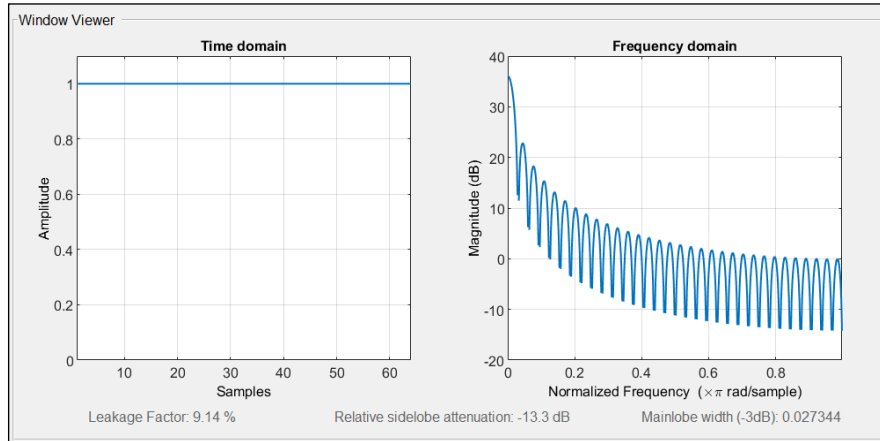


Figura 18 – Ventana rectangular de 64 muestras en el dominio temporal y frecuencial.

La ventana rectangular tiene una amplitud constante de valor unidad en el dominio del tiempo, por lo que equivale a simplemente seleccionar un segmento de la señal, sin modificar ninguna de las características de la señal. Esta señal se suele utilizar en señales periódicas.

Sin embargo, para modificar la amplitud de los lóbulos secundarios, y que así tengan menor peso, se debe modificar la forma de la ventana y utilizar ventanas que no contengan discontinuidades abruptas.

Ventana de Hanning

La ventana de Hanning, llamada por su inventor Von Hann, viene definida de la siguiente forma.

$$w_{hann}[n] = \begin{cases} 0,5 \left(1 - \cos\left(\frac{2\pi n}{M}\right) \right), & 0 \leq n \leq M \\ 0, & \text{resto} \end{cases} \quad (13)$$

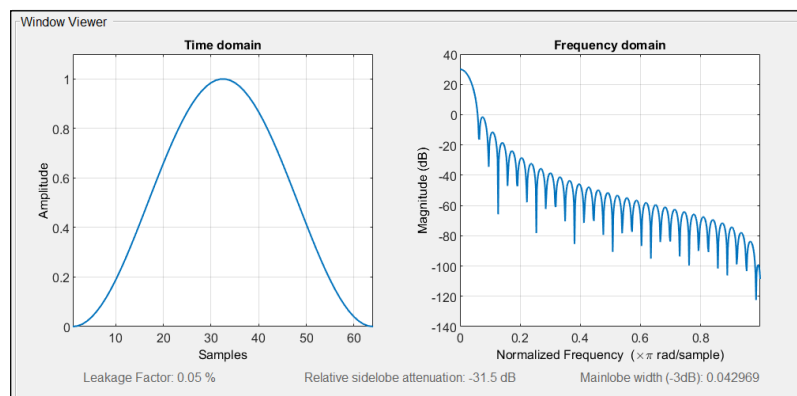


Figura 19 – Ventana de Hanning de 64 muestras en el dominio temporal y frecuencial.

La ventana de Hanning atenúa los extremos de la señal de entrada a cero, por lo que no evalúa el audio de la grabación del dron al inicio ni final del audio. Esta ventana se debe usar con señales continuas y de banda ancha. Además, ofrece buena resolución en frecuencia, aunque distorsiona la señal de entrada en amplitud.

Ventana Hamming

La ventana Hamming viene definida por la siguiente ecuación.

$$w_{hamm}[n] = \begin{cases} 0,54 + 0,46 \cos\left(\frac{2\pi n}{M}\right), & 0 \leq n \leq M \\ 0, & \text{resto} \end{cases} \quad (14)$$

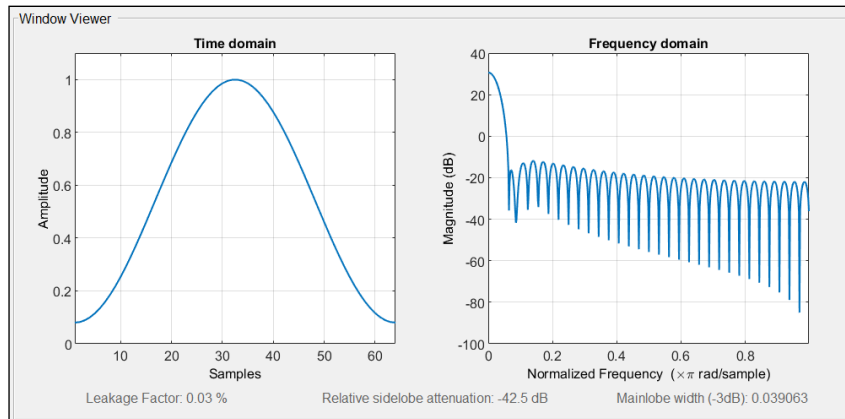


Figura 20 – Ventana de Hamming de 64 muestras en el dominio temporal y frecuencial.

La ventana Hamming genera menor rizado que la de Hanning, es decir, los primeros lóbulos laterales tienen menor amplitud. Sin embargo, los demás lóbulos son más altos que los de Von Hann, por lo que tiene menor selectividad (para tener en cuenta cuando se usa para el diseño de filtros digitales FIR).

Ventana Blackman

La ventana Blackman presenta la siguiente ecuación y forma:

$$w_{blackman}[n] = \begin{cases} 0,42 - 0,5 \cos\left(\frac{2\pi n}{M}\right) + 0,08 \cos\left(\frac{4\pi n}{M}\right), & 0 \leq n \leq M \\ 0, & \text{resto} \end{cases} \quad (15)$$

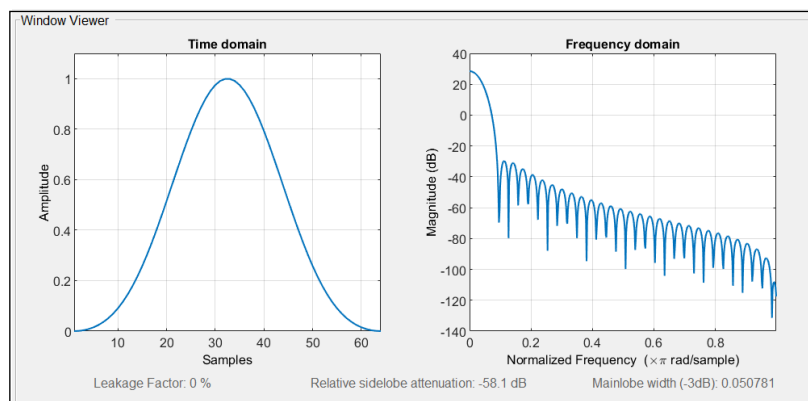


Figura 21 – Ventana de Blackman de 64 muestras en el dominio temporal y frecuencial.

La ventana Blackman tiene una gran ventaja respecto a la amplitud de sus lóbulos laterales, ya que están muy por debajo del principal, en comparación con las ventanas anteriores. No obstante, esta mejora en la amplitud conlleva una mayor anchura en el lóbulo principal.

Finalmente, las ventanas que presentan un lóbulo principal más estrecho, más definido, son las ventanas de Hamming y rectangular. Sin embargo, estas ventanas no atenúan en gran medida los lóbulos secundarios, que son frecuencias parásitas, por lo que dependerá de la aplicación a desempeñar la selección de una u otra ventana. En el caso que se está abordando, se requiere una gran atenuación de las frecuencias parásitas, además de selectividad en el lóbulo principal, para poder seleccionar la frecuencia de mayor energía. Asimismo, se tiene una señal de banda ancha con características pseudoaleatorias. Por lo tanto, se va a seleccionar la ventana de Hanning de 2048 muestras.

2.5.2 Espectrograma de un UAV

Atendiendo a la longitud de cada trama de 20 ms, la longitud del solapamiento de 10 ms, la longitud de la ventana de 2048 muestras y la frecuencia de muestreo que es de 48 kHz, se representan en las figuras siguientes los espectrogramas de los tres drones en cuestión:

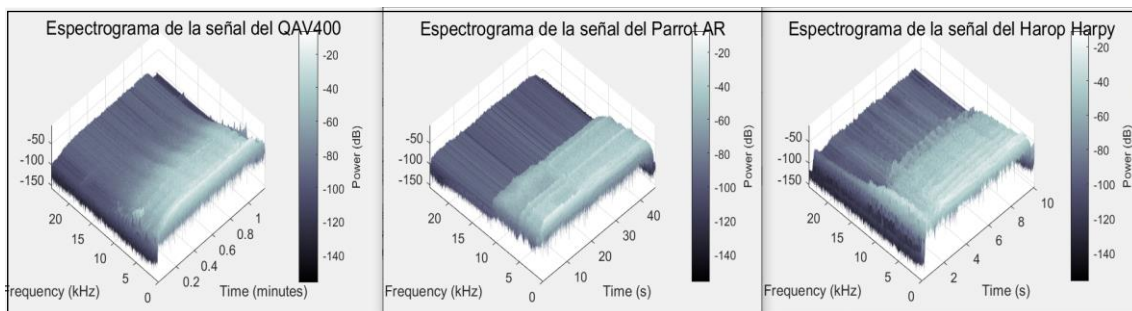


Figura 22 – Espectrograma de los UAVs con vista de $(-45,65)^\circ$.

El análisis anterior se ha realizado con un ángulo de $(-45,65)$ grados, para poder ver en qué rangos de frecuencias son las que se tiene más información espectral. Como se puede observar, en el análisis del espectrograma de los tres drones, la información se condensa en el rango de 0 a 10 kHz.

Por lo tanto, si se vuelve a analizar el espectrograma en este rango se obtiene:

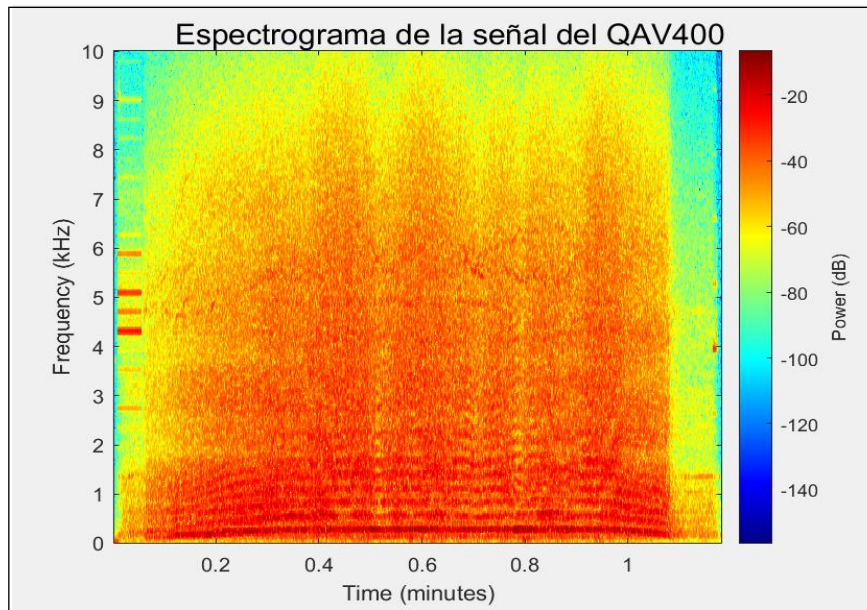


Figura 23 – Espectrograma del QAV400.

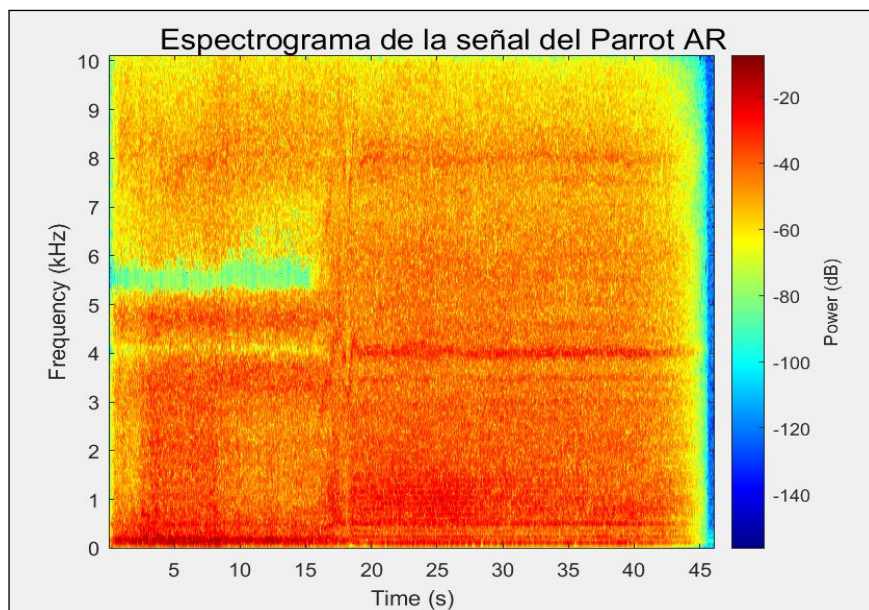


Figura 24 – Espectrograma del Parrot AR.

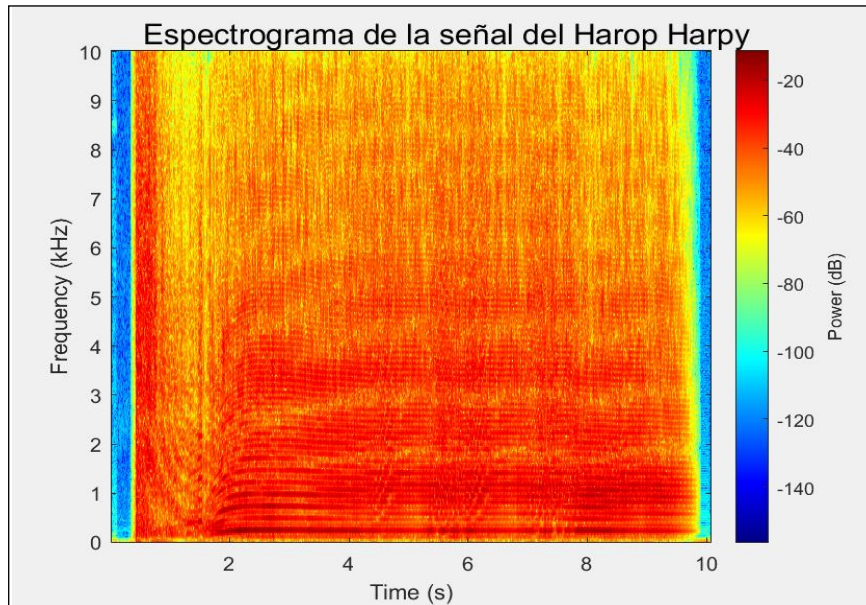


Figura 25 – Espectrograma del Harop Harpy.

En las imágenes anteriores, se puede ver como aparecen líneas horizontales que se repiten casi armónicamente. Estas líneas se denominan HERM (HELicopter Rotor Modulation) y aparecen cuando la longitud de la ventana es suficientemente grande para poder contener varios ciclos de la rotación de las hélices del dron. Debido a que las líneas HERM aparecen equiespaciadas en frecuencia, el pitch o frecuencia fundamental pueden obtenerse de ellas [12]. Además, estas líneas corresponden a la frecuencia fundamental y sus armónicos que van disminuyendo en términos de potencia. Por lo tanto, se puede apreciar como la frecuencia fundamental que se está buscando estará en la banda de 0 a 0,5 kHz.

Sin embargo, en dichos espectrogramas, también se puede observar cómo dependiendo del instante temporal del audio, las bandas frecuenciales crecen, se atenúan o cambian. Por ejemplo, en el espectrograma del QAV400, al inicio y final del audio se obtiene un análisis distinto. Esto es debido a que la señal acústica de este dron proviene de sus cuatro rotores, y va cambiando, dependiendo de qué velocidad lleva el dron o en qué dirección se mueva. Es decir, para que un cuadricóptero pueda volar, tiene que estar constantemente monitorizado y corrigiendo la velocidad de rotación de los motores. Ya que incluso una pequeña brisa de viento hará que el dron necesite una corrección. Esto pasa en cualquier tipo de movimiento (hacia delante detrás, lateral y hacia arriba) [3]. Entonces, si el cuadricóptero se está moviendo, existirán diferencias en las velocidades de rotación de los 4 motores, y, por lo tanto, existirá un sonido distinto en la huella acústica del dron, manifestándose en el espectro la frecuencia de rotación de cada uno de los rotores, que no tienen por qué ser múltiplos de una fundamental. El espectro deja de ser el de una señal pseudo-periódica, pero se sigue apreciando líneas espectrales en las que se concentra la energía. Es por esto mismo por lo que se pueden ver cambios en el análisis espectral en ciertos instantes de tiempo.

El objetivo al calcular el espectrograma era ver si las frecuencias se mantenían constantes independientemente del momento del audio, y salvo algún cambio brusco de velocidad en las hélices del dron, las bandas frecuenciales son prácticamente iguales a lo largo del audio. Por lo tanto, es posible calcular la frecuencia fundamental del dron en términos generales.

2.6 FRECUENCIA FUNDAMENTAL

Con el objetivo de calcular la frecuencia fundamental de un UAV para poder caracterizar la periodicidad de su huella acústica, se puede obtener mediante la autocorrelación o mediante un análisis en cepstrum, debido a que tanto la función de autocorrelación como el cepstrum de una señal periódica, tendrán la misma periodicidad que la señal en cuestión.

2.6.1 Comparación entre dos métodos de estimación del pitch (Autocorrelación y Cepstrum)

Para poder comparar ambos métodos de obtención de la frecuencia fundamental, se va a hacer referencia a los resultados obtenidos de los documentos de la IEEE [13] y [14]. En ambos archivos, los autores utilizan una serie de pruebas diferentes para poder obtener una comparación lo más completa posible entre estos y otros métodos. Esta comparativa trata de cuatro puntos:

- **PRECISIÓN DE LA ESTIMACIÓN DEL PITCH.**

Para estimar la precisión de la obtención del pitch, en los documentos anteriores ([13] y [14]) se ha hecho una comparativa entre la señal de entrada proveniente de una voz masculina y la señal de una voz femenina. En otras palabras, una comparación entre una señal que tuviera la frecuencia fundamental en valores de baja frecuencia y una señal con el pitch en frecuencias altas, respectivamente. Los resultados de esta comparativa muestran que, para señales de alta frecuencia fundamental, el método de la autocorrelación obtiene un mejor resultado, mientras que, en el caso contrario, el método con resultados más precisos será el del análisis cepstral.

En el caso que se está abarcando, la frecuencia fundamental va a estar presente en valores de baja frecuencia, como se ha podido observar tanto en el análisis de la transformada de Fourier, como en el espectrograma de la señal, y es por ello, por lo que, en lo tocante a esta prueba, la estimación del pitch mediante el cepstrum será el más adecuado.

- **ROBUSTEZ ANTE EL RUIDO.**

Para comprobar la robustez ante el ruido, en los documentos de la IEEE [13] y [14], se analizan los detectores de pitch bajo una situación de ruido. Si se añade ruido blanco gaussiano (*AWGN*) a una señal sintetizada con una relación señal a ruido (SNR) de 0 dB, por ejemplo, el método del análisis en cepstrum obtiene resultados adecuados, mientras que el método de autocorrelación es muy susceptible al ruido e incluso para señales de SNR igual a 0 dB, los resultados están muy contaminados para poder realizar una buena estimación de la frecuencia fundamental.

- **COMPLEJIDAD DE LOS ALGORITMOS.**

Con el objetivo de comprobar la complejidad entre un método u otro, basta con observar el análisis matemático necesario para lograr llevarlos a cabo. En el caso del método de obtención de la frecuencia fundamental mediante la autocorrelación, éste se realiza mediante el sumatorio del producto de la señal en cuestión por la misma señal desplazada.

$$R_x[m] = \sum_{n=-\infty}^{\infty} x^*[n]x[n+m] \quad (16)$$

Por otro lado, el método de obtención del pitch mediante el análisis cepstral consiste en realizar la transformada de Fourier de la señal, después, aplicar el logaritmo a dicho espectro y, por último, calcular la transformada inversa del resultado.

$$C(qft) = \mathcal{F}^{-1} \log_{10}(\mathcal{F}(s[n])) \quad (17)$$

Por consiguiente, en principio, la complejidad del análisis cepstral es mucho mayor que la dificultad de la autocorrelación, y esta diferencia aumenta si la señal tiene un número alto de muestras.

Sin embargo, no se debe olvidar que la autocorrelación puede estimarse como la transformada inversa de Fourier del periodograma. Como el periodograma se calcula a partir del módulo al cuadrado de la transformada de Fourier, ambos métodos, el basado en la autocorrelación o en el Cepstrum, tendrían una complejidad semejante. El criterio de complejidad no podrá, por tanto, ser determinante a la hora de decidir por uno u otro algoritmo.

- **PROBLEMAS ASOCIADOS A LA SEGMENTACIÓN DE LA SEÑAL.**

Por último, en [13] y [14] se trata la importancia de la selección de la longitud del segmento a analizar para ambos estimadores, debido a que tanto la autocorrelación como el Cepstrum, calculan la media del periodo de pitch de un segmento. Por lo tanto, para poder hacer este cálculo necesita haber un mínimo de dos periodos de pitch en el segmento sonoro que se está analizando para poder obtener una buena estimación. Sin embargo, también existe cierta limitación en la cantidad máxima de muestras del segmento. Esta última limitación se debe a que los algoritmos propuestos no son capaces de detectar las variaciones no estacionarias de la longitud del periodo de pitch entre un periodo y otro.

Para concluir, el método que se ha seleccionado en el desarrollo de este trabajo para poder localizar la frecuencia fundamental de cada uno de los drones ha sido el Cepstrum. Esta selección se debe al análisis realizado en los puntos anteriores. En primer lugar, a frecuencias bajas, donde está localizado el pitch de un dron, el estimador basado en el análisis cepstral obtiene resultados destacables, además de poseer mayor robustez frente al ruido que el método de la autocorrelación. Después, tocante a los problemas de segmentación, ambos estimadores tienen las mismas restricciones de la longitud de la trama. Por último, la complejidad del cepstrum y del método basado en la autocorrelación son semejantes. No obstante, este punto se tiene menos en consideración en comparación con los anteriores, ya que este análisis se va a llevar a cabo una sola vez, y una pequeña diferencia en complejidad no es importante.

A) CEPSTRUM

Para definir el concepto del cepstrum se va a hacer referencia al documento de ‘*Short-Time "Cepstrum" Pitch Detection*’ de los autores A.M. Noll y M.R. Schroeder [15], donde explican la definición de cepstrum en base a una señal de voz.

Debido a que las periodicidades temporales en una señal de voz pueden causar ondas periódicas en el espectro de amplitud, la transformada de Fourier del espectro nos da la “frecuencia” de la onda, que es inversamente proporcional a la frecuencia fundamental de la señal de voz. Al aplicarse el logaritmo al espectro

de amplitud se obtiene el efecto aditivo entre el tracto vocal (envolvente del espectro) y de la fuente vocal (estructura fina del espectro), de este modo se separan las bajas “frecuencias” de la envolvente del espectro, de las usualmente altas frecuencia de la fina estructura del espectro. La detección del pitch mediante el análisis en cepstrum no es sensible a la distorsión de fase, distorsión de amplitud, ruido aditivo ni a la ausencia de la frecuencia fundamental del habla [15].

En la producción de la señal de voz, la excitación glotal y la respuesta impulsiva del filtro que modela el tracto vocal no se combinan linealmente, sino mediante una convolución. El análisis del cepstrum permite representar la señal en un dominio donde se combinan excitación y respuesta impulsiva linealmente, además de aparecer separadas en el dominio de la cuefrecuencia.

El análisis en cepstrum se define como la transformada de Fourier del logaritmo del espectro de una señal.

$$C(qft) = \mathcal{F}^{-1} \log_{10} \left(\mathcal{F}(s(t)) \right) \quad (18)$$

En la fórmula anterior, se define *qft* como cuefrecuencia, aunque dicha variable tenga unidades de tiempo, se añadió este término para no mezclar el tiempo, la frecuencia y el dominio del análisis cepstral. Además, también se denominó *cepstrum* al espectro del logaritmo del espectro de una onda en el dominio del tiempo, para evitar confundirlo con el *spectrum* en el dominio de la frecuencia.

El análisis cepstral es una representación de la señal en el procesado homomórfico. El tratamiento homomórfico se utiliza para poder transformar señales no lineales a un dominio lineal. El cepstrum funciona de la siguiente forma: en primer lugar, se aplica la transformada de Fourier de tiempo discreto (DTFT) a una señal de audio digital, cuyas características son reales, por lo tanto, al aplicar esta transformada, el espectro resultante será complejo. A continuación, se muestra la ecuación resultante al aplicar la DTFT.

$$S(e^{j\Omega}) = \sum_{n=-\infty}^{\infty} s[n] e^{-j\Omega n} \quad (19)$$

En segundo lugar, se aplica el logaritmo a la ecuación anterior, para poder conseguir el efecto aditivo entre el tracto vocal y la excitación glotal. De esta forma, aplicando la propiedad logarítmica donde se puede transformar un producto a una suma, se consigue separar la parte real de la compleja.

$$\log \left(S(e^{j\Omega}) \right) = \log |S(e^{j\Omega})| + j \arg \{ S(e^{j\Omega}) \} \quad (20)$$

Por último, se aplica la ecuación de síntesis de la DTFT, es decir, la inversa de la DTFT, de forma que se obtiene el cepstrum de la señal en el dominio de la cuefrecuencia, donde la señal está definida entonces en un dominio lineal y se podrán usar técnicas convencionales del procesado de señal.

$$C(qft) = \mathcal{F}^{-1} \log_{10} \left(\mathcal{F}(s(t)) \right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [\log |S(e^{j\Omega})| + j \arg \{ S(e^{j\Omega}) \}] e^{j\Omega n} d\Omega \quad (21)$$

Por lo tanto, se ha comprobado que, para una señal real, la transformada de Fourier de tiempo discreto es compleja y que el logaritmo de este espectro se puede calcular como la suma del módulo más la unidad imaginaria *j* por la fase. Por otra parte, la transformada de Fourier (TF) y su inversa están caracterizadas por la dualidad, es decir, tienen las mismas propiedades básicas.

Por consiguiente, el módulo de la TF de una señal real y par será a su vez real y par, y la TF de una función impar será imaginaria pura.

En el caso que se está tratando, se está buscando obtener el pitch de la señal, y para ello, en este momento no es importante la parte impar de la señal, que, únicamente contiene la información relativa al argumento. Es por ello por lo que se va a trabajar con el cepstrum real, que trabaja con la parte real del cepstrum complejo y sirve para separar la excitación glotal de la envolvente del tracto vocal.

$$C(qft) = Re \left\{ \mathcal{F}^{-1} \log_{10} \left(\mathcal{F}(s(t)) \right) \right\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |S(\Omega)| e^{j\Omega n} d\Omega \quad (22)$$

Con el fin de poder calcular la frecuencia fundamental de la huella acústica del dron, la parte de la señal que es interesante para ello es la excitación, que en el caso que se está abordando, es causada por la rotación de los rotores del UAV. Por lo tanto, se puede usar esta excitación para poder determinar la periodicidad del espectrograma, que está relacionado con la frecuencia de torno de la hélice [16]. Por consiguiente, para poder aislar esta parte de la señal, es necesario realizar un liftado.

El término liftado fue introducido por Bogert et al. para nombrar a un filtrado lineal del logaritmo del espectro. A través del liftado, es posible aislar la excitación y poder obtener la frecuencia fundamental de la señal sonora; para ello, hay que aplicar un filtro paso banda en las frecuencias correspondientes donde aparecían las líneas HERM (líneas horizontales periódicas que se repiten periódicamente en el espectrograma) con mayor intensidad.

El liftado, paso banda, en cuestión, ha de evitar los valores más bajos de las cuefrecias, debido a que en estos valores son donde se concentrará la función de banda ancha correspondiente a la envolvente de la señal, y el objetivo es quedarse con la parte de la excitación de la señal para poder obtener el pitch del sonido del dron.

B) CEPSTRUM DE UN UAV

Con el fin de calcular el análisis cepstral de los drones QAV400, Parrot AR y Harop Harpy se va a utilizar el software de *MATLAB*. En primer lugar, para calcular el cepstrum se usa la función *rceps.m* que calcula directamente el cepstrum real de la señal sonora que se le pase como argumento de entrada.

```
% Cepstrum
cepstrum=rceps(audio);

figure;
plot(tiempo,cepstrum);
title(['Análisis cepstral del dron ' num2str(uavs(ii))]);
xlabel('Cuefrecia');ylabel('Gamnitud');
```

Al calcular la función del cepstrum, como pasaba en el caso de la representación del periodograma, se aplica la FFT (Transformada Rápida de Fourier), que es el algoritmo que calcula la DFT (Transformada Discreta de Fourier) de forma rápida, debido a que la DFT es el resultado de muestrear la DTFT. Esto es debido a que la señal de salida de una DTFT es una función continua y no podría implementarse. De este modo, al aplicar la función del cepstrum y representar el resultado directamente, se obtiene, por ejemplo, en el caso del QAV400 la siguiente representación.

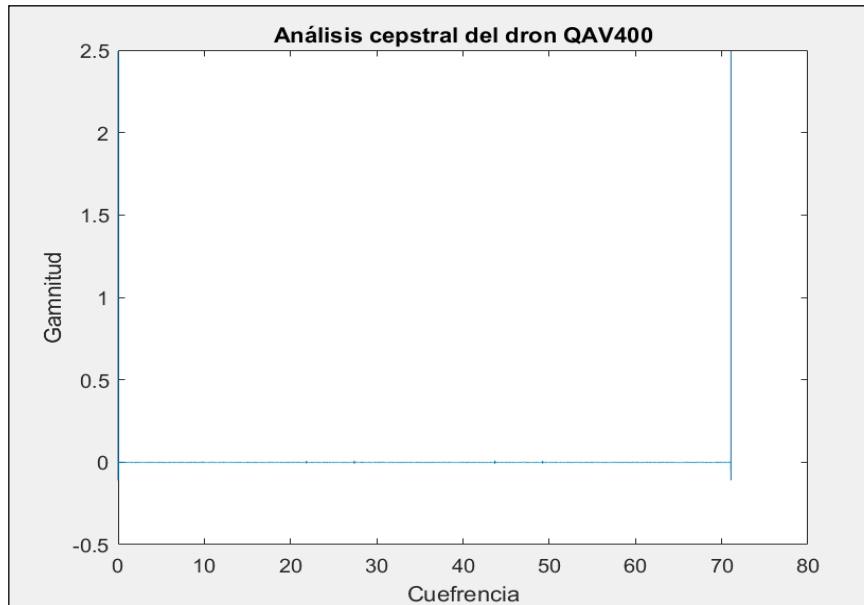


Figura 26 – Análisis cepstral del QAV400.

Esta figura no ofrece ninguna información acerca de la frecuencia fundamental, debido a lo que se comentó en el apartado anterior donde se explicaba el marco teórico del cepstrum. Esto es, el análisis cepstral separa las dos partes de la señal (envolvente y excitación) de forma que se combinan linealmente, pero es necesario aplicar un liftado paso banda para poder seleccionar la parte de la señal que se quiera analizar.

Por lo tanto, para poder localizar la frecuencia fundamental, es necesario aislar la parte de la señal correspondiente a la excitación, que es la parte periódica de la señal producida por la frecuencia de torno de las hélices del dron. Luego es necesario poner a cero la envolvente del cepstrum.

Dicha envolvente, como se comentó anteriormente, aparece reflejada en los valores más bajos de cuefrecencia, por lo tanto, se eliminan estos valores. Hay que tener en cuenta que la unidad en la que se representa la cuefrecencia son los segundos, por lo que es inversamente proporcional a la frecuencia ($Hz = 1/s$). Por consiguiente, la frecuencia máxima del filtro paso banda corresponderá al inicio del filtro paso banda en el dominio de la cuefrecencia. Debido a esto, para poder eliminar la parte de la envolvente de la señal acústica (valores bajos de cuefrecencia), se elige una frecuencia de 500 hercios, es decir, de 0,002 segundos de cuefrecencia. Este valor tiene una explicación y es que la primera línea HERM de los espectrogramas siempre aparece por debajo de 500 hercios, y, además, se intenta limitar al máximo la banda del cepstrum que se está analizando sin coger parte de la señal de excitación. De la misma forma se cogen 50 hercios, es decir, 0,02 segundos como valor máximo del filtro paso banda en el dominio de la cuefrecencia.

Una vez seleccionado el liftado, se representa el resultado, donde el pico más alto del cepstrograma corresponderá a la frecuencia fundamental de la firma de la línea HERM que será N veces el periodo de rotación del rotor donde N es el número de aspas de las hélices [17].

Con el fin de hallar la representación del análisis cepstral en conjunto con el liftado y la selección de la frecuencia fundamental, se implementa el siguiente código en *MATLAB*:

```

% Cepstrum aplicando el liftado
fpb_inic=fs*(1/0.5e3); %Inicio del liftado
fpb_fin=fs*(1/50); %Final del liftado

cepstrum_lift=cepstrum(fpb_inic:fpb_fin);
eje_cuefrecia=(0:length(cepstrum_lift)-1)/fs;

[peaks,Cuefreqs]=findpeaks(cepstrum_lift,eje_cuefrecia);
valor_max=find(peaks==max(peaks));
Freq_fund=fs/(Cuefreqs(valor_max)*fs+fpb_inic);

figure;
plot(eje_cuefrecia,cepstrum_lift);
hold on;
plot(Cuefreqs(valor_max),peaks(valor_max),'md','Linewidth',2);
xlabel('Cuefrecia');ylabel('Gannitud');
suptitle(['Cepstrum del ' num2str(uavs(ii))]);
title(['Frecuencia fundamental = ' num2str(Freq_fund) ' Hz']);
  
```

El código anterior da pie a los siguientes resultados:

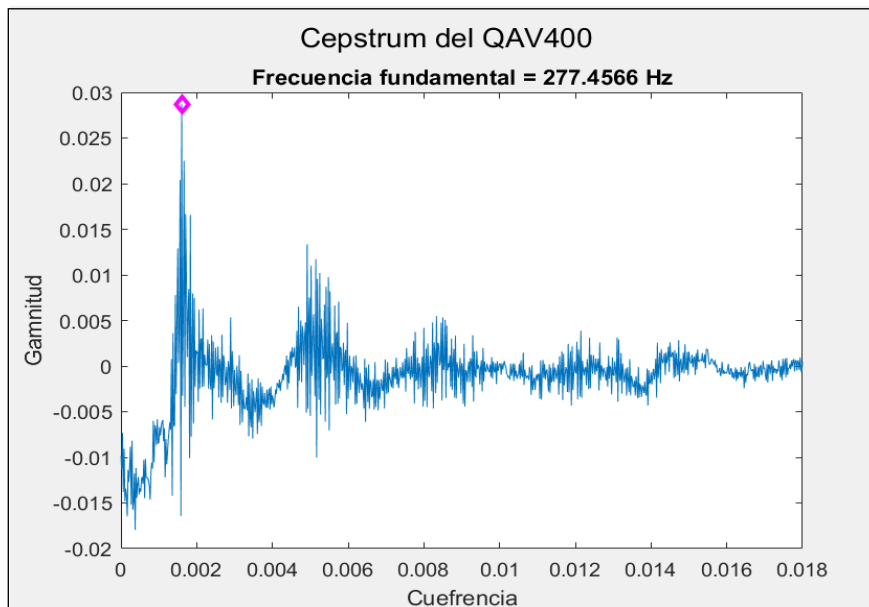


Figura 27 – Análisis cepstral con liftado del QAV400.

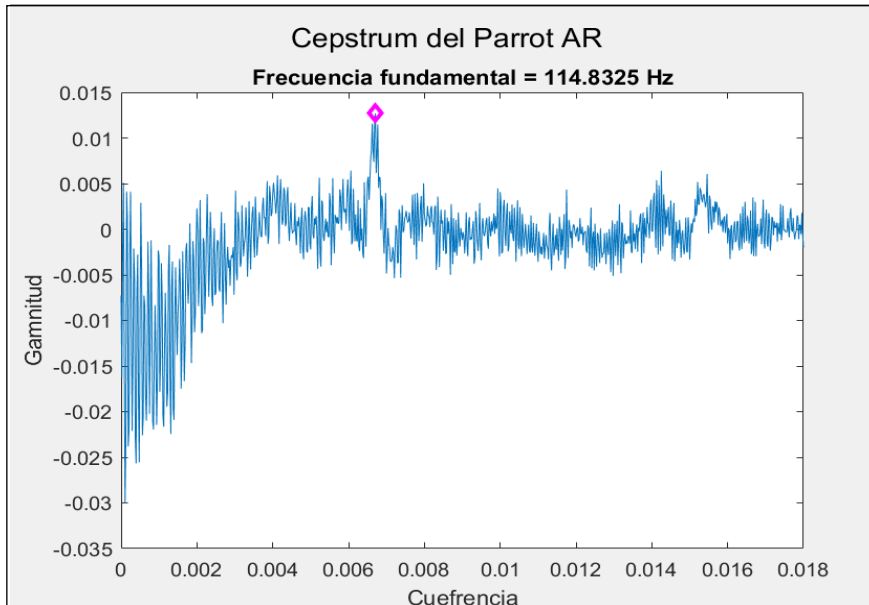


Figura 28 – Análisis cepstral con liftado del Parrot AR.

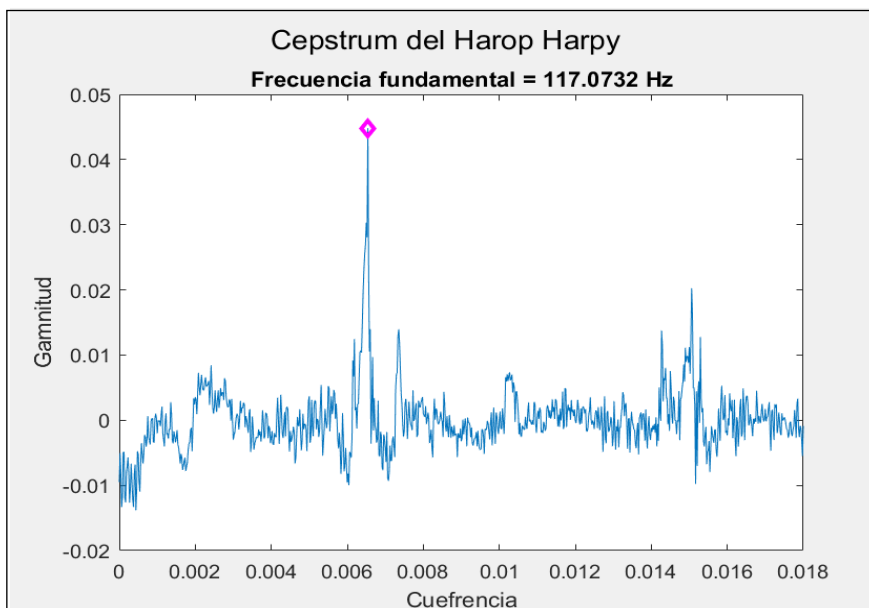


Figura 29 – Análisis cepstral con liftado del Harop Harpy.

En conclusión, aplicando el análisis cepstral con un liftado de 50 a 500 hercios ha sido posible obtener la frecuencia fundamental de los UAVs bajo análisis, mostrada en las figuras 27-29. El siguiente paso será caracterizar el medio de propagación por el que se transmite el sonido. Sin embargo, en primer lugar, se calculan las frecuencias armónicas de la señal de un UAV, con el objetivo de observar si los rotores de los drones tienen la misma velocidad de rotación.

2.6.2 Armónicos de la frecuencia fundamental

Con el objetivo de calcular la distancia a la que podría detectarse un UAV mediante el sonido que produce, se va a calcular la atenuación debida a la propagación del sonido. Esta atenuación viene dada en función de la distancia entre la fuente de sonido y el array de micrófonos, y en función de la frecuencia, como se especificará en la siguiente sección. Por lo tanto, se hace necesario no

solo calcular la frecuencia fundamental de la señal, sino también sus armónicos, para poder elegir estas frecuencias como los centros de los algoritmos de banda de octava que se usarán en el análisis de la atenuación.

Una señal periódica, o pseudo-periódica, está caracterizada en el dominio de la frecuencia por su frecuencia fundamental (pitch), frecuencia en la que la señal tendrá mayor energía, y una serie de componentes armónicos presentes en frecuencias múltiplos de ella, que tienen menor amplitud normalmente. Estas frecuencias vienen dadas por el sonido de la rotación de las aspas del dron, ya que es una señal periódica. Además, la frecuencia fundamental debe ser idéntica a la separación entre las líneas HERM.

Para poder obtener el valor de estas frecuencias armónicas, se va a hacer uso del primer algoritmo descrito en [18], donde se especifican una serie de pasos a llevar a cabo.

En primer lugar, la señal sonora se segmenta en intervalos de tiempo de longitud $T_d = T/K$, siendo T el tiempo que dura la señal y K el número de intervalos. Después, se calcula la transformada rápida de Fourier (FFT) para obtener el espectro de la señal y así tenerla caracterizada en el dominio frecuencial.

En segundo lugar, se escoge un intervalo de búsqueda, restringiendo el primer intervalo de $f_L = 100 \text{ Hz}$ a $f_R = 300 \text{ Hz}$. En este intervalo estará localizada la frecuencia fundamental del UAV, pero, de nuevo, se usa el valor obtenido mediante el análisis cepstral, ya que dará lugar a un valor más preciso. A continuación, se selecciona una ventana que irá desplazándose, para poder ir localizando los armónicos de la señal. Debido a que las frecuencias de los armónicos son aproximadamente múltiplos enteros de la frecuencia fundamental, la ventana vendrá dada de la siguiente manera: $\left[\frac{n}{n-1} f_{n-1} - f_t, \frac{n}{n-1} f_{n-1} + f_t \right]$, donde f_t es un umbral de, por ejemplo, 25 muestras y $n \geq 2$.

Algorithm 1: The Harmonic Frequencies Estimation	
Input:	The sampled data $s(t)$ in time domain;
Output:	The frequency estimation results f_1, f_2, \dots, f_N ;
1	Initialize f_L, f_R, f_t, N ;
2	compute FFT of $s(t)$;
3	search the f_1 in maximum peak value in spectrum between $[f_L, f_R]$;
4	for $n = 2; n \leq N; n++$ do
5	search the f_n in maximum peak value in $\left[\frac{n}{n-1} f_{n-1} - f_t, \frac{n}{n-1} f_{n-1} + f_t \right]$;
6	end
7	return f_1, f_2, \dots, f_N ;

Figura 30 – Algoritmo de estimación de las frecuencias armónicas [18].

En la Figura 31, se han marcado con color rojo las frecuencias detectadas como armónicos del fundamental para los sonidos producidos por los tres UAVs bajo estudio. Estas frecuencias de los armónicos, obtenidas en la Figura 31, serán las que se introducirán en el análisis de la propagación del sonido para poder localizar las características y seleccionar la banda de frecuencia que menor atenuación posea para poder observar la mayor relación señal a ruido posible.

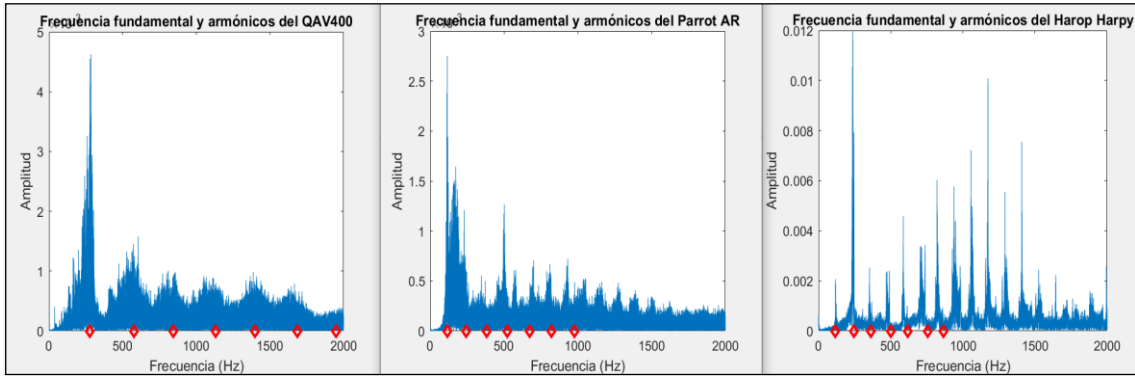


Figura 31 – Cálculo de los armónicos mediante la FFT de las señales acústicas de los UAVs.

3 PROPAGACIÓN DEL SONIDO EN AMBIENTES REALES

La emisión de sonido por una fuente sonora implica el movimiento de las partículas de un medio elástico, que, posteriormente vuelve a su estado original, debido a su propiedad de elasticidad. En el escenario que se está abordando, la vibración de partículas se transmite por el aire, creando una onda de presión sonora definida por la diferencia entre presión instantánea y la presión atmosférica.

Las ondas sonoras en el aire son ondas longitudinales porque las partículas del medio a través del cual se transporta el sonido vibran en paralelo a la dirección en la que se mueve la onda sonora. El resultado de tales vibraciones longitudinales es la creación de compresiones y rarefacciones dentro del aire.

La onda de presión sonora se va degradando y sufriendo diversas atenuaciones en su propagación por un ambiente real. Ya que la propagación del sonido cerca del suelo y al aire libre implica la dispersión geométrica, la absorción del aire, la interacción con el suelo, las barreras, la vegetación y la refracción asociada con el viento y los gradientes de temperatura.

Estas atenuaciones se van a estudiar en los apartados siguientes gracias a la norma de la ISO 9613.

3.1 LA ATENUACIÓN DEL SONIDO: ISO 9613

Con el objetivo de realizar el estudio de la propagación del sonido generado por drones en ambientes reales de forma que se pueda caracterizar la atenuación en función de la distancia para distintas bandas de frecuencia, se va a hacer referencia a la norma ISO 9613.

Las normas ISO son documentos que especifican requerimientos que pueden ser empleados en organizaciones para garantizar que los productos y/o servicios ofrecidos por dichas organizaciones cumplen con su objetivo. Las normas ISO son establecidas por el Organismo Internacional de Normalización (ISO), y se componen de estándares y guías relacionados con sistemas y herramientas específicas de gestión aplicables en cualquier tipo de organización. El objetivo perseguido por estas normas es asegurar que los productos y/o servicios alcanzan la calidad deseada [19].

De este modo, se va a utilizar el estándar 9613 de la ISO “Atenuación del sonido durante la propagación en exteriores” para poder abordar el estudio de la propagación del sonido en ambientes reales y así poder calcular la atenuación que sufrirá la onda sonora dependiendo de la distancia a la que se encuentre del array de micrófonos y de las bandas de frecuencia a las que emita. Estas bandas de frecuencia serán la frecuencia fundamental junto a sus armónicos.

El método especificado consiste en algoritmos de banda de octava, con frecuencias centrales entre 63 Hz y 8 kHz, para calcular la atenuación de sonido el cual se origina a partir de una fuente puntual [20]. Un algoritmo de banda de octava consiste en una banda de frecuencia donde la frecuencia más alta es el doble que la frecuencia más baja, además, las frecuencias van incrementando en octavas. Esto se realiza para poder obtener más información de la señal que se está analizando.

Capítulo 3 – Propagación del sonido en ambientes reales

Para poder caracterizar la atenuación del sonido se van a estudiar los siguientes efectos:

- Divergencia geométrica
- Absorción atmosférica
- Efecto del suelo
- Reflexiones de superficies
- Apantallamiento por obstáculos

Con el objetivo de poder caracterizar dichos efectos mencionados, en primer lugar, se hace necesario conocer ciertos parámetros con respecto a la geometría de la fuente y del ambiente, las características de la superficie del suelo y de la fuerza de la fuente en términos de niveles de presión sonora en bandas de octava.

De este modo, el nivel de presión sonora continuo equivalente viene dado por la siguiente ecuación.

$$L_{AT}(dB) = 10 \log_{10} \left\{ \frac{1}{T} \int_0^T \frac{p_A^2(t) dt}{p_0^2} \right\} \quad (23)$$

En la ecuación (23), $p_A(t)$ es la presión sonora en pascales ponderada la atenuación por banda de octava instantánea, p_0 es la presión sonora de referencia ($20 \cdot 10^{-6} Pa$) y T es el intervalo de tiempo en segundos especificado.

Por otro lado, el nivel de presión sonora continuo equivalente por bandas de octava a favor del viento (*downwind*) viene definido por la ecuación (24), donde p_f es la presión sonora instantánea por bandas de octava a favor del viento.

$$L_{ft}(DW)_{dB} = 10 \log_{10} \left\{ \frac{1}{T} \int_0^T \frac{p_f^2(t) dt}{p_0^2} \right\} \quad (24)$$

Por último, se define la pérdida por inserción, es decir, las pérdidas de una barrera, que se puede calcular como la diferencia entre el nivel de presión sonora en el receptor en una posición específica con la barrera y sin ella.

Las condiciones meteorológicas que se van a tener en cuenta en el ambiente para caracterizar la atenuación del sonido son:

- Condiciones de propagación a favor del viento: viento de $\pm 45^\circ$ de la dirección de propagación, conteniendo el centro de la fuente sonora y el viento soplando desde la fuente hacia el receptor [20].
- Velocidad del viento aproximadamente entre 1 y 5 m/s medido a una altura de 3 a 11 metros sobre el suelo [20].

Una vez definidos los aspectos anteriores, en la ISO 9613 se procede a definir las ecuaciones básicas que se van a utilizar para lograr caracterizar la propagación del sonido en un ambiente real.

El nivel de presión sonora continuo equivalente por bandas de octava a favor del viento $L_{ft}(DW)$ se calcula para cada fuente puntual y sus fuentes imagen y por cada banda de octava con la

frecuencia nominal desde 63 Hz hasta 8 kHz [20]. Esta presión sonora viene definida en la ecuación (25).

$$L_{ft}(DW) = L_W + D_c - A \quad (25)$$

Siendo:

- L_W el nivel de potencia sonora por bandas de octava producido por la fuente sonora puntual relativo a una potencia sonora de referencia de 1 pW [20].
- D_c la corrección por directividad, que describe el grado en que el nivel de presión sonora equivalente de una fuente puntual se desvía en una dirección determinada respecto del nivel sonoro de una fuente puntual omnidireccional [20].
- A es la atenuación por bandas de octava que se produce durante la propagación desde la fuente hasta el receptor. Este factor de atenuación es el que se va a seleccionar para poder obtener la atenuación del ambiente real [20].

En el caso que se está abordando en este proyecto, se supone un escenario con una única fuente, por lo que el factor de corrección por directividad se establece a cero.

$$D_c = 0 \text{ dB} \quad (26)$$

Por otra parte, en lugar del nivel de presión sonora continuo equivalente por bandas de octava a favor del viento ($L_{ft}(DW)$), se necesita el término promedio del nivel de presión sonora ponderado $L_{AT}(LT)$ cuando el intervalo de tiempo es del orden de meses o años. Esto se debe a que este intervalo extenso de tiempo incluirá cambios en las condiciones meteorológicas. Sin embargo, en el caso que se está abordando, se está considerando que se capta el sonido de un dron que está en vuelo tan solo unos instantes de tiempo, por lo tanto, esta especificación no sería necesaria, y con $L_{ft}(DW)$ será suficiente.

Una vez caracterizada la fuente, las potencias y sus definiciones básicas, se procede a obtener el cálculo de los efectos que producen la atenuación nombrados al principio de esta sección. La atenuación en ambientes reales para la propagación de una onda de sonido viene definida por la siguiente ecuación, cuyos términos se definen en los siguientes apartados:

$$A = A_{div} + A_{atm} + A_{gr} + A_{bar} + A_{misc} \quad (27)$$

3.1.1 Divergencia geométrica (A_{div})

El sonido emitido por un dron es el sonido de una fuente omnidireccional, ya que se propaga en todas las direcciones posibles, por lo que las ondas que se emiten son ondas esféricas cuyo nivel de presión sonora coincide a la misma distancia de la fuente. A medida que esta distancia aumenta, la energía del sonido se distribuye por una superficie cada vez mayor, de forma que cada vez que se duplica esta distancia, el nivel de presión sonora disminuye un factor de 6 dB teóricamente.

$$A_{div}(dB) = 20 \log_{10} \left(\frac{d}{d_0} \right) + 11 \quad (28)$$

La ecuación (28) responde a la atenuación por divergencia geométrica, donde d es la distancia desde la fuente al receptor en metros y d_0 es la distancia de referencia cuyo valor es 1 metro.

3.1.2 Absorción atmosférica (A_{atm})

La absorción atmosférica es la atenuación debida al nitrógeno, oxígeno y dióxido de carbono durante la propagación del sonido a medida que recorre una distancia concreta hasta el receptor.

$$A_{atm}(dB) = \alpha d \quad (29)$$

En la ecuación (29), α (dB/km) es el coeficiente de atenuación atmosférica para cada banda de octava. Este coeficiente depende de los siguientes parámetros: la frecuencia del sonido, la temperatura atmosférica ambiente, la humedad relativa del aire y la presión ambiente [20].

A) Frecuencia del sonido

En la mayor parte de los casos, el espectro de un sonido cubre un amplio rango de frecuencias, por lo que se realiza un análisis espectral por filtros de bandas de octava. Sin embargo, para poder realizar este barrido de frecuencias, primero es necesario remontarse al sonido de un tono puro.

Un tono puro es una onda sinusoidal, es decir, un sonido de una frecuencia única en particular. Si un tono puro se propaga a través de la atmósfera sobre una distancia s , la amplitud de la presión sonora p_t disminuye exponencialmente, desde el valor inicial p_i , como resultado del efecto de la absorción atmosférica. Esta absorción se produce de acuerdo con la fórmula de decaimiento para ondas planas en espacio libre.

$$P_t = p_i e^{\frac{-1}{10 \log_{10}(e_2)} \alpha s} \quad (30)$$

Por consiguiente, la atenuación debida a la absorción atmosférica entre el nivel de presión sonora de un tono puro con frecuencia f , a partir del valor inicial a $s = 0$, viene dado por:

$$\delta L_t(f) = 10 \log_{10} \left(\frac{p_{i_2}}{p_{t_2}} \right) dB = \alpha s \quad (31)$$

B) Temperatura atmosférica ambiente

La temperatura de referencia del aire será de 293,15 °K (20°C). Para calcular el nivel de ruido ambiental, el coeficiente de atenuación atmosférica se debe basar en valores promedio determinados por el rango del tiempo ambiental.

C) Humedad relativa del aire

La absorción atmosférica es sensible a la composición del aire, particularmente a la amplia variación de concentración de vapor de agua. En altitudes de al menos 50 km sobre el nivel del mar, el aire seco no tiene influencia en la atenuación por absorción atmosférica; sin embargo, en alturas hasta 10 km, la concentración molar del vapor de agua tiene una influencia considerable y varía ampliamente, debido básicamente a los procesos de combustión además de la generación de ozono. En el caso que se está abordando, se va a calcular la atenuación hasta 400 metros de altura, por lo que se hace necesario estudiar este efecto.

Con el objetivo de calcular la concentración molar del vapor de agua, se hace necesario conocer el volumen del vapor de agua, y para ello se introduce el concepto del diagrama psicrométrico. El diagrama psicrométrico del vapor de agua es un gráfico utilizado como instrumento de cálculo de forma que relaciona ciertos parámetros para poder calcular la humedad del aire en una estancia.

Capítulo 3 – Propagación del sonido en ambientes reales

El diagrama psicrométrico relaciona los siguientes parámetros: temperatura de bulbo seco, temperatura de bulbo húmedo, temperatura de punto de rocío, humedad absoluta y humedad relativa. De este modo, si se conocen dos de estos cinco parámetros se pueden conocer las demás especificaciones del vapor de agua en el aire.

En este diagrama se representa la temperatura como abscisa y la humedad absoluta en ordenadas calculadas relativa del 100% o curva de saturación, en donde se obtiene la humedad del aire saturado en función de su temperatura [21].

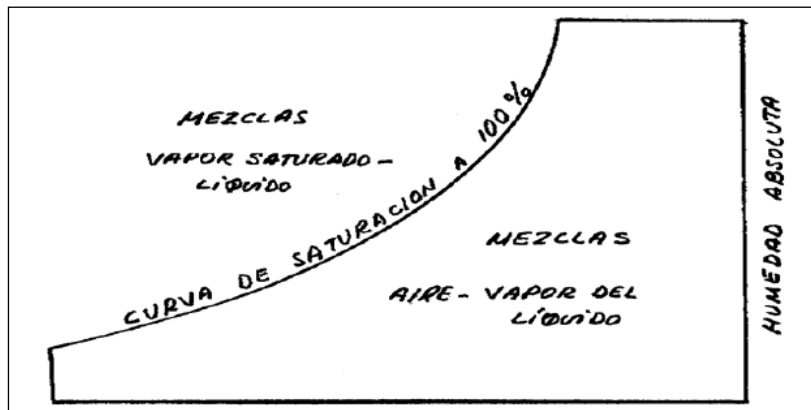


Figura 32 – Localización de las distintas zonas en un diagrama psicrométrico [21].

Los puntos localizados a la izquierda de la curva de saturación a 100% representan mezclas de aire saturado con agua en estado líquido, lo que provoca bruma y son condiciones muy inestables. Cualquier punto localizado a la derecha de la curva de saturación, representa una mezcla bien definida de aire y vapor de agua [21].

Si, por ejemplo, se evalúa el caso de la siguiente figura en el que la temperatura es de 20 °C y la humedad relativa del 40 %, se obtiene, con ayuda del diagrama psicrométrico, un volumen del vapor de agua de $v = 0.838 \frac{m^3}{kg}$ (d. a).

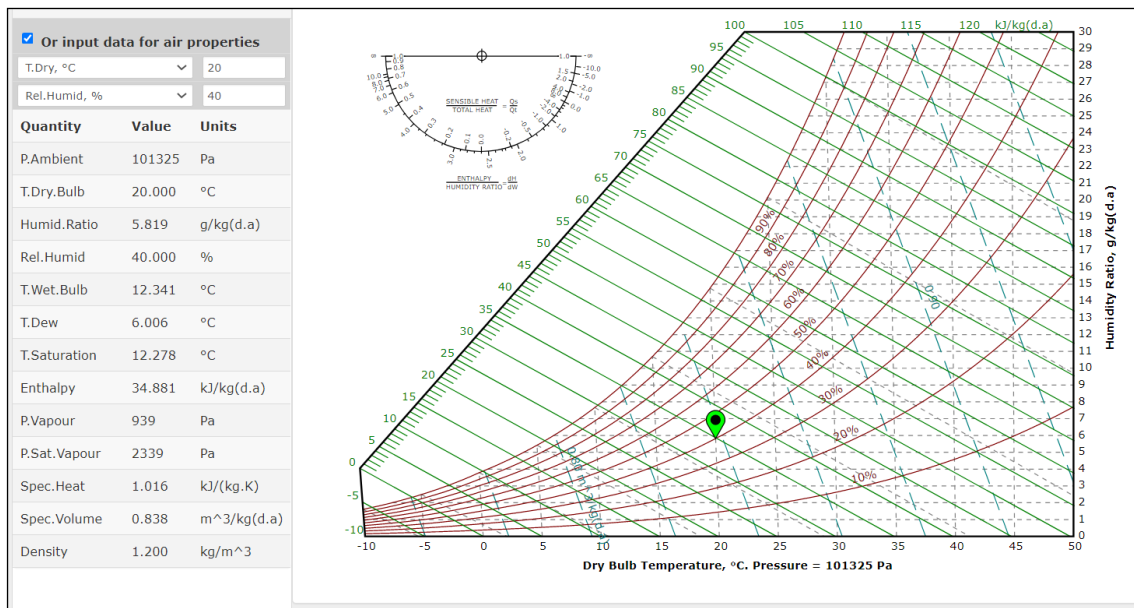


Figura 33 – Diagrama Psicrométrico a 20°C y 40% de humedad [22].

Una vez obtenido el volumen del vapor de agua, se calcula la presión parcial del vapor de agua, y finalmente se obtiene la concentración molar del vapor de agua como la presión parcial entre la presión atmosférica.

$$\text{Presión parcial del vapor de agua } \left(P_v = \frac{R_v T}{v} \right) \quad (32)$$

Siendo R_v la constante de vapor de agua, T la temperatura en grados Kelvin y v el volumen del vapor de agua calculado mediante el diagrama psicrométrico.

$$\text{Concentración molar del vapor de agua } \left(h = \frac{P_v}{P_a} \right) \quad (33)$$

En esta última ecuación P_a hace referencia a la presión atmosférica ambiente.

D) Presión atmosférica ambiente

La presión atmosférica ambiente de referencia p_r , se toma del Estándar Internacional de Atmósfera, que da un valor de 101,325 kPa.

Una vez calculados todos los parámetros del coeficiente de atenuación se muestra dicho coeficiente en forma de tablas para un rango de variables comúnmente encontrado en la predicción de propagación del sonido en exteriores. Este rango de variables se obtiene de la ISO 9613-1 y es el siguiente:

- Frecuencias desde 50 Hz hasta 10 kHz.
- Temperaturas desde -20°C hasta 100 °C.
- Humedad relativa desde 10% hasta 100%.
- Presión de 101,325 KPa.

		Coeficiente de absorción atmosférica α [dB/km]							
Temperatura	Humedad Relativa	Frecuencia central nominal [Hz]							
°C	%	63	125	250	500	1000	2000	4000	8000
10	70	0,1	0,4	1,0	1,9	3,7	9,7	32,8	117
20	70	0,1	0,3	1,1	2,8	5,0	9,0	22,9	76,6
30	70	0,1	0,3	1,0	3,1	7,4	12,7	23,1	59,3
15	20	0,3	0,6	1,2	2,7	8,2	28,2	88,8	202
15	50	0,1	0,5	1,2	2,2	4,2	10,8	36,2	129
15	80	0,1	0,3	1,1	2,4	4,1	8,3	23,7	82,8

Figura 34 – Rango de valores del coeficiente de absorción atmosférica [20].

Sin embargo, para poder calcular la atenuación por absorción atmosférica en *MATLAB*, se van a usar las siguientes fórmulas obtenidas de la norma ISO 9613-2, de forma que se puedan modificar las variables de entrada y se pueda conseguir el resultado del coeficiente de atenuación exacto.

En primer lugar, se calcula la frecuencia de relajación del oxígeno (34) y del nitrógeno (35). La frecuencia de relajación es la frecuencia a la que un elemento se mantiene en estado de equilibrio final.

$$f_{rO} = \frac{p_a}{p_r} \left(24 + \left(4,04 \cdot 10^4 h \left(\frac{0,02+h}{0,391+h} \right) \right) \right) \quad (34)$$

$$f_{rN} = \frac{p_a}{p_r} \left(\frac{T}{T_0} \right)^{-\frac{1}{2}} \left(9 + \left(280 h e^{-4,17 \left(\left(\frac{T}{T_0} \right)^{-\frac{1}{3}} - 1 \right)} \right) \right) \quad (35)$$

Por último, el coeficiente de atenuación viene expresado como:

$$\alpha \left(\text{dB}/\text{km} \right) = 8,686 f^2 \cdot \left(\left(1,84 \cdot 10^{-11} \cdot \left(\frac{p_a}{p_r} \right)^{-1} \cdot \left(\frac{T}{T_0} \right)^{\frac{1}{2}} \right) + \left(\frac{T}{T_0} \right)^{-\frac{5}{2}} \cdot \left(0,01275 \cdot e^{-\frac{2239,1}{T}} \cdot \left(f_{rO} + \frac{f^2}{f_{rO}} \right)^{-1} + 0,1068 \cdot e^{-\frac{3352}{T}} \cdot \left(f_{rN} + \frac{f^2}{f_{rN}} \right)^{-1} \right) \right) \quad (36)$$

Por consiguiente, ya se puede calcular la atenuación por absorción atmosférica como el producto del coeficiente de atenuación con la distancia recorrida por la onda sonora, reflejado en la ecuación (29).

3.1.3 Efecto del suelo (A_{gr})

La atenuación por efecto del suelo es debida, principalmente, al sonido reflejado por la superficie del suelo que interfiere con la propagación del sonido directamente desde la fuente al receptor.

La atenuación se produce cuando la fuente o el receptor están cercanos a la superficie del suelo. El método de cálculo que se detalla es aplicable solamente al suelo plano, horizontal o con una pendiente constante [20].

Para poder caracterizar la atenuación del sonido debido al efecto del suelo, se van a diferenciar tres regiones distintas:

- **REGIÓN DE LA FUENTE.** Esta región se extiende una distancia desde la fuente hasta 30 veces la altura de la fuente (h_s), con un valor máximo de d_p (distancia transmisor-receptor proyectada en el suelo).
- **REGIÓN DEL RECEPTOR.** Es la región desde el receptor hasta 30 veces la altura del receptor (h_r), con un valor máximo también de d_p .
- **REGIÓN MEDIA.** Distancia entre la región de la fuente y del receptor. En el caso de $d_p < (30h_s + 30h_r)$, las regiones de la fuente y del receptor estarán sobreestimadas y no existirá región media [20].

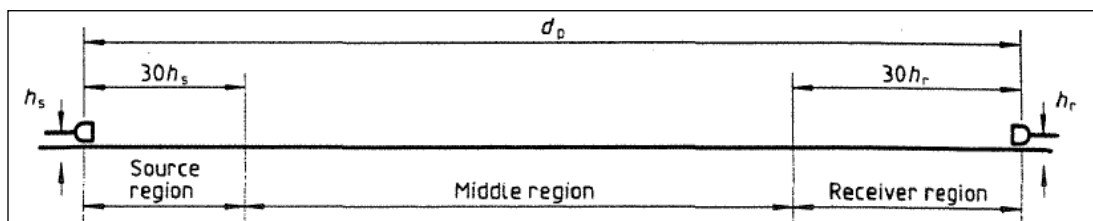


Figura 35 – Regiones de atenuación por efecto del suelo [20].

Capítulo 3 – Propagación del sonido en ambientes reales

Las propiedades acústicas de cada región de suelo son tomadas en cuenta a través de un factor G , existiendo tres categorías de superficies reflectantes [20]:

- Suelo duro ($G = 0$). Por ejemplo: pavimento, agua o hielo
- Suelo poroso ($G = 1$). Como superficies de vegetación o cultivo.
- Suelo mixto ($0 < G < 1$).

Una vez especificadas estas propiedades, ya se puede calcular la atenuación por efecto del suelo como la suma de tres variables de atenuación que dependen de las regiones especificadas anteriormente.

$$A_{gr} = A_s + A_r + A_m \quad (37)$$

De esta forma, la atenuación por efecto del suelo es equivalente a la suma de las atenuaciones de las regiones de suelo duro, poroso y mixto presentes en el escenario a analizar. En la siguiente tabla se muestra la forma de calcular estas variables para unas bandas de frecuencia centradas en frecuencias de mayor uso.

Frecuencia central nominal [Hz]	A_s o $A_r^{1)}$ dB	A_m dB
63	-1,5	$-3q^{2)}$
125	$-1,5 + G \times a'(h)$	$-3q (1 - G_m)$
250	$-1,5 + G \times b'(h)$	
500	$-1,5 + G \times c'(h)$	
1000	$-1,5 + G \times d'(h)$	
2000	$-1,5 (1 - G)$	
4000	$-1,5 (1 - G)$	
8000	$-1,5 (1 - G)$	
NOTAS $a'(h) = 1,5 + 3,0 e^{-0,12(h-5)h^{-5}} (1 - e^{-dp/50}) + 5,7 e^{-0,09 h h} (1 - e^{-2,8 \times 10^{-6} \times dp \times dp})$ $b'(h) = 1,5 + 8,6 e^{-0,09 h h} (1 - e^{-dp/50})$ $c'(h) = 1,5 + 14,0 e^{-0,9 h h} (1 - e^{-dp/50})$ $d'(h) = 1,5 + 5,0 e^{-0,9 h h} (1 - e^{-dp/50})$		

Figura 36 – Cálculo de variables para el cálculo del efecto por suelo para determinadas frecuencias [20].

Además, en la norma ISO 9613 se especifican dos aspectos para poder realizar los cálculos incluidos en la tabla de la Figura 36.

- Para calcular la atenuación de la región de la fuente, se toman los datos del suelo y altura de la fuente y lo mismo para el caso del receptor.
- Con relación a la región media, la variable q puede tomar varios valores:

$$\begin{cases} q = 0, & d_p < 30(h_s + h_r) \\ q = 1 - \frac{30(h_s + h_r)}{d_p}, & d_p > 30(h_s + h_r) \end{cases} \quad (38)$$

Por otro lado, aunque existe otra ecuación (39) que permite obtener la atenuación por efecto del suelo de forma más sencilla, su funcionamiento está especificado únicamente para distancias

largas y con superficie porosa o mixta, pero no suelo duro, como puede ser, en el caso que se está analizando, un aeropuerto.

$$A_{gr}(dB) = 4,8 - \left(2 \frac{h_m}{d}\right) \cdot \left[17 + \left(\frac{300}{d}\right)\right] > 0 \quad (39)$$

En esta ecuación, h_m sería la altura media del camino de propagación sobre el suelo en metros y d la distancia desde la fuente al receptor, también en metros.

Sin embargo, podría darse el caso de analizar un escenario donde sería válido usar la ecuación (39), es decir, un tono que tuviera más de una frecuencia propagándose por suelo en su mayoría poroso y con una distancia extensa entre el receptor y la fuente. En este caso, habría que incluir en el factor de corrección por directividad, el término D_0 para considerar el aparente incremento de atenuación en el nivel de potencia sonora de la fuente debido a reflexiones desde el suelo cercano a la fuente [20].

$$D_0(dB) = 10 \log_{10} \left(1 + \frac{d_p^2 + (h_s - h_r)^2}{d_p^2 + (h_s + h_r)^2}\right) \quad (40)$$

3.1.4 Apantallamiento por efecto de barreras (A_{bar})

Un objeto debe ser tomado en cuenta como un obstáculo de apantallamiento (barrera) si: cumple con una densidad superficial de al menos 10 kg/m^2 , tiene una superficie cerrada sin grandes grietas o huecos y la dimensión horizontal del objeto perpendicular a la línea que une emisor-receptor es mayor que la longitud de onda de la frecuencia central de interés λ [20].

En la ISO 9613 se calcula la atenuación por apantallamiento mediante las pérdidas de inserción.

En el caso de tener la propagación del sonido a favor del viento, se calcula la difracción en el borde del obstáculo como:

$$A_{bar}(dB) = D_z - A_{gr} > 0 \quad (41)$$

Y, por otro lado, la difracción que se obtiene alrededor del obstáculo viene dada por:

$$A_{bar}(dB) = D_z > 0 \quad (42)$$

Siendo D_z la atenuación para cada banda de octava. Con el objetivo de calcular el apantallamiento, se va a asumir un solo camino de propagación del sonido, para poder facilitar los cálculos. Entonces, la ecuación de la atenuación para cada banda de octava es la definida en (43).

$$D_z(dB) = 10 \log_{10} \left(3 + \frac{C_2}{\lambda} C_3 z K_{met}\right) \quad (43)$$

En la ecuación (43):

- $C_2 = 20$ e incluye el efecto de reflexión del suelo.
- $C_3 = \begin{cases} 1, & \text{para difracción simple} \\ 1 + \frac{5\lambda}{e^2} \frac{1}{3} + \frac{5\lambda}{e^2}, & \text{para difracción múltiple} \end{cases} \quad (44)$
- λ (m) es la longitud de onda de la frecuencia central de cada banda de octava.

- z es la diferencia entre los caminos directo y difractado.

$$z = \begin{cases} [(d_{ss} + d_{sr})^2 + a^2] - d, & \text{si difracción única} \\ [(d_{ss} + d_{sr} + e)^2 + a^2]^{\frac{1}{2}} - d, & \text{si difracción doble} \end{cases} \quad (45)$$

Donde, $d_{ss}(m)$ es la distancia desde el emisor hasta el primer borde del obstáculo, $d_{sr}(m)$ es la distancia desde el segundo borde del obstáculo hasta el receptor y $a(m)$ es la componente de distancia, paralela al borde de la barrera entre la fuente y el receptor [20].

- K_{met} es el factor de corrección para efectos meteorológicos.
- e es la distancia entre los bordes de difracción del obstáculo en el caso de haber difracción doble.

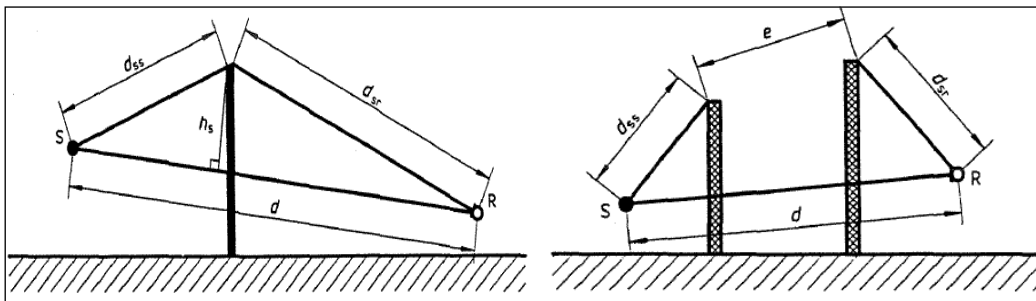


Figura 37 – Efecto por apantallamiento [20]: a) Difracción simple; b) Difracción doble.

Por último, en el caso de que hubiera más de dos barreras produciendo difracción en el camino fuente-receptor, se escogerían las dos barreras más predominantes, es decir, que produzcan más pérdidas para así considerar el caso peor y despreciando los efectos de las demás.

3.1.5 Otros efectos similares (A_{misc})

Por último, en la ecuación (27) aparece la atenuación por otros efectos similares, que lo constituyen:

- **ATENUACIÓN POR FOLLAJE (A_{fol})**. Este tipo de atenuación es mínimo y solo se considera si la vegetación es muy densa e interrumpe el camino directo desde la fuente al receptor. Asimismo, solo se considera la parte de vegetación cerca de la fuente y cerca del receptor, pero no la de la mitad del camino.

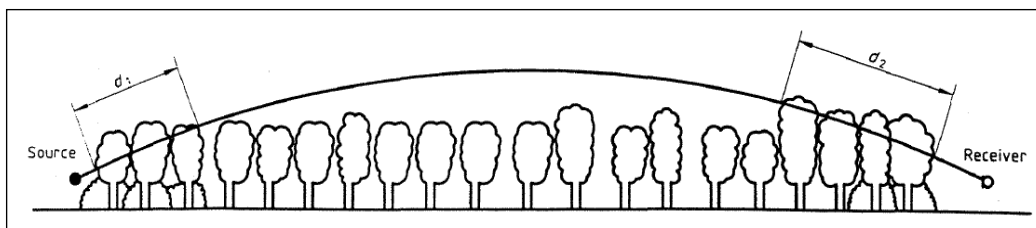


Figura 38 – Representación de atenuación por follaje [20].

- **ATENUACIÓN POR SITIOS INDUSTRIALES (A_{site}).** Este tipo de atenuación ocurre debido a la dispersión de instalaciones u otros objetos en zonas industriales. Este tipo de atenuación depende completamente del tipo de sitio industrial por el que se esté propagando el sonido. Además, la atenuación incrementa linealmente con la longitud del trayecto que sigue la onda a través de la zona industrializada (d_s).

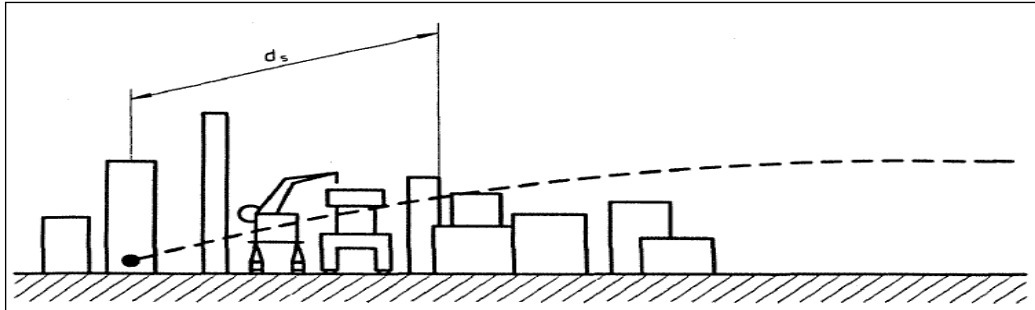


Figura 39 – Representación de atenuación por sitios industriales [20].

- **ATENUACIÓN POR VIVIENDAS (A_{hous}).** La atenuación por viviendas sucede cuando la fuente, el receptor o bien ambos están localizados en una zona urbanizada, al ocurrir debido a la proyección de los edificios. Sin embargo, este efecto se compensa en gran medida por la propagación entre casas y las reflexiones de otros edificios de alrededor, ya que, al ser las fachadas verticales, pueden incrementar el nivel de presión sonora en el receptor. La atenuación por efecto de edificios es totalmente dependiente de la zona y los edificios que hay alrededor, por lo que no se puede dar un valor exacto del efecto que vaya a tener esta atenuación en el total, aunque sí se puede confirmar que dicho efecto no va a ser mayor de 10 dB.

3.2 ANÁLISIS MEDIO DE PROPAGACIÓN PARA UN UAV

El escenario que se va a tener en cuenta para la propagación del sonido es el de un aeropuerto. Además, el UAV en cuestión va a volar hasta una distancia de 400 metros del array de micrófonos. Así, se limitan las características atenuantes anteriores según lo siguiente:

- La mayor parte de la atenuación vendrá dada por la divergencia geométrica, debido a que, las ondas esféricas del dron van disminuyendo hasta 6 dB cada vez que se duplica la distancia desde la fuente.
- La absorción atmosférica no tendrá gran peso en la atenuación del sonido, ya que se están suponiendo condiciones de temperatura de 20 °C con un 40 % de humedad y una presión atmosférica ambiente de 101,325 kPa.
- La atenuación por efecto del suelo será nula. Esto se debe a que el suelo de un aeropuerto es duro, por lo que el factor $G = 0$. Entonces, $G_s = G_r = 0$, por lo que en la Figura 36, $A_s = A_r = -1,5 \text{ dB}$.

Por otro lado, habría que calcular A_m (atenuación en la región media) cuyo cálculo implica lo siguiente:

$$A_m(\text{dB}) = -3q(1 - G_m) \quad (46)$$

Capítulo 3 – Propagación del sonido en ambientes reales

Donde G_m es nulo debido al suelo duro y, además, q también sería nulo, debido a que el array de micrófonos está en el suelo y el dron en el aire, por lo que siempre la relación de la ecuación (38) dará lugar a $q = 0$.

Sin embargo, una atenuación no puede ser negativa, ya que estaría dando lugar a una falsa ganancia del medio de propagación, por lo que, se deja a cero.

- Al suponerse este tipo de escenario, no habrá ningún tipo de apantallamiento por barreras, ya que, el objetivo de este trabajo es localizar acústicamente a los drones para que no interfieran en el despegue y aterrizaje de los aviones, por lo que la detección se hará en una etapa previa.
- Por último, en la atenuación por efectos diversos, se obtendrá una atenuación nula al estar en un escenario aeroportuario.

De este modo se calcula la atenuación en función de la distancia y de la frecuencia para los tres tipos de drones en cuestión: QAV400, Parrot AR y Harop Harpy.

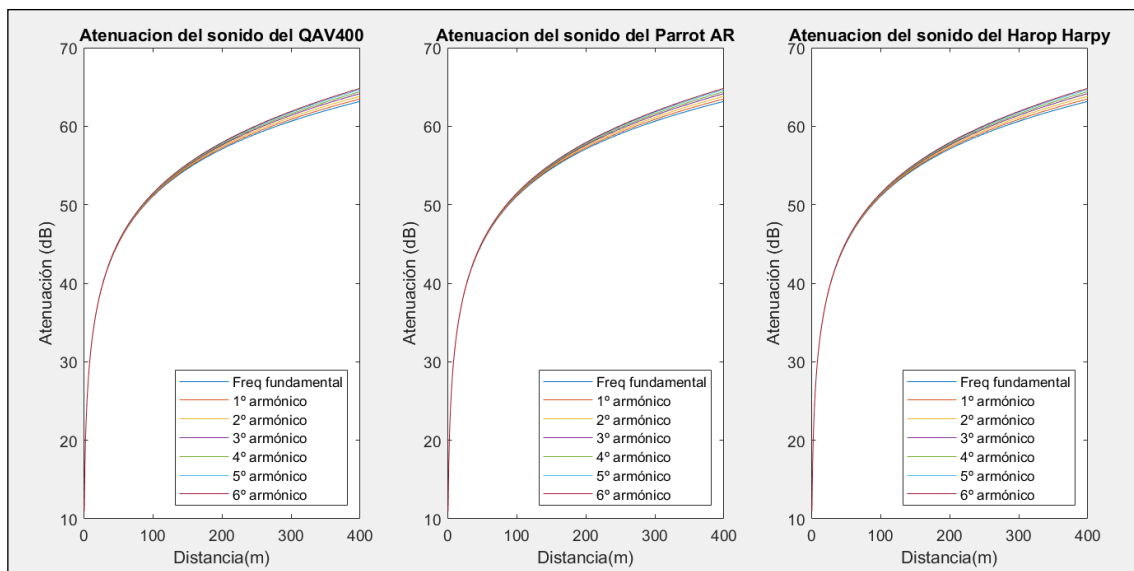


Figura 40 – Atenuación total en función de la frecuencia y distancia de los UAVs.

Como se puede comprobar en la Figura 40, la atenuación aumenta a la vez que la distancia entre fuente y receptor, esto se debe al efecto de la divergencia geométrica, ya que es la atenuación que mayor peso tiene en la propagación del sonido.

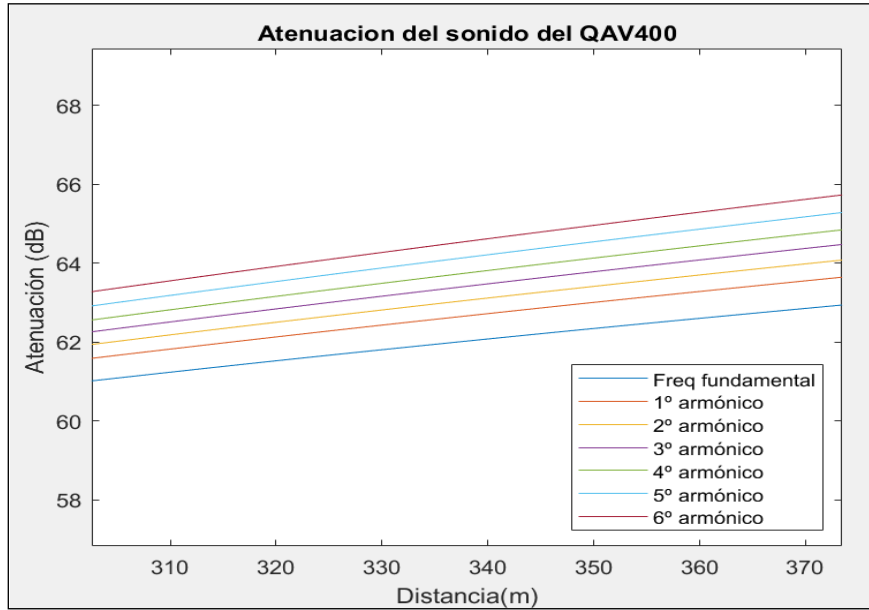


Figura 41 – Ampliación de la atenuación total del QAV400.

Por otro lado, en la Figura 41, se puede observar como la atenuación aumenta proporcionalmente con la frecuencia, por lo que, a frecuencias menores, la atenuación será menor. De este modo, la frecuencia fundamental será la frecuencia central óptima para la detección del dron, ya que además de ser la que menos atenuación presenta, es la que mayor energía posee, por lo que se detectara una señal con un valor de SNR mayor a las demás bandas.

4 RECEPTORES ACÚSTICOS

Una vez seleccionadas las bandas de frecuencia más favorables, se va a estudiar como detectar los drones acústicamente usando arrays de micrófonos. El objetivo final de este trabajo es encontrar un algoritmo mediante el cual el array de micrófonos pueda vigilar el espacio aéreo y detectar los pequeños UAVs, de forma que no interfieran con la actividad que se desarrolla en espacios como aeropuertos.

Sin embargo, en primer lugar, se debe abordar un estudio del receptor acústico o micrófono.

4.1 MICRÓFONO

Las ondas sonoras son ondas de presión sonora, provocadas al producir la vibración de un medio elástico, ya sea sólido, líquido o gaseoso, como el aire. La presión sonora se genera al variar la presión atmosférica mediante esta presión instantánea. Estas vibraciones generadas por la fuente sonora se convierten en ondas mecánicas y posteriormente en ondas eléctricas en el dispositivo receptor, en este caso, un micrófono.

Las ondas sonoras se propagan longitudinalmente y producen zonas de compresión de las partículas del medio y zonas de descompresión (rarefacción) de las partículas de transmisión.

4.1.1 Elementos del receptor acústico

Los elementos que debe tener un micrófono para poder realizar la conversión de ondas mecánicas a energía eléctrica y así obtener la señal para su posterior análisis son los siguientes:

- DIAFRAGMA. El diafragma es una membrana que capta las vibraciones mecánicas producidas por la fuente sonora. Debido a la fragilidad del diafragma, este debe estar protegido por una rejilla para evitar daños por golpes de sonido o físicos.
- TRANSDUCTOR ELECTROMECAÁNICO. El transductor es un dispositivo que realiza la transformación entre energía mecánica y energía eléctrica.
- ALIMENTACIÓN Y/O PREAMPLIFICADOR EXTERNO. Si se está usando un micrófono de condensador, al tener elementos electrónicos activos, se necesita de cierta alimentación. Además, si así se necesita, es posible el uso de un preamplificador externo para incrementar la amplitud de la señal antes de realizar la conversión de la señal analógica a digital.
- CONECTOR DE SALIDA. Para poder manejar la señal eléctrica y así llevarla al conversor analógico-digital, es posible o llevarla mediante un conector XLR, o, en modelos inalámbricos, mediante un transmisor de radiofrecuencia.
- EQUIPO GRABADOR. Este equipo realiza la conversión de la señal analógica a una señal digital para así poder realizar el análisis mediante la señal obtenida.

4.1.2 Características del receptor acústico

Un receptor acústico o micrófono dispone de las siguientes características, las cuales son imprescindibles para poder realizar la recepción de la señal sonora.

A) Sensibilidad

La sensibilidad de un micrófono indica la capacidad de dicho micrófono para captar señales de poca intensidad. Por ejemplo, si se dispone de un micrófono de condensador, que tiene una sensibilidad aproximada de -34 dB, se podrán recibir las señales superiores a dicho valor.

La sensibilidad se calcula de la siguiente forma:

$$S(\text{dB}) = 20 \log_{10} \left(\frac{V_{mic}}{P_a} \right) \quad (47)$$

Donde V_{mic} es la tensión de salida del micrófono en valores eficaces y P_a es la presión sonora a la entrada del micrófono en pascales.

B) Respuesta en frecuencia

La respuesta en frecuencia es la función que representa la variación de la sensibilidad con la frecuencia. La respuesta en frecuencia es la función que representa la respuesta del sistema ante entradas sinusoidales en un margen de frecuencia determinado. En un entorno ideal, se requeriría un micrófono cuya respuesta en frecuencia fuera homogénea para todo el rango de frecuencias de interés.

C) Impedancia de salida

Cualquier micrófono tiene siempre una impedancia eléctrica de salida. Esta es la que se conecta directamente con el preamplificador o con el elemento de grabación para poder transmitir la señal. Por otro lado, debido a que el micrófono debe usar una resistencia mucho menor a la resistencia del elemento receptor, existirá cierta adaptación de impedancias.

D) Ruido eléctrico y distorsión

Un receptor acústico no solo recibe la señal deseada, sino efectos adversos como reverberación, interferencias o ruido. Además, los micrófonos generan cierto ruido eléctrico y, también, debido a la falta de linealidad, se generan señales eléctricas indeseadas a la salida del micrófono.

E) Patrón de directividad

La directividad de un micrófono es la sensibilidad que tiene el micrófono al recibir ondas sonoras, además, es totalmente dependiente de la dirección de captación de la onda, puesto que el micrófono será más sensible a unas direcciones que a otras. La directividad se puede mostrar gracias a un diagrama polar, el cual consta de un lóbulo principal, un lóbulo posterior y lóbulos secundarios. El lóbulo principal es la dirección de máxima directividad, el posterior es el que se encuentra a 180 grados del principal, es decir, su inverso, y, por último, los lóbulos secundarios son todos los demás que aparecen en el patrón de directividad. Con el objetivo de ejemplificar el diagrama polar de la directividad, se considera la siguiente figura.

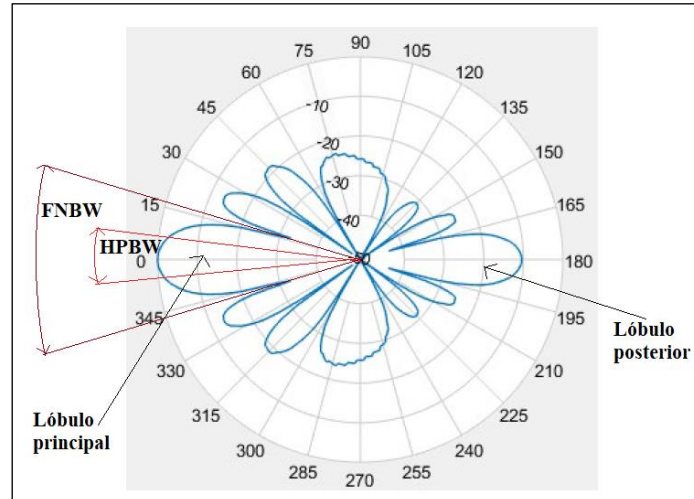


Figura 42 – Diagrama de radiación de una antena.

Este diagrama polar se denomina diagrama de radiación o patrón de directividad (*beam pattern*). En él, el lóbulo principal, es decir, donde el micrófono tiene mayor sensibilidad a las ondas que recibe, está en torno a 0 grados. Por otro lado, el lóbulo posterior, está en 180 grados, y, los lóbulos secundarios son todos los demás.

Por otro lado, en la Figura 42, aparecen dos parámetros: HPBW (*Half-Power Beam Width*), corresponde al ancho de haz a 3 dB, y mide la anchura de haz a la cual la ganancia del micrófono tiene máxima directividad, hasta que se reduce 3 dB del valor de ganancia máximo. Y, por otro lado, FNBW (*First Null Beam Width*) es el ancho de haz entre nulos, y mide la anchura del haz de directividad máxima hasta que llega a un ángulo donde la ganancia es mínima.

Azimut y elevación

El diagrama de radiación de un micrófono puede tener información tanto en azimut como en elevación. El ángulo en azimut es un corte horizontal y se refiere al ángulo en la horizontal de la dirección en la que se recibe una señal con respecto a un eje de referencia. El azimut tiene una cobertura de -180 a 180 grados. En cambio, la elevación es el ángulo que forma la dirección de llegada con el plano horizontal, y su valor varía entre -90 y 90 grados, y corresponde a un corte en vertical, informando sobre la altura en la que se está recibiendo la señal. El diagrama de radiación puede mostrar la directividad en azimut o en elevación; por ejemplo, si se quiere conocer el azimut, la elevación será nula y viceversa.

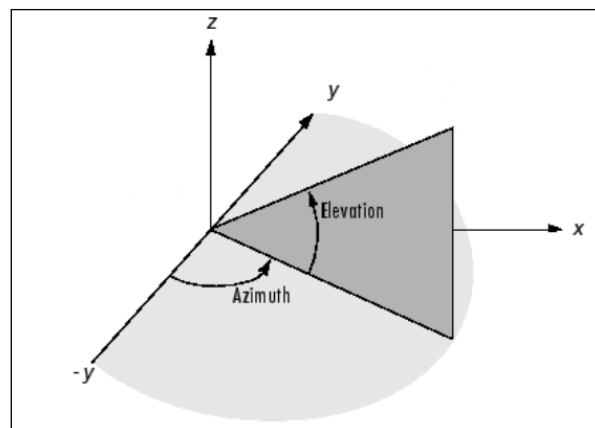


Figura 43 – Diagrama de azimut y elevación [24].

Por último, el patrón de directividad de un micrófono es similar al de un array de micrófonos.

4.2 ARRAYS DE MICRÓFONOS

Un array de micrófonos consiste en un conjunto de micrófonos posicionados de una forma en la que la información espacial sea bien capturada. La información disponible gracias a la diversidad necesita ser procesada. Luego, el principal objetivo del procesado de señal mediante un array de micrófonos es la estimación de algunos parámetros o la extracción de señales de interés, usando información espaciotemporal y frecuencial disponible a la salida del array de micrófonos [25].

Las propiedades que ofrecen estos arrays son [26]:

- Directividad espacial (selectividad espacial).
- Configuración geométrica (se trata en el siguiente apartado).
- Ganancia del array mayor que la de un único micrófono.
- Dependencia de la respuesta en frecuencia con el tamaño del array. Cuanto mayor sea la relación entre el receptor y la longitud de onda, mayor será la directividad.

4.2.1 Geometría de los arrays

Los micrófonos de un array pueden estar dispuestos en geometrías particulares, de forma que cada sensor está localizado en una posición con respecto al punto de referencia. La geometría del array de micrófonos depende de la aplicación que se vaya a llevar a cabo, por ejemplo, en los ordenadores o los móviles, la geometría es lineal para poder integrarla más fácilmente en el dispositivo [27]. Sin embargo, cuando la señal deseada puede venir de cualquier ángulo, la mejor geometría para el array es la circular, ya que el array lineal no tiene la misma respuesta a diferentes direcciones. En este documento, se han seleccionado tres geometrías de micrófonos distintas para evaluar posteriormente una serie de algoritmos que estiman el ángulo de llegada del dron, y así poder seleccionar la combinación que mejores resultados ofrezca.

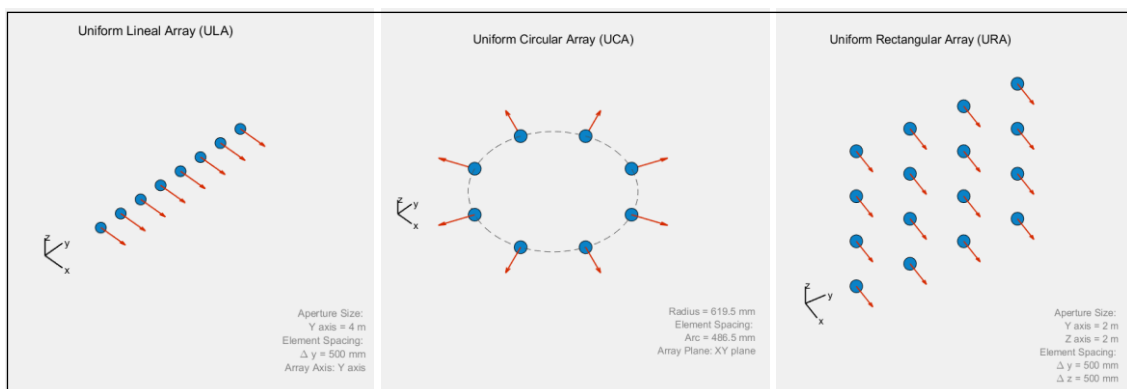


Figura 44 – Arrays de micrófonos con geometría uniforme y : a) Lineal; b) Circular; c) Rectangular.

Las combinaciones de la Figura 44 son respectivamente: array lineal uniforme (ULA) con ocho micrófonos, array circular uniforme circular (UCA) con ocho micrófonos y, array rectangular uniforme (URA) con 16 micrófonos.

Los patrones de directividad de los arrays tanto en elevación como en azimut para un array de, por ejemplo, ocho micrófonos con una distancia entre ellos de 0,5 metros son los siguientes:

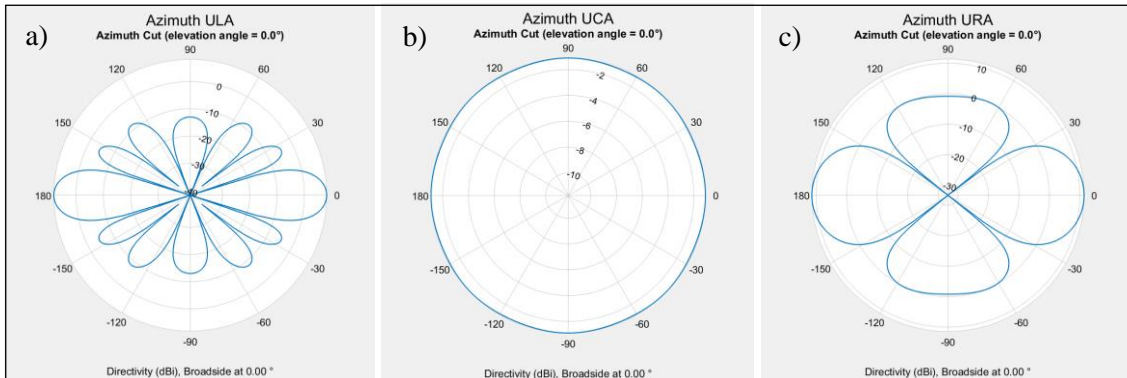


Figura 45 – Patrón de directividad en azimut de un array con geometría: a) ULA; b) UCA; c) URA.

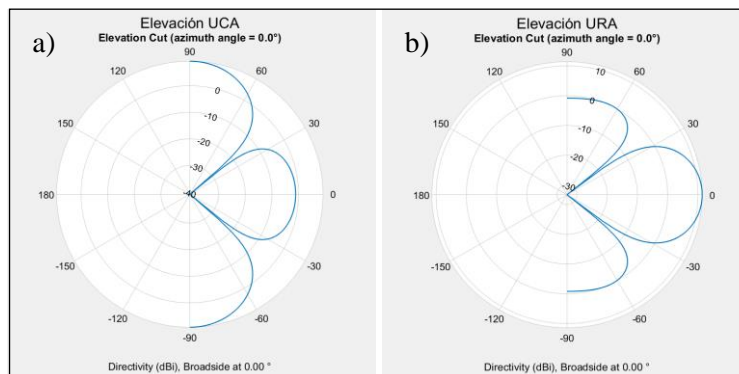


Figura 46 – Patrón de directividad en elevación de un array con geometría: a) UCA; b) URA

En relación con las figuras 45 y 46, se mencionan las siguientes características. En primer lugar, el array lineal uniforme (ULA) está formado por una serie de micrófonos idénticos situados en una línea, y, separados entre ellos una distancia d . El array lineal uniforme destaca por su simplicidad, directividad y por obtener el haz principal más estrecho en comparación con el resto de las geometrías. Sin embargo, un ULA no opera de la misma forma para todas las direcciones, su destreza y resolución son significativamente menores en el límite del rango del micrófono, por lo que no destaca su habilidad para localizar un objetivo sin saber a priori el ángulo de llegada. Además, un ULA no proporciona ninguna información acerca de la elevación, por tanto, no podrá dar información completa acerca del posicionamiento del dron en 3D.

Después, en las figuras 45 y 46, aparece el array circular uniforme (UCA), este tipo de array está formado por un cierto número de micrófonos idénticos M distribuidos uniformemente sobre una circunferencia de radio r y separados entre ellos una distancia d .

$$r = \frac{d/2}{\arcsen(\pi/M)} \quad (48)$$

La representación de los ángulos de llegada mediante este array UCA se realiza a través de un sistema de coordenadas esféricas, donde el origen de coordenadas es el centro del array. El patrón de directividad en azimut de este tipo de array es omnidireccional, es decir, tiene la misma directividad para cualquier dirección, por lo que le hace óptimo en circunstancias donde no se puede saber la dirección de la señal de llegada. Por otro lado, un array circular uniforme sí tiene información en elevación y su diagrama de radiación en elevación está compuesto por un lóbulo

principal y un cierto número de lóbulos secundarios, que dependen del número de micrófonos del array.

Por último, en las figuras 45 y 46 se encuentra el array rectangular uniforme (URA) que está formado por un cierto número de micrófonos idénticos distribuidos uniformemente formando un rectángulo con la misma distancia d entre ellos. Al igual que el caso anterior, se utiliza un sistema de coordenadas esféricas para representar las direcciones de las ondas incidentes. Un URA es parecido a una composición de arrays lineales, tanto de forma horizontal como vertical, es por ello por lo que puede ofrecer información tanto en azimut como en elevación. En ambos planos angulares el diagrama de radiación depende de la distancia entre los micrófonos y el número de micrófonos. En un caso común donde no existe aliasing, el patrón de directividad tiene un lóbulo principal y algún lóbulo secundario. Además, en el caso del ángulo en azimut, el patrón también tiene un lóbulo posterior, factor que puede afectar a la hora de localizar el dron sin saber previamente su posición. En un patrón de directividad, siempre es interesante que el lóbulo posterior sea en un gran orden menor que el principal, pero en este caso son de la misma magnitud, lo que puede afectar negativamente en la localización del UAV.

4.2.2 Características del array

El patrón de directividad de los arrays, además de la configuración en la que estén puestos los micrófonos, se ven modificados por componentes como la distancia entre micrófonos, el número de micrófonos y en el caso del array circular, el radio de dicho array. En el caso que se va a analizar en este proyecto, se van a variar las configuraciones de estas características, modificando:

- El número de micrófonos: Desde 4 hasta 16 en pasos de dos micrófonos.
- La distancia entre sensores: Desde 0,3 hasta 0,6 metros en pasos de 0,05 metros.

En primer lugar, se caracteriza el patrón de directividad de los micrófonos. Para ello, se va a mostrar el corte en azimut del array lineal uniforme con distintas configuraciones, ya que es el array de mayor simplicidad. Primero, se deja constante una distancia entre sensores de 0,3 metros y se muestra el patrón del array con 4, 8 y 16 micrófonos. Además, los análisis que se van a hacer son a la frecuencia fundamental del dron QAV400, es decir, de $f = 277,4566 \text{ Hz}$.

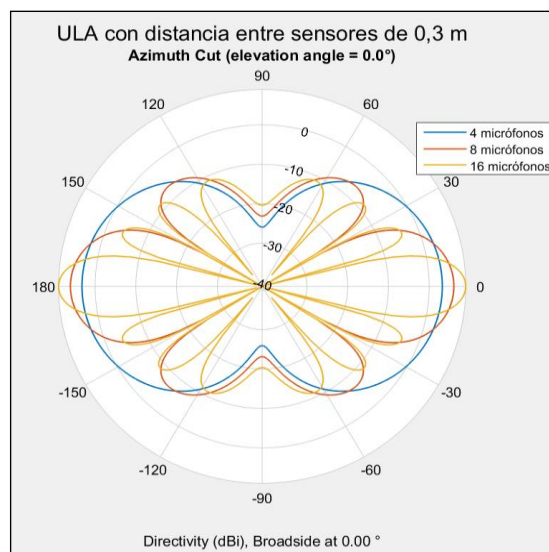


Figura 47 – Patrón de directividad de un array ULA al variar el número de micrófonos, y, con una distancia entre sensores de 0,3 metros.

En la Figura 47, según va aumentando el número de micrófonos dejando fija tanto la distancia entre ellos como la frecuencia, los lóbulos del patrón de directividad se van haciendo más estrechos. Es decir, el ancho de banda es menor y esto hace que el array sea más selectivo, tal y como representa la siguiente ecuación.

$$B \text{ (Hz)} = \frac{2c}{Ndf} \quad (49)$$

La ecuación anterior calcula el ancho de banda de un array de micrófonos (B) en hercios en función de la velocidad del sonido ($c = 343 \text{ m/s}$), el número de micrófonos N , la distancia entre sensores d y la frecuencia fundamental f . Asimismo, el array va aumentando en tamaño, debido a la siguiente relación:

$$L = Nd \quad (50)$$

Además, no se producirá aliasing (solapamiento), siempre y cuando la distancia entre sensores sea menor a la mitad de la longitud de onda del sonido recibido. Como $\lambda = \frac{c}{f}$, y $d < \frac{\lambda}{2}$, la distancia entre micrófonos no podrá exceder a 0,6 metros en ningún caso.

Por otro lado, si se deja fijo el número de micrófonos, por ejemplo, un array de 8 micrófonos, y se varía la distancia entre elementos del array, se obtienen los siguientes patrones de directividad.

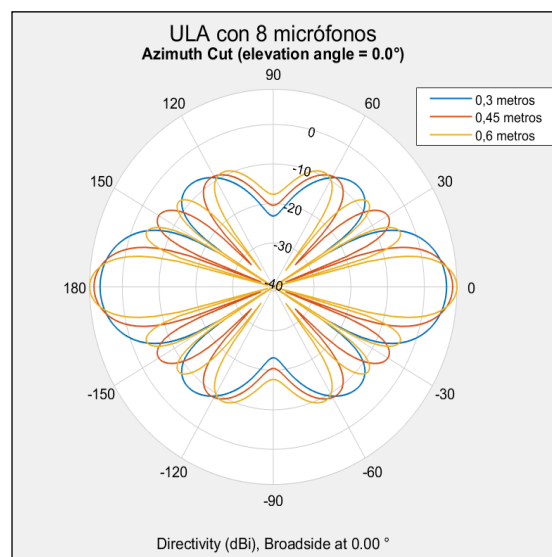


Figura 48 – Patrón de directividad de un array ULA al variar el número de distancia entre micrófonos, y, con un número de 8 sensores

Según la ecuación (49), el ancho de banda se va a hacer menor a medida que aumenta el número de micrófonos, dejando fija tanto la distancia como la frecuencia, por lo que el ancho de haz será menor y el array será más selectivo a mayor distancia entre elementos. Sin embargo, como la distancia entre elementos tiene que ser menor que la mitad de la longitud de onda, según va creciendo la distancia, aunque el array sea más directivo, cuando llegue a $d = \frac{\lambda}{2}$, se producirá solapamiento y la señal recibida estará corrupta, ya que existirán varios ángulos igual de directivos.

En cuanto a la modificación de la frecuencia, si la frecuencia creciera, a mayor frecuencia, el ancho de banda también aumentaría según la ecuación (49), además, el espacio entre micrófonos debe ser menor que un cierto umbral correspondiente a la mitad de la longitud de onda. Por lo

tanto, la frecuencia no podría aumentar indefinidamente, sino hasta el umbral dado por $\frac{\lambda}{2}$. En el caso de este documento, las frecuencias fundamentales de los drones UAV, son del orden de entre los 100 y 300 hercios, con lo cual esta restricción no nos afectaría.

En el entorno *MATLAB* donde se va a realizar la prueba de concepto, se van a utilizar micrófonos *phased.CustomMicrophoneElement* con una respuesta en frecuencia calculada por interpolación del vecino más cercano y que operará para valores frecuenciales en un rango de 0 a 20 kHz. Además, serán micrófonos omnidireccionales con una sensibilidad de 0 dB, es decir, no son nada directivos, por lo que captan la señal con la misma ganancia en cualquier ángulo.

5 DETECCIÓN DE ÁNGULOS DE LLEGADA

Una vez se ha definido qué es un array de micrófonos, así como las características y geometrías uniformes que existen, se va a explicar cómo detectar el sonido de un dron, es decir, el algoritmo necesario para detectar si la señal captada por los micrófonos proviene de un dron.

Para llevar a cabo la detección de los ángulos de llegada de una señal, existen dos tipos de procesados: procesado de señales de arrays de autoadaptación y estimación del espectro espacial. Mientras que el primer tipo se ha usado en varias aplicaciones prácticas, la estimación del espectro espacial fue desarrollado posteriormente y está aún en desarrollo [28].

Sin embargo, antes de entrar a explicar los distintos algoritmos de detección de llegada (DOA) que se han usado en este trabajo, se debe explicar las características y diferencias entre las señales de banda ancha y de banda estrecha. Debido a que muchos algoritmos de DOA se han diseñado para señales de banda estrecha, debe estudiarse la forma de pre-procesar las señales de banda ancha que puedan ser aplicadas a estos algoritmos de DOA.

5.1 SEÑALES DE BANDA ANCHA Y ESTRECHA

El ancho de banda de una señal es el rango de frecuencias donde la señal concentra la mayor potencia. Dependiendo de la relación entre el ancho de banda y la frecuencia central, las señales pueden ser de banda ancha o estrecha. Es decir, si la frecuencia central es mucho mayor que la mitad del ancho de banda, la señal será de banda estrecha $\Delta w \ll w_c$, mientras que, si la frecuencia central es del mismo orden que la mitad del ancho de banda $\Delta w \cong w_c$, la señal será de banda ancha.

En cuanto al periodograma, las señales de banda estrecha están caracterizadas por tener un espectro que concentra la energía en un margen muy estrecho de frecuencias. Por otro lado, las señales de banda ancha están caracterizadas por tener un espectro que ocupa un rango amplio de frecuencias.

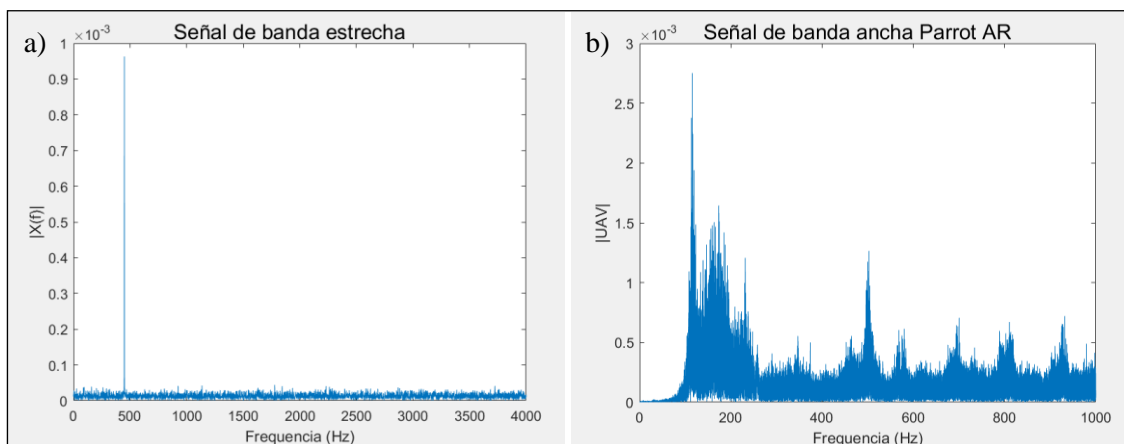


Figura 49 – Periodograma señal de: a) Banda estrecha; b) Banda ancha.

La señal acústica producida por un dron UAV es una señal de banda ancha como se evaluó en capítulos anteriores cuando se trataba el análisis frecuencial o el espectrograma. Sin embargo, la

mayoría de los algoritmos de estimación del ángulo de llegada, son algoritmos que deben recibir como entrada, una señal de banda estrecha. Esto se debe a que la dimensión del subespacio de una señal es mayor que el número de fuentes, por lo que los algoritmos de banda estrecha de alta resolución no funcionan bien con las señales de banda ancha.

A continuación, se va a explicar un método para adaptar una señal de banda ancha, para así poder usar los algoritmos DOA de alta resolución, que se explicarán más adelante. Para este objetivo, se va a utilizar la transformada de Fourier de tiempo reducido (STFT). En el capítulo 2.5 se explicó que esta técnica consiste en dividir la señal en segmentos cortos de tiempo de forma que la señal de entrada en cada segmento sea estacionaria, y, después, se calcula la transformada de Fourier de cada uno de ellos, así se puede obtener un análisis en el dominio del tiempo-frecuencia donde exista una frecuencia de mayor amplitud en cada instante de tiempo. Para poder calcular el análisis de Fourier en cada uno de los segmentos, se utiliza una función ventana, la cual puede ser cualquiera de las ventanas disponibles definidas en el capítulo 2.5 (rectangular, Hamming, Hanning o Blackman), que se seleccionará dependiendo de la señal de entrada y el objetivo que se esté persiguiendo.

Esta transformada STFT tiene la ecuación matemática descrita en la ecuación (11), aunque también se puede describir de la siguiente manera:

$$STFT\{x[n]\} = S(m, k) = \sum_{n=-\infty}^{\infty} (x[n] h[n - m]) e^{-jk\frac{2\pi}{N}n} \quad (51)$$

Donde $x[n]$ es la señal por analizar, $h[n]$ es la ventana de análisis, N es el orden de la DFT, k es el índice de la frecuencia (banda de frecuencia) y m es el punto donde se centra la ventana.

Al mismo tiempo, en la ecuación (51) se puede variar el orden de los productos obteniendo una convolución entre la señal de entrada modulada exponencialmente a cada frecuencia, y la ventana.

$$S(m, k) = \sum_{n=-\infty}^{\infty} (x[n] e^{-jk\frac{2\pi}{N}n}) h[n - m] \quad (52)$$

Si en la ecuación (52) se denomina $v[n]$ a la señal $x[n] e^{-jk\frac{2\pi}{N}n}$, obtenemos la siguiente relación debido a la definición de la convolución:

$$S(m, k) = \sum_{n=-\infty}^{\infty} v[n] h[n - m] = v[n] * h[-n] \quad (53)$$

La STFT de la componente k_0 se obtendría entonces mediante la operación que indica la figura siguiente:

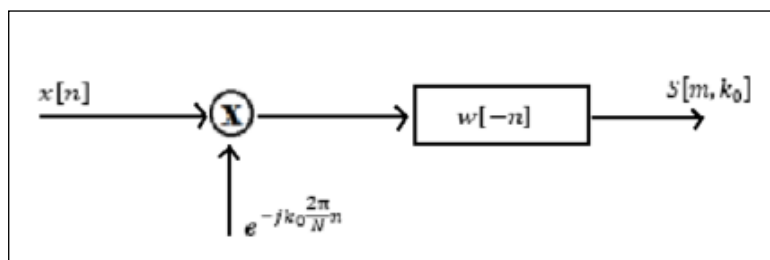


Figura 50 – Diagrama de convolución para cada valor de k_0 (banda de frecuencia).

La señal resultante $S[m, k_0]$ es una señal compleja y para recuperar una señal real en el dominio del tiempo con frecuencia de tiempo discreto $\Omega = k_0 \frac{2\pi}{N}$, es preciso considerar también la componente en $-k_0$. Esto se debe a que la parte real de una señal se define de la siguiente forma:

$$\text{Re}\{s[n]\} = \frac{s[n]+s^*[n]}{2} \quad (54)$$

En la STFT, para cada componente frecuencial se obtiene una señal compleja. Si se suma la señal de salida de k_0 con la de salida en $-k_0$, se obtendrá una señal real, que será la señal subbanda, que se explicará más adelante. Esta propiedad es debida a la hermeticidad de la transformada de Fourier.

$$\frac{S[m,k_0]+S[m,-k_0]}{2} \quad (55)$$

A modo de ejemplo, se muestra una tabla donde se pueden observar una señal de este tipo.

73i	0.6102 - 0.5495i	-0.5376 + 0.0667i	0.1015 + 0.0645i	0.0383 + 0.0000i	0.1015 - 0.0645i	-0.5376 - 0.0667i	0.6102 + 0.5495i	-0.2364 - 0.5073i	1.0020
168i	0.5847 - 0.5686i	-0.5347 + 0.0836i	0.1015 + 0.0615i	0.0368 + 0.0000i	0.1015 - 0.0615i	-0.5347 - 0.0836i	0.5847 + 0.5686i	-0.2168 - 0.5168i	1.018E
257i	0.5585 - 0.5865i	-0.5315 + 0.1003i	0.1014 + 0.0586i	0.0352 + 0.0000i	0.1014 - 0.0586i	-0.5315 - 0.1003i	0.5585 + 0.5865i	-0.1969 - 0.5257i	1.0301
340i	0.5315 - 0.6031i	-0.5280 + 0.1169i	0.1009 + 0.0557i	0.0333 + 0.0000i	0.1009 - 0.0557i	-0.5280 - 0.1169i	0.5315 + 0.6031i	-0.1767 - 0.5340i	1.035E

Figura 51 – Representación del valor real central y las parejas de complejos conjugados.

Esta tabla ha sido obtenida directamente después de aplicar la STFT a una señal acústica para poder obtener la información frecuencia-temporal en cada segmento. Como se puede observar, existe una muestra real marcada en rojo, mientras que las de los lados se van a compensar cuando se realice la semisuma de los valores complejos conjugados.

A continuación, se va a explicar la forma en la que se ha procesado la señal de banda ancha para obtener una señal real de banda estrecha, para poder usarla de forma adecuada en el análisis de las señales acústicas que producen los drones.

En primer lugar, se va a remuestrear la señal de entrada a 8 kHz, debido a que las señales acústicas que se están analizando, están muestreadas a 48 kHz. Esta configuración se debe a que la señal sonora en cuestión tiene información útil hasta los 4 kHz aproximadamente, como se puede ver en la Figura 22. Por lo tanto, se realiza este remuestreo para reducir la carga computacional, ya que no serán necesarios los 40 kHz de muestreo sobrantes.

Adicionalmente, como se ha visto, las frecuencias fundamentales obtenidas gracias al análisis cepstral están en el orden entre los 100 y los 300 hercios, siendo en esa banda donde se concentra la energía de la señal. El análisis de la atenuación del sonido con la frecuencia muestra que, en bajas frecuencias, la atenuación es más reducida y nos interesa entonces centrarnos en una banda de baja frecuencia para la estimación de la DOA.

Gracias al software de *MATLAB* con el siguiente código es posible extraer las subbandas de interés y así poder obtener las señales de banda estrecha.

```
% Conversión banda ancha-estrecha

num_subbandas=2*nextpow2(length(audio)/256);

for nmic=1:M
    S=stft(x(:,nmic),8000,'Window','hann(512)','OverlapLength',511,'FFTLenght',512,'Centered',true);
    for sub=1:num_subbandas
        subband(nmic,sub,:)=(S(256+4+sub,:)+S(256-4-sub,:))/2;
```

```
end  
end
```

Los parámetros elegidos para el cálculo de la STFT son los siguientes:

- En primer lugar, se escoge una longitud de 512 muestras para los segmentos en los que se va a dividir la señal y a los que se va a aplicar la transformada rápida de Fourier (FFT). Este valor permite obtener una resolución espectral de $\frac{8000}{512} = 15,62$ Hz para cada subbanda, suficiente para esta aplicación. Con 16 subbandas se abarcaría un ancho de banda aproximadamente de 250 Hz, que ocupa la mayor parte de los valores de frecuencia fundamental encontrados en las señales de UAV analizadas.
- Las frecuencias fundamentales se encuentran en el margen entre 100 y 300 Hz, según el análisis desarrollado en capítulos previos. Incluyendo un margen de guarda, se selecciona la banda entre 70 y 320 Hz, por lo que en la selección de las subbandas de la STFT se comienza en $k=5$ y se termina en $k=5+16$.
- Se selecciona la ventana de Hann para el análisis.
- La frecuencia de muestreo será de 8 kHz, después del cambio antes comentado, ya que la señal aparece muy sobre muestreada en las grabaciones originales.
- Como el objetivo es recuperar una señal en el dominio del tiempo, real, con la información en la banda de frecuencia de interés, y con la misma frecuencia de muestreo de 8 kHz, se deben solapar los segmentos analizados, de forma que el avance entre uno y el siguiente sea de una sola muestra. Esto se consigue con la función *stft.m* de *MATLAB*, definiendo el parámetro '*OverlapLength=511*', siendo la duración del segmento de 512 muestras.
- Se define el parámetro '*Centered=true*', para que se devuelva una transformada bilateral, con valores en frecuencia en $(-\pi, \pi]$. En este caso, la componente para $k=512/2=256$, corresponde a la subbanda centrada en 0 Hz. Esto se ilustra en la Figura 52.

Después, se crea un bucle para ir recorriendo el número de micrófonos, ya que las señales captadas por los micrófonos son diferentes entre sí, y el procesado anterior debe aplicarse a cada una de ellas.

Debe comentarse que el procesado anterior no introduce retardo dependiente de la frecuencia, y podría sustituirse por el filtrado paso banda con un filtro FIR de fase lineal paso banda. En ese caso se introduciría cierta distorsión de amplitud en la banda de interés, mientras que, con el procesado propuesto, no hay distorsión en la banda de interés. La opción elegida e implementada permite la experimentación al ser muy simple el cambio de la banda de interés, sin rediseños adicionales. Sin embargo, la complejidad computacional sería mayor que la que se tendría con el filtrado FIR. Esta opción sería seguramente la elegida en una implementación en tiempo real, tras la fase de diseño previa.

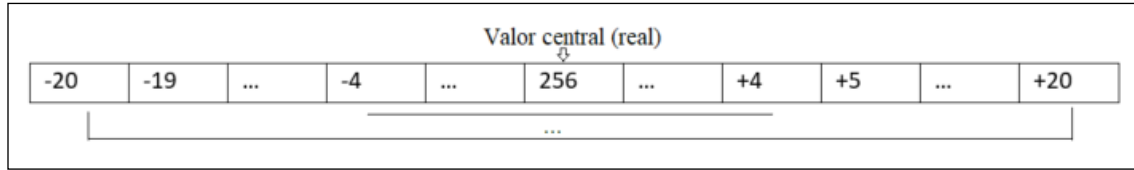


Figura 52 – Valores para formar las subbandas de la señal de entrada.

Como se puede ver en la Figura 52, cada una de las subbandas se calcula de la siguiente forma:

$$Subbanda = \frac{S(256+4+i)+S(256-4-i)}{2} \quad (56)$$

Donde S se refiere a la DFT de cada uno de los segmentos e i se refiere al número de subbanda que se esté calculando (desde 1 hasta 16 subbandas). En definitiva, la conversión de una señal de banda ancha a banda estrecha mediante este método se puede resumir de la siguiente forma:

- 1) Cada uno de los micrófonos recibe una señal de banda ancha (el sonido de un dron).
- 2) A cada señal de banda ancha, se le aplica la STFT, consiguiendo dividir la señal en segmentos de 512 muestras a los que se aplica la DFT, moviendo únicamente una muestra la ventana, para poder obtener toda la información frecuencial posible.
- 3) Entonces, en cada uno de estos segmentos, se puede aplicar la fórmula (54), para poder obtener una señal real, ya que los elementos tanto a la izquierda como a la derecha del valor central de cada uno de los segmentos son complejos conjugados. En este caso, se ha seleccionado un número de 16 valores para realizar esta operación, dejando un margen de 4 muestras desde el valor central, por lo que se realiza la semisuma desde el quinto al veinteavo elemento desde dicho valor. El resultado de estas 16 semisumas corresponde a las 16 subbandas que se están buscando.
- 4) De forma que ahora, es posible dar como argumento de entrada al estimador del algoritmo del ángulo de llegada, que se estudiará posteriormente, la frecuencia de operación del estimador, puesto que depende del valor de cada subbanda, en lugar de las 512 muestras complejas resultantes del cálculo de las DFT, que correspondían a una señal de banda ancha.

$$frecuencia\ de\ operación = \frac{(subbanda_i+4)*f_s}{512} \quad (57)$$

5.2 ALGORITMOS DE ESTIMACIÓN DEL ÁNGULO DE LLEGADA (DOA)

Una vez explicado el procedimiento seguido para obtener una señal de banda estrecha a partir de una de banda ancha, lo cual es útil para tres de los cuatro algoritmos que se van a introducir con el objetivo de poder detectar a una aeronave no tripulada, se van a tratar estos algoritmos para después poder explicar el análisis llevado a cabo para poder extraer la mejor configuración del array junto con el algoritmo que mejor se adapte en el escenario bajo análisis.

Para estimar el ángulo de llegada, los algoritmos DOA, se clasifican de la siguiente manera [29]:

- **TÉCNICAS CONVENCIONALES.** Se basan en técnicas de *beamforming* (conformación de haces). Esto es, primero, se ajustan los pesos del array, y, después, se calcula el espectro que tiene máximos en los ángulos donde la energía de la señal es mayor, por lo tanto, se

buscan los picos y así es posible detectar a las fuentes sonoras. Este tipo de técnicas DOA son las más fáciles de implementar, por lo que suelen dar menor precisión. Un ejemplo es la técnica MVDR (*Minimum Variance Distorsionless Response*).

- **TÉCNICAS BASADAS EN EL SUBESPACIO.** Estas técnicas también consisten en obtener el espectro, pero ahora se obtiene mediante la auto descomposición de la matriz de covarianza de las señales de entrada. Esta descomposición hace que la carga computacional sea algo mayor que en las técnicas convencionales. Después, se seleccionan las fuentes, según los picos de este espectro espacial. Un ejemplo es la técnica MUSIC (*MULTiple Signal Classification*).
- **TÉCNICAS BASADAS EN LA MÁXIMA VEROSIMILITUD.** Este tipo de técnicas asumen un modelo para las señales y un marco estadístico para la generación de datos y luego estiman la DOA utilizando el criterio de máxima verosimilitud [29]. Estos algoritmos tienen una gran carga computacional, aunque son muy útiles cuando la distancia es pequeña entre la fuente sonora y el array.
- **TÉCNICAS DE AJUSTE DEL SUBESPACIO.** Se refieren a técnicas paramétricas que se basan en el procesamiento del subespacio. Estos métodos tienen propiedades similares tanto a las técnicas basadas en la máxima verosimilitud, como las basadas en el subespacio.

Por lo tanto, los dos primeros métodos están basados en el espectro, puesto que consisten en calcular el espectro espacial de las señales recibidas y seleccionar el DOA como la localización de los picos del espectro. Por otro lado, los dos últimos métodos se denominan métodos de procesamiento paramétrico de arrays que directamente estiman la DOA sin calcular el espectro. Estos algoritmos tienen mejor rendimiento en términos de precisión y resolución. El coste por esta mejora es la mayor complejidad y número de operaciones computacionales [29].

Para abordar el análisis de este estudio, se van a utilizar únicamente los métodos basados en el espectro, y, en concreto, los cuatro algoritmos que se van a estudiar son: MUSIC, MVDR, Beamscan y GCC-PHAT. Los tres primeros, son algoritmos que deben recibir como entrada una señal de banda estrecha, mientras que el último debe recibir una señal de banda ancha.

5.2.1 Algoritmos de banda estrecha

Anteriormente se ha definido la forma de obtener una señal de banda estrecha a partir de una de banda ancha, debido a que se van a estudiar tres algoritmos de estimación del ángulo de llegada que deben recibir como señal de entrada esta clase de señal. Además, los tres algoritmos con los que se van a trabajar (MUSIC, MVDR y Beamscan), son algoritmos de estimación del espectro espacial, por lo que antes de entrar en el análisis de cada uno de ellos, se van a definir ciertas características en común a los tres algoritmos de banda estrecha.

A) Estimación del espectro espacial

La estimación del espectro espacial es una tecnología de estimación de señales que está especializada en utilizar matrices espaciales para lograr los parámetros de la señal. El espectro espacial presenta la distribución de energía de las señales en todas las direcciones del espacio. La dirección de máxima energía será la dirección de llegada de la fuente [28].

Las señales recibidas en el array de micrófonos se pueden expresar como una combinación lineal de señales incidentes y ruido gaussiano de media nula y con varianza σ^2 . Estas señales de entrada a los micrófonos se suponen que llegan siguiendo una línea de visión directa, sin ninguna obstrucción en su propagación, y no correladas con el ruido [29]. De esta forma, la señal recibida se expresa de la siguiente manera:

$$\mathbf{x}(t) = \sum_{m=1}^M \mathbf{a}(\theta_m) s_m(t) + \mathbf{q}(t) = \mathbf{A}\mathbf{S} + \mathbf{Q}_n \quad (58)$$

Donde, M representa el número de señales incidentes en el array, $s_m(t)$ es un vector de la forma de onda para la fuente m en la dirección θ de máxima directividad, $\mathbf{a}(\theta_m)$ es el vector dirección de cada ángulo θ_m , N es el número de sensores en el array y $\mathbf{q}(t)$ corresponde al ruido y las interferencias.

En la segunda parte de la igualdad, se obtiene:

$$\mathbf{A} = [\mathbf{a}(\theta_1) \ \mathbf{a}(\theta_2) \ \dots \ \mathbf{a}(\theta_p)], \quad \mathbf{a}(\theta) = [e^{-j\theta} \ \dots \ e^{-j(N-1)\theta}]^T$$

$$\mathbf{Q}_n = [\mathbf{q}_1 \ \mathbf{q}_2 \ \dots \ \mathbf{q}_p] \quad (59)$$

Donde \mathbf{A} es la matriz de múltiples matrices que incluyen los vectores de dirección $\mathbf{a}(\theta)$, y \mathbf{Q}_n corresponde a la matriz de vectores del subespacio de ruido.

Las técnicas para la estimación del espectro espacial conllevan el cálculo de la matriz de correlación espacial de la señal recibida \mathbf{R}_{xx} . Tal y como se explicó en el capítulo 2.3, la correlación trata de calcular el grado de semejanza entre dos señales.

$$\mathbf{R}_{xx} = E[\mathbf{x}(t) \cdot \mathbf{x}(t)^H] \quad (60)$$

En la ecuación (60), $E[\cdot]$ es el operador esperanza y H es el operador que se refiere a calcular la traspuesta conjugada de la señal. Si se relacionan las ecuaciones (58) y (60), la correlación espacial de la señal recibida se puede expresar de la forma siguiente.

$$\mathbf{R}_{xx} = E[\mathbf{A} \mathbf{s}(t) \cdot \mathbf{s}(t)^H \mathbf{A}^H] + E[\mathbf{q}(t) \cdot \mathbf{q}(t)^H] = \mathbf{A} \mathbf{R}_{ss} \mathbf{A}^H + \sigma_n^2 \mathbf{I} \quad (61)$$

Donde \mathbf{R}_{ss} es la matriz de correlación de la señal e \mathbf{I} es la matriz identidad.

Sin embargo, dado que los micrófonos no pueden recibir señales de corriente directa, los valores medios de las señales recibidas con el ruido se hacen nulos y, por tanto, la matriz de correlación obtenida en (61) se refiere como la matriz de covarianza [29].

$$\text{Matriz de covarianza espacial} = \mathbf{A}^H \mathbf{Q}_n \quad (62)$$

La estimación del espectro espacial se puede clasificar en dos categorías: técnicas de beamforming y técnicas basadas en el subespacio.

l) Beamforming

Las técnicas de beamforming fueron el primer intento de localización automática de fuentes sonoras usando arrays de micrófonos o antenas. La traducción literal al castellano de este término sería ‘formación de haces’, cuyo nombre tiene sentido según su forma de operar. El array de micrófonos apunta a una dirección en particular cada instante de tiempo y mide en esa dirección

la respuesta en potencia. Luego, las direcciones de apuntamiento que hayan dado una potencia máxima serán las estimaciones de los ángulos donde se encuentra la fuente.

La respuesta del array se dirige formando una combinación lineal de las salidas de los sensores [30].

$$y(t) = \sum_{t=1}^M \mathbf{w}_t^* \mathbf{x}_t(t) = \mathbf{w}^H \mathbf{x}(t) \quad (63)$$

La potencia de salida se mide mediante:

$$P(\mathbf{w}) = \frac{1}{N} \sum_{t=1}^N |y(t)|^2 = \frac{1}{N} \sum_{t=1}^N \mathbf{w}^H \mathbf{x}(t) \mathbf{x}^H(t) \mathbf{w} = \mathbf{w}^H \mathbf{R}_{xx} \mathbf{w} \quad (64)$$

Bartlett beamforming (Beamscan)

Beamscan es el nombre con el que el software *MATLAB* se refiere a la técnica de beamforming convencional o beamforming de Bartlett. Esta técnica de DOA es una extensión natural del clásico análisis espectral de Fourier de los arrays de datos [30]. Este algoritmo maximiza la potencia de salida del beamforming para cualquier señal recibida por un array de geometría arbitraria.

En [30] se estudia este algoritmo suponiendo que se intenta maximizar la salida a una determinada dirección θ . Así, los autores exponen la forma de operar de dicho algoritmo.

Se recibe una señal en la dirección θ , sin embargo, esta señal de entrada al array de micrófonos está corrupta con ruido aditivo.

$$\mathbf{x}(t) = \mathbf{a}(\theta)s(t) + \mathbf{n}(t) \quad (65)$$

Entonces, para poder maximizar la potencia de salida, existe el siguiente problema: al intentar maximizar la potencia de salida, para comprobar si el ángulo al que se está apuntando con el array es la dirección donde está presente una fuente, se está maximizando también el ruido aditivo. Esto se muestra en la siguiente expresión:

$$\max_{\mathbf{w}} E\{\mathbf{w}^H \mathbf{x}(t) \mathbf{x}^H(t) \mathbf{w}\} = \max_{\mathbf{w}} \{E|s(t)|^2 |\mathbf{w}^H \mathbf{a}(\theta)|^2 + \sigma^2 |\mathbf{w}|^2\} \quad (66)$$

En [30] continúan desarrollando dicho tema, donde para obtener una solución no trivial cuando se usa la maximización anterior, los pesos de los micrófonos se restringen a un valor de módulo unidad, resultando la solución de la ecuación (67).

$$\mathbf{w}_{BF} = \frac{\mathbf{a}(\theta)}{\sqrt{\mathbf{a}^H(\theta)\mathbf{a}(\theta)}} \quad (67)$$

El vector de pesos anterior se puede interpretar como un filtro espacial emparejado con la señal de entrada. Intuitivamente, la ponderación de la matriz iguala los retrasos (y, posiblemente, las atenuaciones) experimentados por la señal en varios sensores para combinar al máximo sus respectivas contribuciones [30].

Así, si se combinan las ecuaciones (64) y (67), se obtiene el espectro espacial del beamformer convencional:

$$P_{BF}(\theta) = \frac{\mathbf{a}^H(\theta)\mathbf{R}_{xx}\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\mathbf{a}(\theta)} \quad (68)$$

El mayor inconveniente de esta técnica de DOA es que el espectro espacial tiene la misma limitación de resolución que el periodograma. Por lo tanto, las fuentes que estén localizadas en posiciones más próximas que el ancho de haz del array de micrófonos, no podrán ser diferenciadas y se notarán como una única fuente.

MVDR

MVDR (*Minimum Variance Distorsionless Response*), también llamado beamformer de Capon, es una técnica de beamforming, por lo que pertenece a las técnicas convencionales. MVDR es conocida como la técnica óptima de beamforming, ya que no requiere conocimiento previo de las direcciones ni los niveles de potencia de las interferencias o del ruido, para maximizar la relación señal a ruido de salida. Además, es capaz de abordar la limitación del beamforming de Bartlett, ya que ofrece una resolución superior de forma que es capaz de separar dos fuentes posicionadas más cerca que un ancho de haz.

La optimización que ofrece MVDR es minimizar la potencia aportada por el ruido y las interferencias no deseadas, mientras que mantiene una ganancia fija en la dirección de apuntamiento, generalmente igual a la unidad [30]. Esto se puede expresar de la siguiente forma:

$$\begin{aligned} \min E[|y(\theta)|^2] &= \min \mathbf{w}^H \mathbf{R}_{xx} \mathbf{w} = \min P(\mathbf{w}) \\ \mathbf{w}^H \mathbf{A}(\theta_0) &= 1 \end{aligned} \quad (69)$$

Donde $y(\theta)$ denota la salida del array y \mathbf{w} denota el vector de pesos. \mathbf{w} viene dado, por la ecuación siguiente.

$$\mathbf{w} = \frac{\mathbf{R}_{xx}^{-1} \mathbf{A}}{\mathbf{A}^H \mathbf{R}_{xx}^{-1} \mathbf{A}} \quad (70)$$

Los pesos del array son seleccionados asumiendo todas las fuentes como una interferencia, y el procesador se puede denominar filtro de máxima verosimilitud (ML), al encontrar la estimación de la potencia de la señal de ML con la suposición anterior [31].

Finalmente, los picos del espectro angular vienen dados por los valores máximos obtenidos de la siguiente ecuación, que ocurren cuando el vector de direccionamiento y el subespacio de ruido son ortogonales entre sí [29].

$$P_{MVDR} = \frac{1}{\mathbf{A}^H \mathbf{R}_{xx}^{-1} \mathbf{A}} \quad (71)$$

El espectro espacial ofrecido por MVDR supera el del beamformer convencional debido a que usa cada grado de libertad disponible para concentrar la energía recibida en la dirección de interés. En la ecuación (69), la minimización de potencia también se puede interpretar como el sacrificio de cierta capacidad de supresión de ruido para obtener una “anulación” más enfocada en las direcciones donde están presentes otras fuentes. Por lo tanto, la fuga espectral de fuentes poco espaciadas se reduce, aunque la capacidad de resolución de MVDR aún depende de la apertura del conjunto y de la SNR [30].

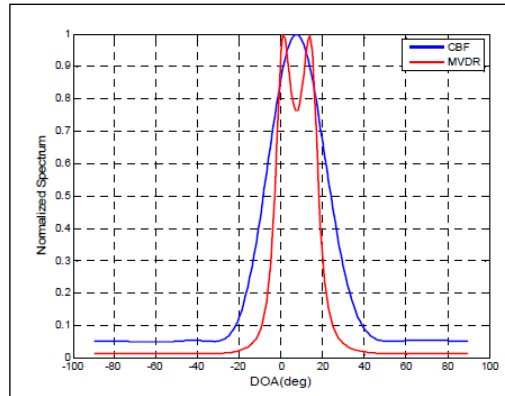


Figura 53 – Comparación entre beamformer convencional y MVDR en la estimación de dos fuentes próximas [32].

MVDR es un algoritmo DOA que acepta cualquier configuración del array de micrófonos que captan la señal de entrada. Además, es una técnica que representa el beamformer de mayor optimización y obtiene una DOA de poca carga computacional.

En definitiva, la idea de los métodos de beamforming es apuntar en una dirección en particular cada instante de tiempo y medir la respuesta en potencia en cada dirección. Las direcciones de apuntamiento que den una potencia máxima serán las estimaciones DOA. Sin embargo, con la técnica de beamforming convencional existían ciertas limitaciones, como la proximidad de dos fuentes, ya que darían lugar a percibir una única fuente. No obstante, con MVDR esta limitación se reduce y es posible percibir dos fuentes por separado cuando éstas están incluso a una distancia menor del ancho de haz del array de micrófonos. Por ello, ofrece una resolución mucho mayor a cualquier otra técnica convencional.

II) Técnicas basadas en el subespacio

En el pasado, muchos métodos espectrales han recurrido implícitamente a la descomposición espectral de una matriz de covarianza para llevar a cabo el análisis espectral [30].

Los primeros enfoques en estas técnicas incluyen el análisis factorial de componentes principales y el análisis de series de tiempo de errores en las variables. Sin embargo, una de las contribuciones más significativas se produjo cuando se invocó explícitamente la estructura propia de la matriz de covarianza y sus propiedades intrínsecas se utilizaron directamente para proporcionar una solución a un problema de estimación subyacente para un proceso observado dado [30].

El enfoque subspectral debe su interés, principalmente, a la introducción del algoritmo MUSIC (*Multiple Signal Classification*), el cual se introdujo directamente como un algoritmo DOA, en lugar de una técnica previa a la estimación del ángulo de llegada.

MUSIC

MUSIC (*MUltiple Signal Classification*) es un algoritmo basado en el subespacio que recibe como señal de entrada una señal de banda estrecha. Es un método de alta resolución que está basado en las propiedades espaciales y la habilidad de descomponer la señal en dos subespacios ortogonales (el subespacio de señal y el subespacio de ruido).

Para la estimación del ángulo de llegada, MUSIC utiliza las propiedades del subespacio de la matriz de covarianza espacial. Entonces, primero se estima la autocorrelación o autocovarianza

de un modelo de señal más ruido y este modelo se utiliza para formar una matriz cuya estructura propia da el subespacio señal a ruido [33]. Por otro lado, la autocovarianza determina si existe dependencia entre las variables de un proceso en cada par de instantes.

MUSIC opera de la siguiente forma: si llega una cantidad de señales al array de micrófonos, los vectores propios que tienen los autovalores más grandes de la matriz de covarianza abarcan el subespacio de la señal, mientras que los vectores propios correspondientes a los autovalores más pequeños de la matriz de covarianza abarcan el subespacio de ruido. Así, los vectores de dirección que forman la matriz a son ortogonales al subespacio de ruido, y, por tanto, los vectores de dirección se encuentran en el espacio de señal [29]. Esto se consigue obteniendo los picos del espectro de MUSIC mediante la siguiente fórmula:

$$P_{MUSIC}(\theta) = \frac{1}{A^H Q_n Q_n^H A} \quad (72)$$

En otras palabras, si llega un ángulo θ donde está la fuente, $a(\theta)$ será ortogonal al subespacio de ruido, de forma que el denominador de la ecuación (72) se hará nulo, y, por lo tanto, $P_{MUSIC}(\theta)$ se hará máximo, dando lugar a un ángulo DOA. Sin embargo, para operar el algoritmo MUSIC de esta forma, se necesita cierto conocimiento previo acerca de la respuesta del sistema, número de señales y el nivel de ruido e interferencia. Con lo cual, en la mayoría de los escenarios, el procesador no conoce variables como la matriz de autocorrelación o el ruido, y ha de hacer una estimación, con lo cual habrá ciertas disparidades.

MUSIC es un algoritmo DOA, que acepta cualquier tipo de configuración en la geometría o número de micrófonos/distancia entre elementos del array. Además, está basado en el subespacio, proporciona alta resolución y precisión, y, tiene una baja carga computacional. Sin embargo, necesita de cierto conocimiento previo para ofrecer las mejores estimaciones, y, además, si las señales están correladas, se necesita de una especie de suavizado del subespacio para poder contar con su alta resolución.

5.2.2 Algoritmos de banda ancha

La señal acústica proveniente de un dron es una señal de banda ancha, y, por tanto, se pueden usar directamente algoritmos que reciban como señal de entrada una señal de banda ancha. Para ello se va a introducir el algoritmo GCC-PHAT.

GCC-PHAT

GCC-PHAT (*Generalized Cross Correlation with Phase Transform*) es una técnica de estimación del ángulo de llegada que recibe como señal de entrada una señal de banda ancha. Este algoritmo consiste en utilizar la correlación cruzada generalizada para estimar la diferencia del tiempo de llegada de la señal a cada uno de los sensores y así poder localizar de qué dirección proviene la señal. La correlación cruzada generalizada trata de realizar una comparación entre dos señales obteniendo su grado de semejanza para así poder obtener las características desconocidas de una de ellas. El array de micrófonos recibe una señal como la ecuación (65). Esta señal se recibe en los micrófonos del array en instantes de tiempo ligeramente distintos. Este tiempo de retraso entre la recepción de la señal en un micrófono y otro se denomina TDOA (*Time Difference Arrival*). Para estimar el valor de este retardo de tiempo se debe encontrar el tiempo de retraso que haga máxima la función de correlación cruzada entre las señales recibidas en cada par de micrófonos $x_1(t)$ y $x_2(t)$.

$$R(\tau) = E\{x_1(t) x_2(t + \tau)\}$$

$$\hat{D} = \arg \max_{\tau} R(\tau) \quad (73)$$

En la ecuación anterior, $R(\tau)$ es la función correlación cruzada entre dos señales recibidas en un par de sensores, D se refiere al retardo de tiempo y τ es el máximo de correlación.

Esta función tiene una característica importante y es que cuanto más puntiaguda sea la correlación cruzada entre un par de sensores, el rendimiento mejorará. Por tanto, para poder realizar dicha mejora, se usa la técnica de GCC (*Generalized Cross-Correlation*), que consiste en usar una función de ponderación para suavizar las señales de entrada. Además, una función de ponderación particular PHAT (*Phase Transform*) normaliza la densidad espectral de la señal por la magnitud del espectro [24] y es útil en casos donde existe reverberación, para evitar que los ecos en paredes, objetos, etc., sean interpretados como fuentes originales.

En [34], se representa la ecuación de la correlación cruzada del algoritmo GCC-PHAT de la siguiente forma:

$$R_{il}(\tau) = \int \psi_{il}(f) X_i(f) X_l^*(f) e^{j2\pi f \tau} df \quad (74)$$

Donde i es el índice del micrófono y ψ_{il} es el filtrado previo donde se define PHAT como:

$$\psi_{il}(f) = \frac{1}{|X_i(f) X_l^*(f)|} \quad (75)$$

En conclusión, GCC-PHAT es un algoritmo DOA que recibe como entrada una señal de banda ancha. GCC-PHAT utiliza la función de correlación cruzada generalizada para calcular el retardo de la señal de entrada entre un par de micrófonos. Para ello, utiliza un filtrado previo, que pondera la densidad espectral de la señal por la magnitud del espectro. Obteniendo este retardo entre los micrófonos es posible obtener el ángulo donde se encuentra la fuente sonora, ya que será aquella para la cual el retardo sea mínimo.

6 ANÁLISIS Y RESULTADOS

Una vez se han obtenido las frecuencias fundamentales de las señales acústicas de los drones, la atenuación que va a experimentar el sonido durante su propagación y se han explicado tanto las características de los arrays de micrófonos (el patrón de directividad, los ángulos en los que emiten y las distintas geometrías), como los algoritmos de estimación de llegada (DOA), se va a realizar un análisis para seleccionar la mejor configuración del array junto con el algoritmo DOA para detectar una señal acústica proveniente de un dron UAV. Dicho análisis consta de tres pasos esenciales.

- En primer lugar, se van a utilizar todas las geometrías disponibles (ULA, URA y UCA) de los arrays de micrófonos junto con todos los algoritmos DOA definidos en el capítulo anterior (Beamscan, MVDR, MUSIC y GCC-PHAT). Para este primer paso, se va a ubicar al dron en un ángulo tanto en azimut como en elevación de valores ordinarios, y se va a ir variando tanto el número de micrófonos de cada array como la distancia entre sensores, para poder obtener la configuración de cada geometría del array que presente una mejor respuesta para cada uno de los algoritmos DOA bajo estudio.
- Después, en la segunda etapa, se va a realizar un barrido de los ángulos en azimut y elevación, con las configuraciones obtenidas en el primer paso. Este análisis trata de obtener los mejores resultados generales para el caso de que el dron pueda estar en cualquier ángulo. Por lo tanto, se descartarán combinaciones que sobrepasen un máximo de error absoluto medio.
- Por último, con las configuraciones seleccionadas del segundo paso, se va a aplicar a la señal acústica del dron, la atenuación en un ambiente real caracterizada en el capítulo 3, además de cierto porcentaje de ruido blanco gaussiano. Esta implementación se realiza con el objetivo de obtener un escenario más próximo a la realidad y así poder confirmar la combinación de array y algoritmo que mejor operen tanto en condiciones ideales como reales. Finalmente, para la configuración seleccionada, se volverá a modificar las características del array, para observar si alguna configuración presentaba mejores resultados para un ambiente real.

El análisis en simulación se ha realizado gracias a las funciones y algoritmos del software *MATLAB* [24].

6.1 ELECCIÓN DEL ARRAY PARA CADA DOA

En primer lugar, como se ha explicado anteriormente, se van a seleccionar varias configuraciones para poder abordar este estudio. Para ello, se va a utilizar tanto la geometría del array circular uniforme como la del array rectangular uniforme. El array lineal uniforme (ULA) no se va a considerar en este análisis, puesto que, como se ha explicado en el marco teórico de los arrays de micrófonos, el ULA no ofrece ninguna información acerca del plano de elevación, y el objetivo de este trabajo es encontrar el punto exacto en 2D donde exista un dron para que no interfiera, por ejemplo, en un estacionamiento aeroportuario.

Por otro lado, se utilizan estas geometrías (UCA y URA) con cuatro algoritmos distintos de estimación del ángulo de llegada (Beamscan, MVDR, MUSIC, y GCC-PHAT) para poder

seleccionar las configuraciones que ofrezcan mejores resultados para cada uno de los DOA bajo análisis. Para ello, se va a variar el número de micrófonos, así como, la distancia entre ellos de forma progresiva.

Además, como se ha visto en el marco teórico, los algoritmos de MUSIC, MVDR y Beamscan deben recibir como señal de entrada una señal de banda estrecha, mientras que en GCC-PHAT la señal de entrada es de banda ancha, como la señal acústica de un UAV. Por lo tanto, para estos algoritmos de banda estrecha, debe obtenerse una señal de banda estrecha en la banda de interés, tal y como se indicó en la sección 5.1.

De esta manera, el primer paso de este análisis es introducir variables como el tipo de micrófonos de los que consta el array, la velocidad del sonido o la posición donde se encuentra el dron, ya que como se indicó a principio de capítulo, se sitúa el dron UAV en una posición ordinaria donde es posible que se encuentre normalmente (40° en azimut, 30° en elevación). Además, la señal sonora que se introduce se va a remuestrear de una frecuencia de 48 kHz a una frecuencia de 8 kHz, puesto que la señal únicamente tiene información hasta los 4 kHz aproximadamente. Asimismo, las frecuencias fundamentales obtenidas son del orden de entre los 100 y 300 hercios, con lo cual, se realiza este remuestreo para reducir la carga computacional.

```
% Se remuestrea la señal
audio=resample(audio_48,8000,48000);
```

Por otro lado, las señales acústicas que se han utilizado para el resto del trabajo eran del orden del minuto y tenían en cuenta el despegue y aterrizaje de los drones (inicio y fin del audio). Como el objetivo último de este análisis es detectar un dron en el aire, se ha acortado el audio seleccionando únicamente una porción de 5 segundos de la parte central donde el dron esté o bien moviéndose hacia cualquier dirección, o bien sosteniéndose en el aire.

```
% Se seleccionan 5 segundos del audio
t = 0:1/8000:5;
audio = audio(round(length(audio)/2)+1:round(length(audio)/2+length(t)));
```

Una vez introducidas las variables y configuraciones, se va a iniciar un bucle en el que se selecciona un número distinto de micrófonos en cada vuelta. Dentro de éste, se introduce otro bucle en el que esta vez se varía la distancia entre los micrófonos del array. Así, se puede obtener para cada número de micrófonos todas las distancias entre sensores posibles, de forma que nunca superen el umbral de $d = \frac{\lambda}{2}$, donde se empezará a producir aliasing y se perderá la información útil de los algoritmos DOA. Siguiendo estas indicaciones, el número de micrófonos se varía desde 4 a 16 micrófonos, en pasos de dos micrófonos, y, la distancia entre sensores desde 0,3 a 0,6 metros, en pasos de 0,05 metros.

Entonces, dentro de estos bucles, para cada configuración del array, se introduce la geometría del array, URA o UCA, que depende del número de micrófonos y de la distancia entre sensores o el radio del array, respectivamente. Además, se define el colector de banda ancha (*phased.WidebandCollector* en *MATLAB*) que recibirá la señal acústica en la dirección que se le haya indicado, y que consta de los siguientes elementos para ello: el array (UCA o URA), la velocidad a la que llega la onda (c), la tasa de muestreo (f_s) y el número de subbandas de procesado (1024 subbandas).


```
collector=phased.widebandCollector('Sensor',ura,'PropagationSpeed',c,...
    'SampleRate',8000,'ModulatedInput',false,'NumSubbands',1024);
```

El siguiente paso, en el caso de GCC-PHAT, sería definir el estimador. Para ello, se llama a la función de *MATLAB* ‘*phased.GCCEstimator*’, y se le pasan los argumentos de entrada del array de micrófonos, la tasa de muestreo, la velocidad del sonido y se indica que el puerto de salida de los valores de retardos muestre su resultado.

Además, se divide la señal acústica recibida en segmentos de 2000 muestras de longitud (0,25 segundos), debido a que, en la práctica, se debe trabajar con ventanas de datos de duración finita, por lo que no se puede obtener la señal completa directamente, y, además, se elige esta configuración dado que se están usando algoritmos que llevan a cabo el cálculo de correlaciones, por lo que es necesario que el número de muestras sea relativamente grande.

Gracias a la función de *MATLAB* ‘*step*’, el estimador devuelve la dirección de llegada de la onda sonora recibida cada 2000 muestras de la señal. De esta forma, se recibirán un total de $\frac{\text{longitud señal}}{2000 \text{ muestras}} = 20$ estimaciones de ángulos de llegada para cada configuración. Estos resultados formarán una matriz de dos filas (azimut y elevación) y 20 columnas, cada una de ellas con una estimación.

```
estimator=phased.GCCEstimator('SensorArray',ura,'SampleRate',...
    8000,'PropagationSpeed',c,'DelayOutputPort',true,'CorrelationOutputPort',true);

for m=1:fix(length(x)/2000)
    [ang(:,m),tau(:,m)]=step(estimator,x((m-1)*2000+1:m*2000,:));
end
```

Por otro lado, para el caso de Beamscan, MVDR y MUSIC, antes de llegar al estimador, se realiza la conversión de banda ancha a banda estrecha, calculando las subbandas, como se indicó en el capítulo 5.1. De esta forma, para cada una de las subbandas se define el estimador del DOA, que obtiene como argumentos de entrada: el array de micrófonos, la frecuencia de operación, la velocidad del sonido, y el número de señales, entre otras variables. A partir de aquí, se divide la señal acústica en segmentos de 2000 muestras como en el caso de GCC-PHAT y se calcula el ángulo de llegada de la onda sonora recibida por cada 2000 muestras y en cada una de las 16 subbandas de la señal. Así, en estos casos, se formarán dos matrices (azimut y elevación) y en cada una de ellas habrá 16 filas correspondientes a las subbandas de la señal y 20 columnas correspondientes a las estimaciones de ángulos de llegada para cada una de las subbandas.

Una vez obtenidas las estimaciones de los ángulos donde se encuentra la señal acústica del dron, se va a calcular el error medio absoluto de cada configuración (número de micrófonos y distancia entre sensores) con el fin de encontrar la mejor configuración del array de micrófonos para cada uno de los algoritmos DOA propuestos. El error absoluto medio se calcula como el módulo de la diferencia entre el valor real del ángulo donde se encuentra el dron y la media de las mediciones obtenidas.

De esta manera, para el algoritmo Beamscan con UCA, por ejemplo, se obtienen los errores en azimut para cada una de las configuraciones disponibles presentes en la tabla de la Figura 54. Asimismo, se obtendría una tabla similar para los errores medios absolutos en el ángulo de elevación.

ERROR ABSOLUTO BEAMSCAN UCA (AZIMUT)							
Distancia entre micrófonos (m) → Número de micrófonos ↓	0,3	0,35	0,4	0,45	0,5	0,55	0,6
4	NaN	NaN	NaN	1,21	0,89	0,50	0,26
6	0	0,70	2,01	2,48	2,85	3,30	3,69
8	0,01	0,01	0,02	0,01	0,01	0,12	0,27
10	0	0,01	0,03	0,04	0,19	0,02	0,61
12	0,01	0,01	0,41	0,41	0,01	0,09	0,39
14	0	0,9	0,47	0	0	0,92	3,11
16	0,39	0,48	0,01	0,70	2,24	2,43	1,17

Figura 54 – Error medio absoluto en azimut del algoritmo BeamsScan con un array UCA, donde cada fila corresponde al número de micrófonos y cada columna a la distancia entre ellos.

Por último, una vez se ha recorrido todo el bucle para cada configuración del array, se calcula la mejor combinación tanto del número de micrófonos como de la distancia entre ellos, seleccionando la configuración que menor error suponga teniendo en cuenta la detección en azimut y elevación por separado. Como resultado se obtiene la configuración que mejor resultados ofrece para un ángulo en azimut y elevación ordinarios, y para cada uno de los DOA, mostradas en la tabla de la Figura 57. Además, puesto que el algoritmo GCC-PHAT no proporciona ángulo de elevación para un array circular, se descartará dicha configuración, al igual que pasaba con el caso del array lineal uniforme.

		N.º micrófonos	Distancia (m)	Error (dB)
BEAMSCAN	UCA	8	0,5	0,03
	URA	6	0,55	0,14
MVDR	UCA	4	0,3	0,15
	URA	8	0,3	0
MUSIC	UCA	16	0,4	0,16
	URA	12	0,35	0,29
GCC PHAT	UCA	-	-	-
	URA	12	0,35	0,01

Figura 55 – Selección de configuraciones para cada geometría y algoritmo DOA.

6.2 SELECCIÓN CONFIGURACIÓN MEDIANTE BARRIDO DE ÁNGULOS

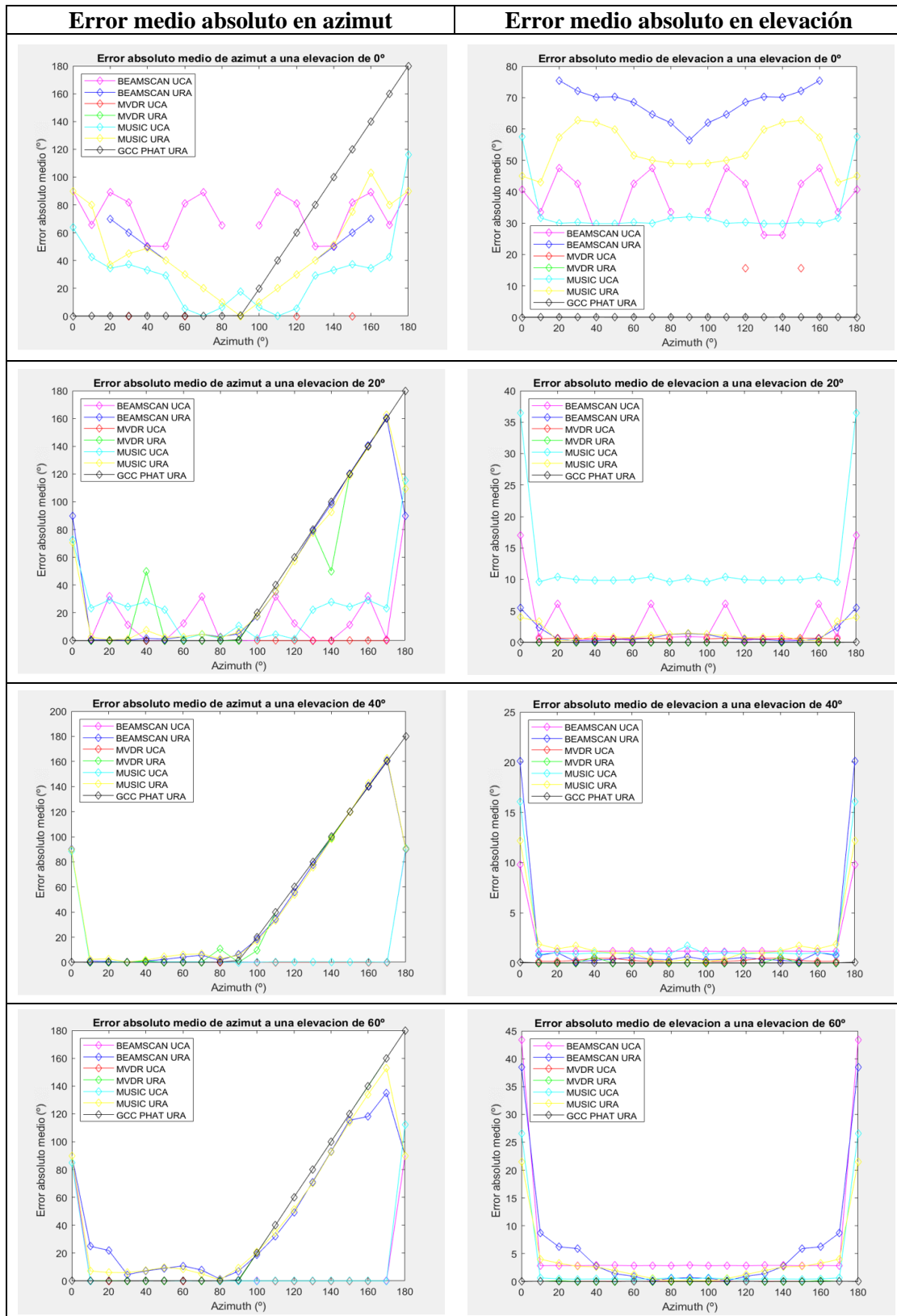
A continuación, se va a realizar una simulación en la que el dron va a ir cambiando de posición tanto en azimut como en elevación para las configuraciones seleccionadas en el primer paso de este análisis. Para ello, se va a ir seleccionando un valor de elevación (0 °, 20 °, 40 °, 60 ° y 80 °) cada vez y, para dicho valor, se va a realizar un barrido en azimut desde 0 ° hasta 180 ° en pasos de 10 grados.

En esta sección del análisis, las variables de entrada son iguales que en el primer paso, al igual que la llamada a los algoritmos, colectores y cálculos de los errores absolutos medios. Sin embargo, en este caso, en lugar de crear un bucle en el que se vaya variando las características del array de micrófonos, el bucle se utilizará para poder realizar el barrido en azimut para cada una de las elevaciones comentadas.

Los siguientes resultados representan el error medio absoluto obtenido en función del ángulo en azimut del UAV, de forma que a la izquierda de la Figura 56 se encuentran las representaciones

Capítulo 6 – Análisis y resultados

del error medio absoluto respecto a la estimación de la dirección en azimut y a la derecha respecto a la dirección en elevación para cada una de las distintas elevaciones.



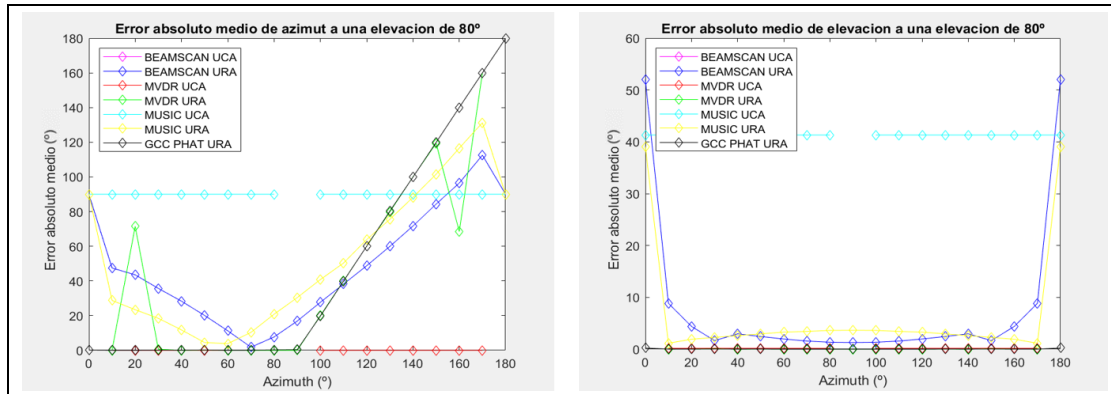


Figura 56 – Error medio absoluto en función de un barrido de ángulos azimut para elevaciones de 0, 20, 40, 60 y 80 grados. (Izquierda: Error en azimut; Derecha: Error en elevación).

Con relación a las figuras anteriores, se exponen las características obtenidas de cada uno de los algoritmos de detección del ángulo de llegada en las distintas localizaciones.

BEAMSCAN

El algoritmo de Beamscan (beamformer convencional), ofrece buenos resultados con la configuración del array circular uniforme para cualquier configuración de azimut, mientras que el plano de elevación no supere los 60 grados, ya que, a partir de la elevación de 60 grados, el algoritmo Beamscan no ofrece ninguna estimación de dirección con la configuración de UCA tanto en azimut como elevación.

En cambio, si se utiliza una configuración del array URA con Beamscan, es posible obtener información para cualquier ángulo de azimut y elevación. Sin embargo, a partir de un valor de azimut de 90 grados, el error al detectar la dirección en azimut empieza a crecer linealmente y a gran medida, de forma que a partir de los 140/160 grados el error supera el centenar de grados. Por el contrario, el error en la detección del ángulo de elevación ofrece valores bajos, menores a 10°, para cualquier posición espacial, excepto a 0 grados de elevación.

MVDR

MVDR, que corresponde a la optimización de las técnicas de beamforming, obtiene, con la geometría UCA, resultados del error cercanos a cero para la mayoría de los ángulos tanto en azimut como en elevación. Sin embargo, cuando se da el caso que el dron está localizado a una elevación nula, MVDR no detecta ninguna señal. Este inconveniente carece de importancia en el escenario que se está analizando, debido a que se está suponiendo que el dron siempre está elevado en el aire. Además de este pequeño inconveniente, MVDR con la geometría circular uniforme, no puede detectar la señal acústica en los ángulos de azimut de 0, 90 y 180 grados.

Por otro lado, si se utiliza la geometría rectangular uniforme con el algoritmo MVDR, tampoco ofrece ninguna información a una elevación de 0 grados, pero, además, al igual que en el caso de Beamscan, a partir de 90° el error medio absoluto del algoritmo empieza a aumentar linealmente y acaba superando los 100° cuando se trata de detectar la dirección en azimut únicamente. Sin embargo, la detección del ángulo de elevación se puede realizar sin problemas con errores casi nulos para todo el rango espacial.

MUSIC

El algoritmo MUSIC produce errores altos de estimación cuando el sonido proviene de elevaciones de 0 y 80 grados. A una elevación de 20 grados, el error en la detección del ángulo

de azimut fluctúa entre valores de 0° y 20° y el error en la detección del ángulo de elevación se mantiene constante a 10° . Luego, para el resto de las elevaciones, las estimaciones producidas por MUSIC son buenas, ya que son casi nulas y solo fallan a valores de 0 y 180 grados de azimut. Por otro lado, si se escoge una geometría del array rectangular, comenzará a fallar a partir de posiciones de 90 grados de azimut cuando se estima la posición en azimut, tal y como ocurría con esta geometría en los algoritmos de beamforming.

GCC-PHAT

Por último, el algoritmo de estimación de la dirección del dron que recibe como entrada una señal de banda ancha, GCC-PHAT, se ha evaluado únicamente con la configuración del array rectangular uniforme debido a la falta de estimación en elevación del UCA comentada en el apartado anterior.

Este algoritmo DOA, ofrece errores cercanos a cero para todos los valores de elevación en la detección de la elevación del dron. Sin embargo, al utilizarse un array URA, a partir de un azimut de 90 grados, en la estimación del ángulo de azimut, el algoritmo pierde fuerza y comienza a producir errores muy significativos.

En definitiva, en condiciones ideales, depende de los objetivos perseguidos, la utilización de un algoritmo u otro. Por ejemplo, si se necesitase información en todos los ángulos de azimut y para la mayor parte de elevaciones, se utilizaría el algoritmo MVDR con la configuración UCA, debido a que, aunque llevase a perder alguna estimación momentánea, este algoritmo ofrecería los mejores resultados, obteniendo siempre una estimación cercana a la posición real del dron. Si, por el contrario, se necesitase información a cualquier valor de elevación y en azimut hasta los 90 grados, se utilizaría el algoritmo GCC-PHAT con la geometría URA. Esto se debe a que ofrece las mejores estimaciones para estos valores, además de no tener que realizar la conversión del sonido del UAV a banda estrecha que pudiera dar lugar a cierto error.

Con respecto a la continuación del análisis para obtener la mejor configuración, se debe agregar cierta atenuación y ruido al canal para comprobar que funcionen de forma óptima también en condiciones no ideales. Para ello, se selecciona una configuración de array para cada tipo de DOA, realizando así el análisis del último paso con la siguiente selección (Beamscan con UCA, MVDR con UCA, MUSIC con UCA y GCC-PHAT con URA), debido a que las demás configuraciones suponían un inconveniente mayor en gran parte de los escenarios.

6.3 SELECCIÓN ARRAY Y DOA EN UN AMBIENTE REAL

A continuación, se realiza un análisis de la respuesta de los algoritmos DOA en condiciones de ambiente reales. Para ello, se debe calcular tanto la atenuación del canal como agregarle cierto porcentaje de ruido blanco gaussiano.

La atenuación que se va a añadir al canal viene calculada para cada una de las frecuencias presentes en la señal sonora de un dron. Esta atenuación se calcula de la misma forma que se calculaba en el capítulo 3 para las frecuencias fundamentales y sus armónicos, y es función de la frecuencia y distancia entre la fuente y el array de micrófonos. De esta forma, se obtiene la atenuación que sufre la huella acústica de un dron en un ambiente real. La señal atenuada resultante tanto en el dominio temporal como frecuencial se ve reflejada en las ecuaciones (76) y (77).

$$\text{Señal atenuada en dB} = 10 \log_{10} S(w) - At \text{ (dB)} \quad (76)$$

$$\text{Señal atenuada temporal} = \text{Re} \left\{ \text{TF}^{-1} \left(\frac{\text{audio}}{10^{\frac{At(dB)}{10}}} \right) \right\} \quad (77)$$

```

% Atenuación aplicada a la señal del dron
L=length(audio);
audio_fft=fft(audio,L,1);
audio_freq=10*log10(abs(audio_fft/L));

[atenuacion]=Atenuacion_total(d,L); % Atenuacion en función de la distancia y frecuencia
atenuacion_nodB=10.^(atenuacion/10);

for d=1:400
    audio_fft_at_dB(:,d)=audio_freq-atenuacion(:,d); % Espectro atenuado en decibelios

    audio_fft_at(:,d)=audio_fft./atenuacion_nodB(:,d);
    audio_at(:,d)=ifft(audio_fft_at(:,d),L,1); % Señal atenuada en el dominio temporal
end

audio_at_real=real(audio_at);
  
```

Como resultado se obtiene la señal acústica de la fuente sonora atenuada en función de la distancia y la frecuencia, es decir, se obtiene una matriz en la que las filas son las frecuencias del UAV y las columnas corresponden a la distancia entre la fuente y el array de micrófonos, ya que la señal se irá deteriorando según vaya aumentando esta distancia, sobre todo debido a la divergencia geométrica comentado en el capítulo de la propagación del sonido en ambientes reales.

Por otro lado, a la señal resultante del UAV atenuada se le añade un porcentaje del 10% de ruido, de forma que siga predominando el sonido del dron, pero se vean afectadas sus componentes por dicho ruido aleatorio.

```

% Ruido
noise = 0.1*(randn(size(collector))+1i*randn(size(collector)));
  
```

Con relación a calcular la estimación de la dirección de llegada del dron, se va a seguir el procedimiento que consistía en la introducción del colector, estimador y la definición de características comentadas al inicio del capítulo, en el apartado 6.1. No obstante, los bucles para ir calculando las estimaciones, en este caso, van a consistir en la variación: del azimut, para poder tener diferentes posiciones del dron, y, de la distancia entre el dron y el array de micrófonos, de forma que se pueda caracterizar el posicionamiento en 3D. La distancia entre emisor y receptor acústicos se va a variar entre 1 y 400 metros, debido a que un dron no puede volar a gran distancia, y, a mayor distancia, la relación señal a ruido de la señal acústica del dron será tan baja que será inapreciable en condiciones reales en las que haya presencia de otras fuentes sonoras.

Con este fin, se va a comprobar las cuatro configuraciones seleccionadas en el apartado 6.2 para tres elevaciones distintas: 20 grados, 40 grados y 60 grados.

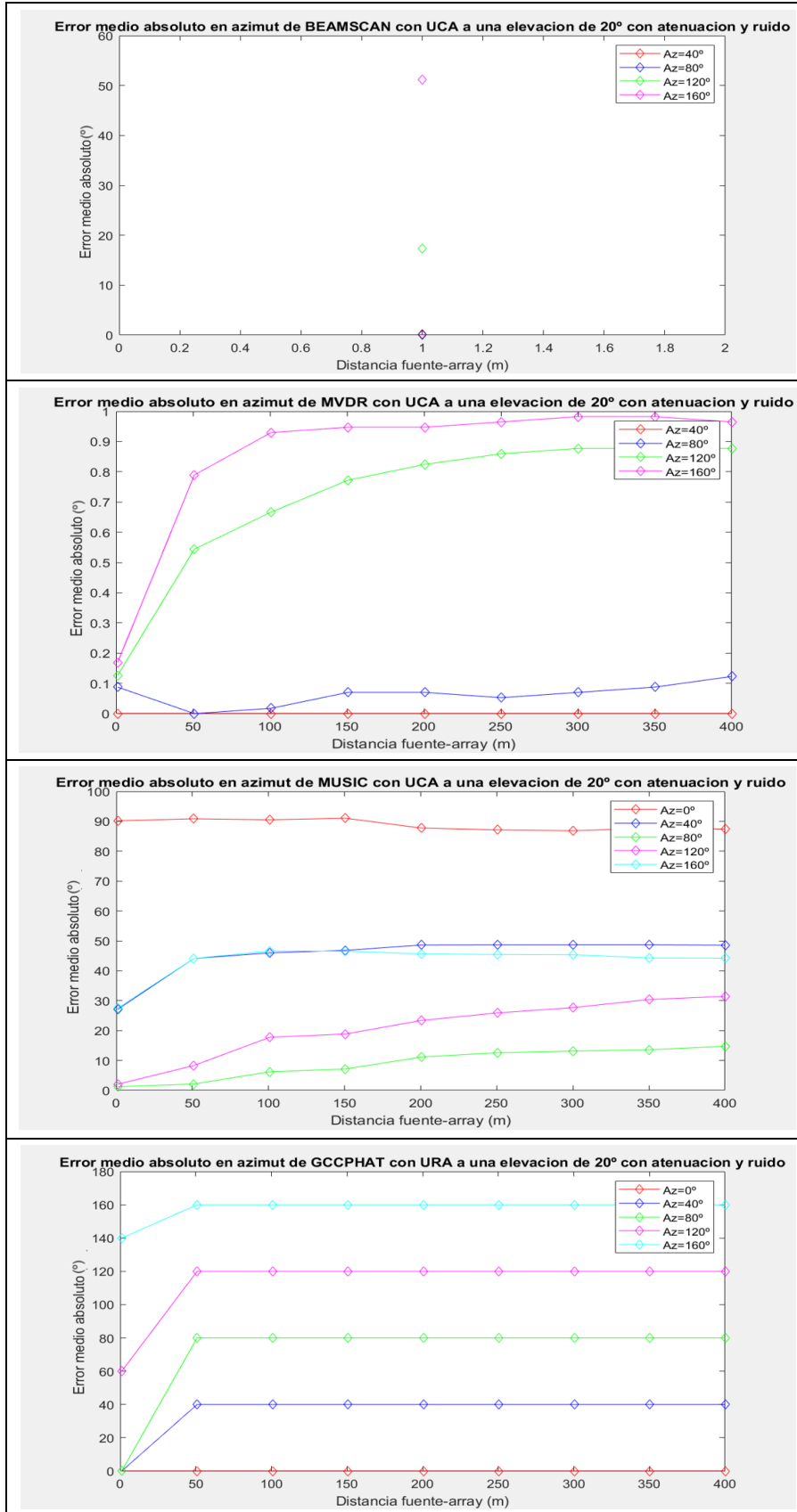


Figura 57 – Representación del error medio absoluto de azimut en función de la distancia entre el dron y el array para distintos ángulos en azimut y a una elevación 20 grados.

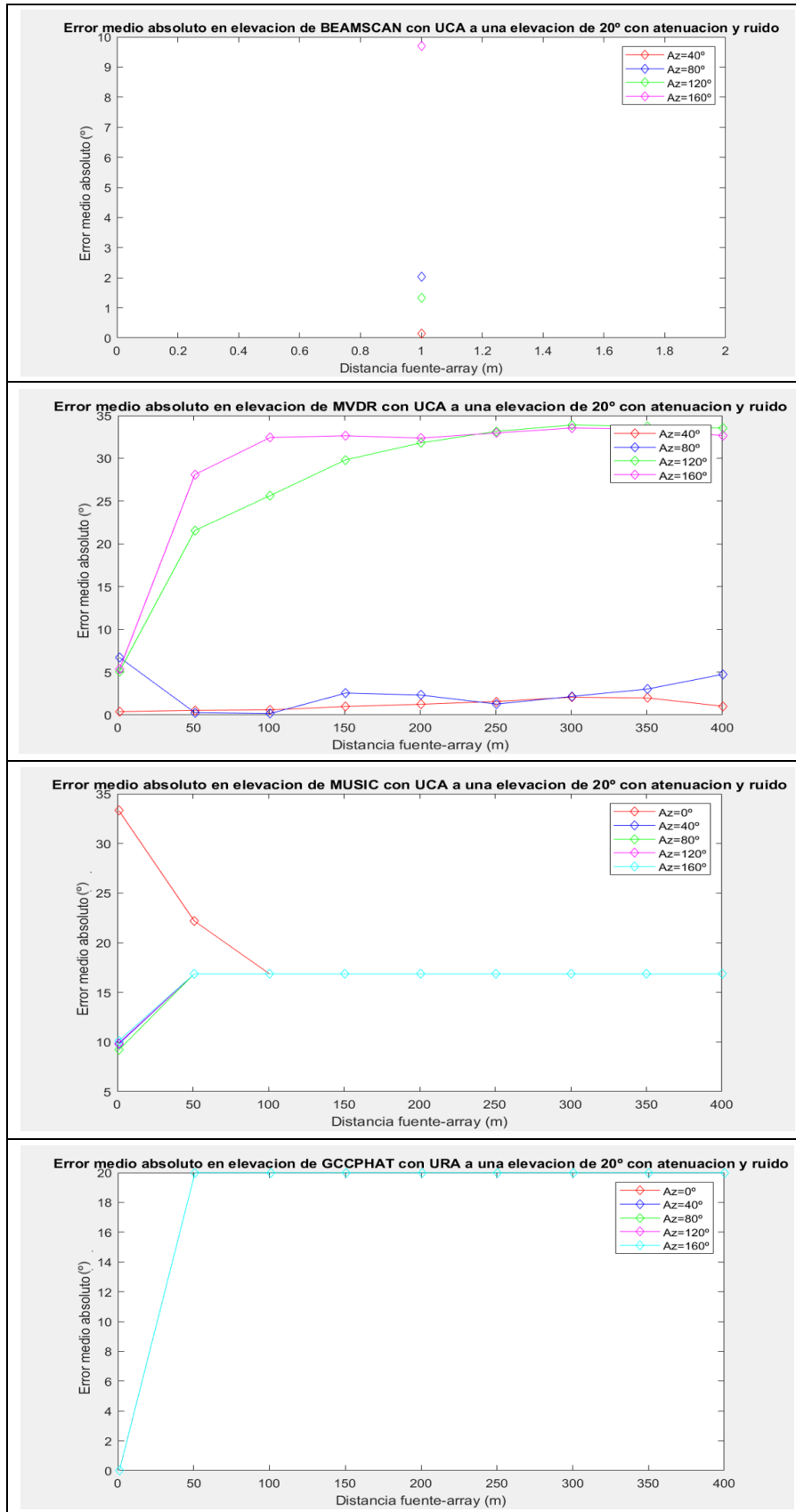


Figura 58 – Representación del error medio absoluto de elevación en función de la distancia entre el dron y el array para distintos ángulos en azimut y a una elevación 20 grados.

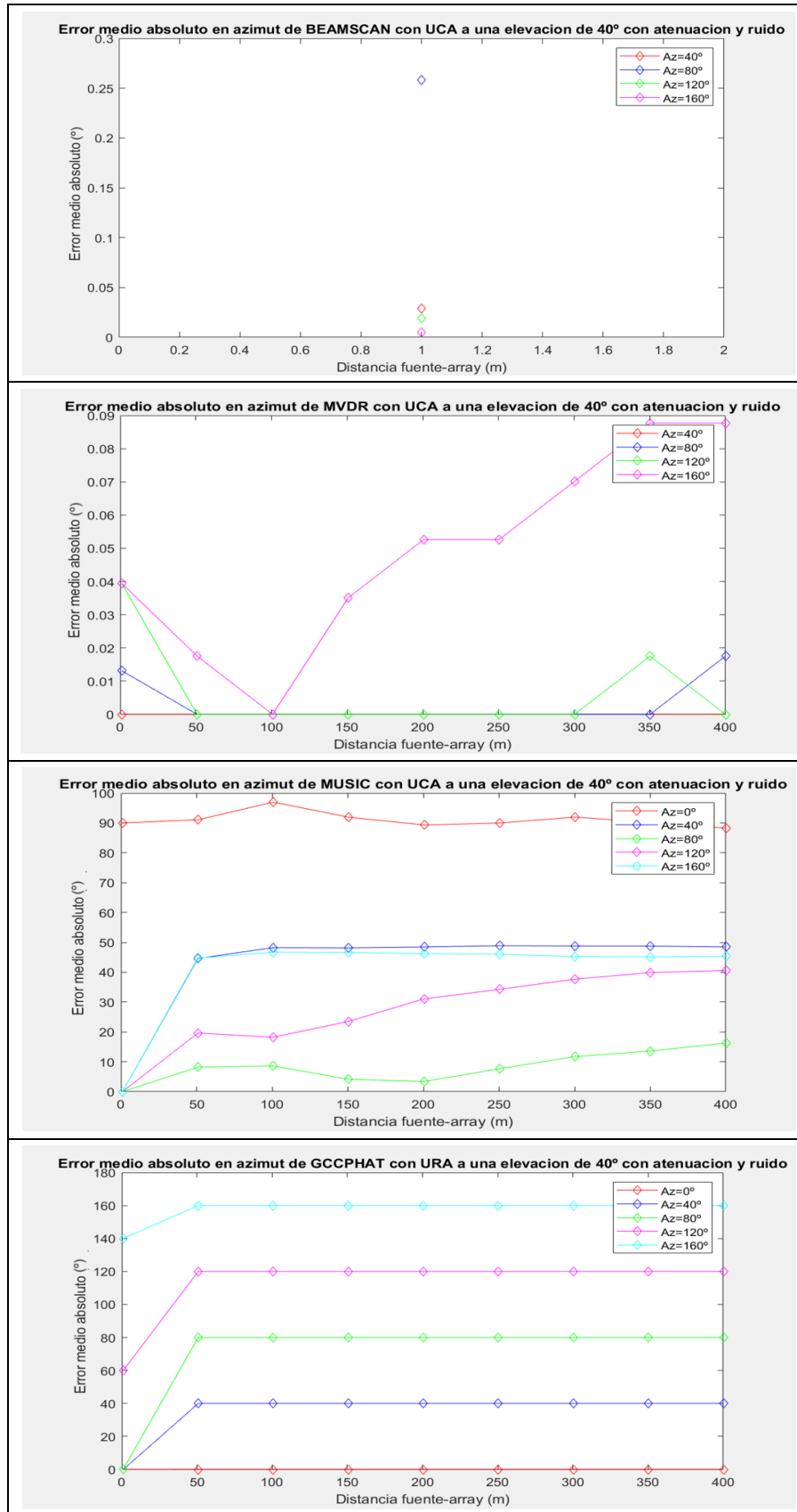


Figura 59 – Representación del error medio absoluto de azimut en función de la distancia entre el dron y el array para distintos ángulos en azimut y a una elevación 40 grados.

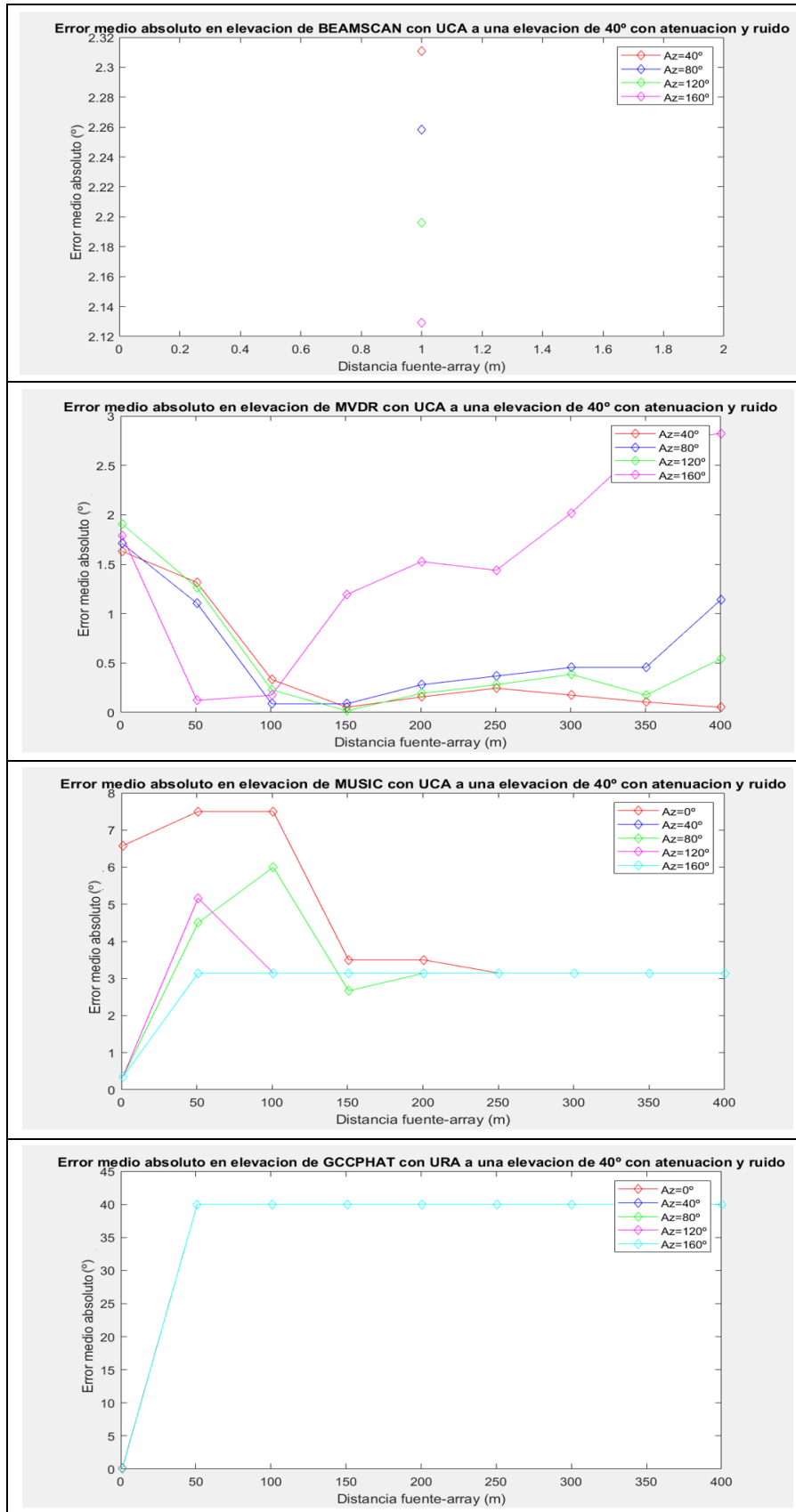


Figura 60 – Representación del error medio absoluto de elevación en función de la distancia entre el dron y el array para distintos ángulos en azimut y a una elevación 40 grados.

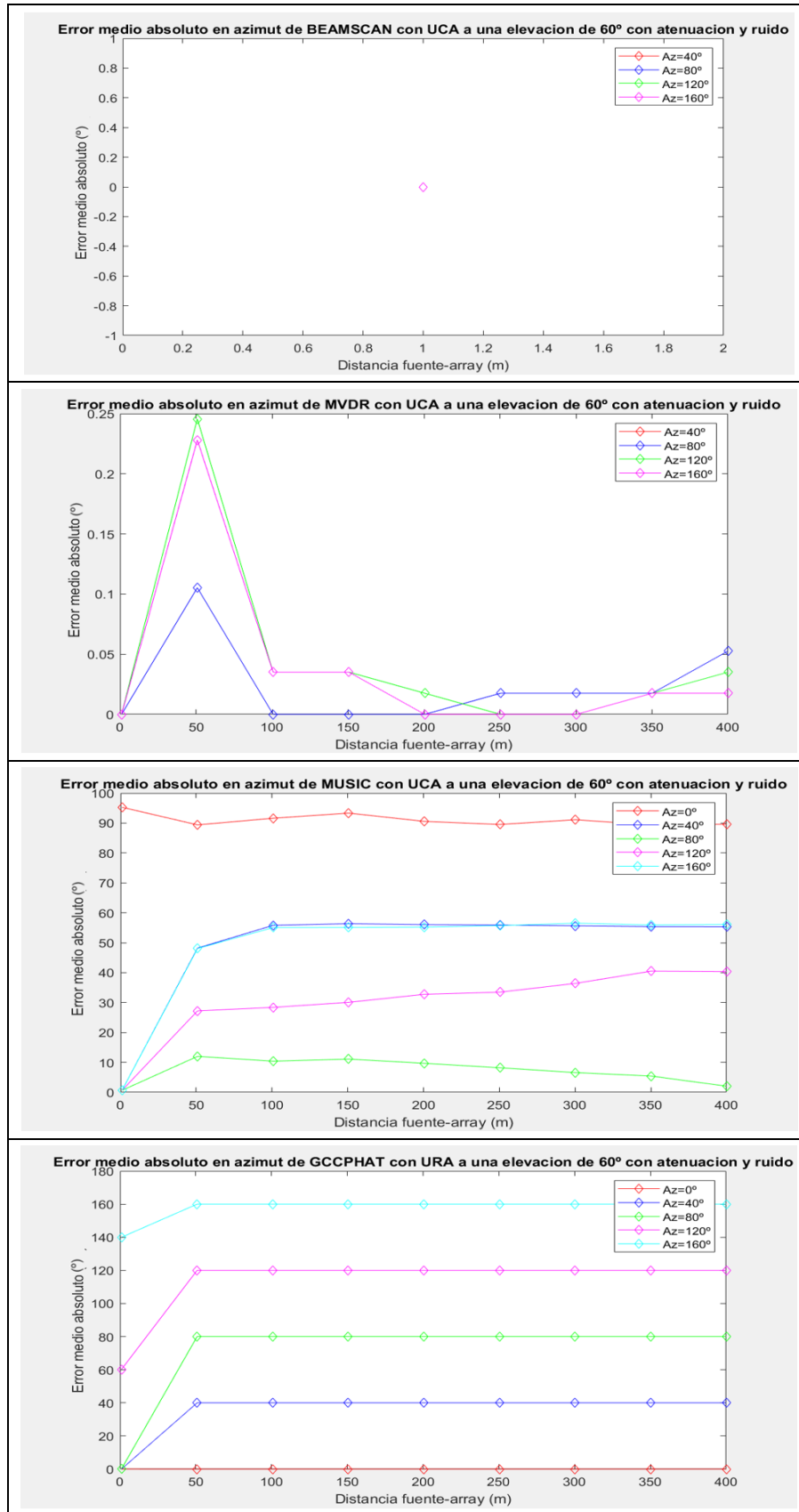


Figura 61 – Representación del error medio absoluto de azimut en función de la distancia entre el dron y el array para distintos ángulos en azimut y a una elevación 60 grados.

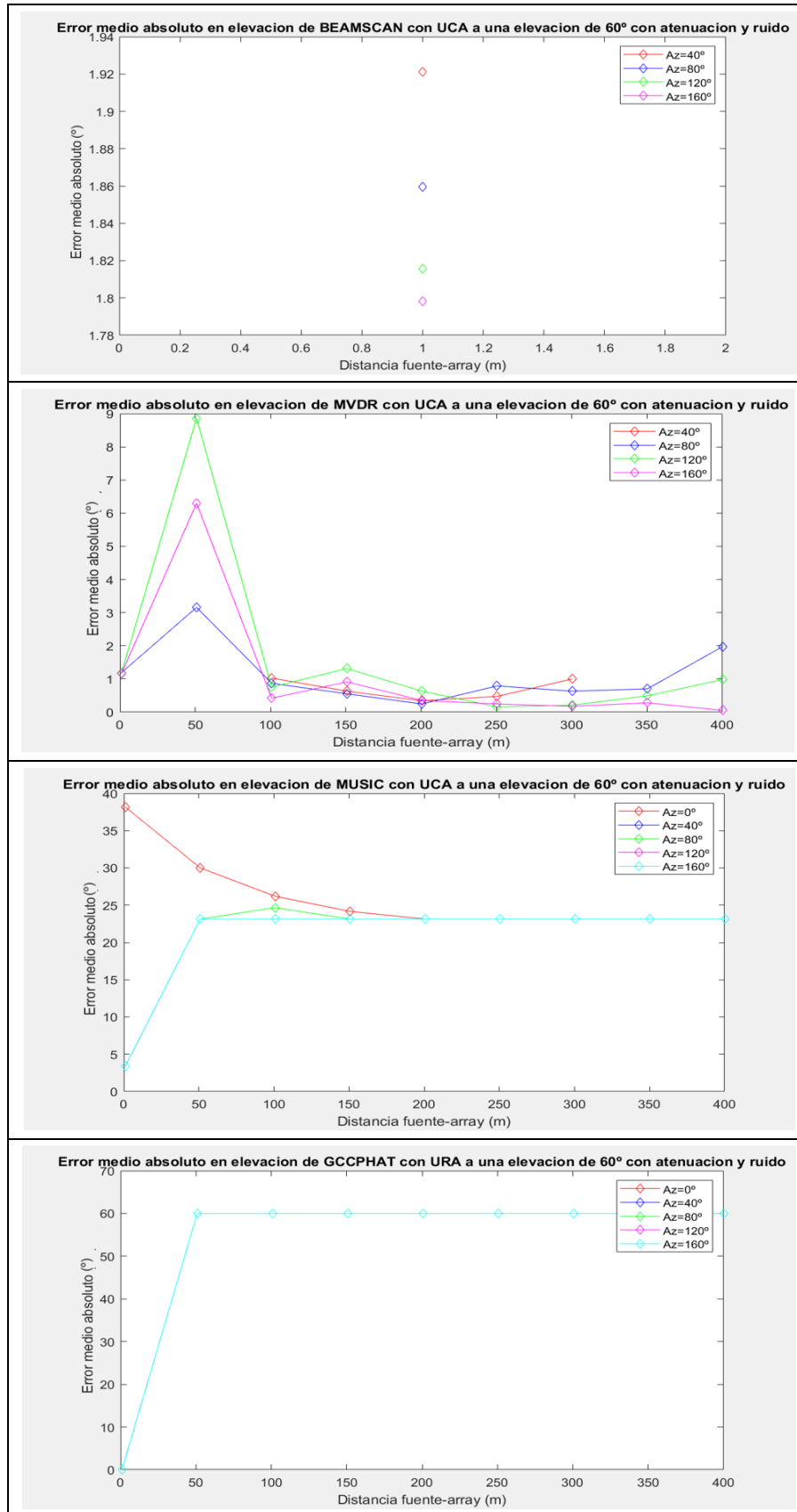


Figura 62 – Representación del error medio absoluto de elevación en función de la distancia entre el dron y el array para distintos ángulos en azimut y a una elevación 60 grados.

En primer lugar, en las Figuras 57 y 58, se observa la representación de los errores de los algoritmos DOA en función de la distancia entre el array de micrófonos y el UAV cuando se tiene una elevación de 20 grados. En estas figuras, se puede observar como, por un lado, Beamscan UCA no obtiene ninguna información cuando la distancia entre el dron y el array de micrófonos es mayor a 1 metro. Por otro lado, MUSIC y GCC-PHAT obtienen errores significativos para cualquier valor de azimut en ambas dimensiones. Y, por último, MVDR obtiene errores cercanos a cero cuando se trata de la estimación del azimut, pero, en la detección en elevación obtiene errores casi nulos para posiciones del UAV en azimut entre 40 y 80 grados, mientras que estima con un error de aproximadamente 30° cuando el dron está posicionado en un azimut entre 120 y 160 grados.

Después, en las Figuras 59 y 60 se observa la representación de los errores medios absolutos respecto a la distancia a una elevación de 40 grados. En estas condiciones, el algoritmo Beamscan sigue sin obtener información para distancias mayores a 1 metro entre el dron y el array de micrófonos. GCC-PHAT estima errores elevados respecto a la dimensión de azimut y no obtiene información en la detección del ángulo de elevación para la mayor parte de posiciones. MUSIC ofrece errores considerables y muy fluctuantes en la estimación del ángulo en azimut, aunque en elevación los errores no sobrepasan los 8° . Y, por último, MVDR obtiene errores casi nulos para el ángulo en azimut y menores de 3° en elevación, incluso a distancias de 400 metros con un 10% de ruido.

Finalmente, cuando el UAV se encuentra en una elevación elevada (60 grados), los resultados de los algoritmos son semejantes a los casos anteriores de elevación baja y media (20 y 40 grados, respectivamente). Por consiguiente, se descarta la utilización del algoritmo Beamscan debido a su imposibilidad de detectar el dron en condiciones más cercanas a la realidad. Asimismo, se descarta GCC-PHAT, al ofrecer un error grande en la detección de la posición del UAV en azimut, ya que el error aumenta a medida que el dron está en posiciones de azimut mayores, y, en el escaneo en elevación no es posible obtener una estimación de la posición de la aeronave en la mayoría de las posiciones espaciales. MUSIC, también se va a descartar, debido a que falla en la detección del UAV en cuanto a la estimación en elevación, y, por otro lado, en azimut los errores son grandes. Por el contrario, MVDR ofrece buenos resultados en la mayoría de los escenarios, independientemente de la elevación y el azimut.

Por lo tanto, debido a las conclusiones anteriores, para llevar a cabo la detección de una aeronave no tripulada de pequeñas dimensiones se va a utilizar el algoritmo MVDR junto con la geometría circular uniforme.

El DOA seleccionado, junto a su geometría, se ha realizado con los valores obtenidos en el apartado 1 de este capítulo, es decir, utilizando el algoritmo MVDR junto con un array circular uniforme que consta de 4 micrófonos y una distancia entre ellos de 0,3 metros. Sin embargo, puesto que el rendimiento de la detección del ángulo de llegada puede haber variado al cambiar las condiciones ideales por unas de ambiente real, se va a volver a realizar un análisis en el que se prueban las distintas configuraciones del array UCA.

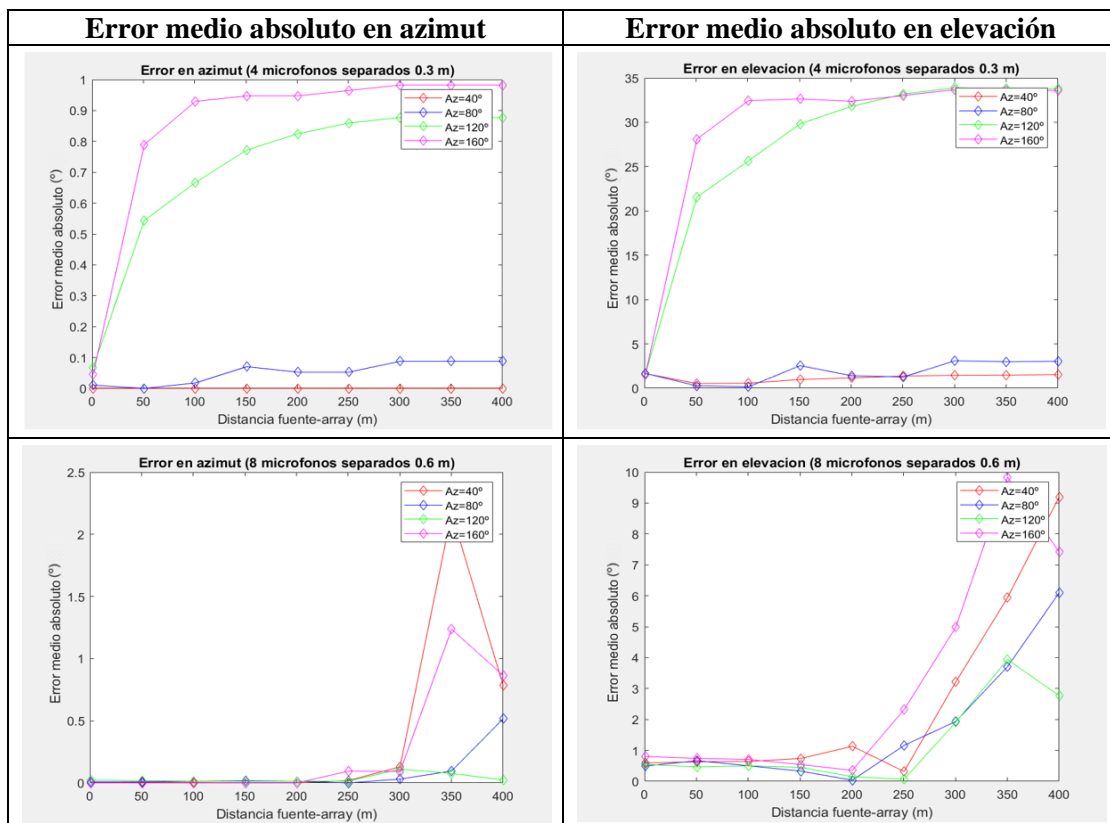
Para aligerar el proceso, en lugar de repetir el estudio para todas las posibles configuraciones, se van a escoger aquellas configuraciones donde para cada número de micrófonos se haya obtenido la distancia entre sensores que produjera un menor error en condiciones ideales teniendo en cuenta ambas dimensiones (azimut y elevación). Es decir, por ejemplo, si se tuviera la tabla de la Figura 54, cada fila representa un número distinto de micrófonos y cada columna una distancia entre ellos. Entonces, se escoge en cada fila, la columna donde el error sea mínimo teniendo en cuenta ambas dimensiones. Así se consiguen 7 configuraciones distintas en lugar de las 49 anteriores. Siendo estas:

Capítulo 6 – Análisis y resultados

- 4 micrófonos separados una distancia de 0.3 metros.
- 6 micrófonos separados una distancia de 0.55 metros.
- 8 micrófonos separados una distancia de 0.6 metros.
- 10 micrófonos separados una distancia de 0.6 metros.
- 12 micrófonos separados una distancia de 0.3 metros.
- 14 micrófonos separados una distancia de 0.5 metros.
- 16 micrófonos separados una distancia de 0.5 metros.

Con estas configuraciones se vuelve a realizar el estudio anterior, donde se variaba el azimut y la distancia de propagación a elevaciones de 20, 40 y 60 grados. El análisis se va a realizar de tal forma que se van a seleccionar las configuraciones que ofrezcan mejores resultados a elevaciones bajas y medias en las Figuras 63 y 64, y, después, con dicha selección se elegirá la configuración definitiva a elevaciones altas en la Figura 65.

A una elevación de 20 grados, se obtienen los siguientes resultados de la Figura 63. En dicha figura, se representa a la izquierda el error medio absoluto en la obtención de la dirección en azimut y, a la derecha, en elevación. La configuración de 6 micrófonos con distancia de 0,55 metros entre sensores se ha descartado debido a que no ofrecía resultados en la mayoría de los ángulos en azimut.



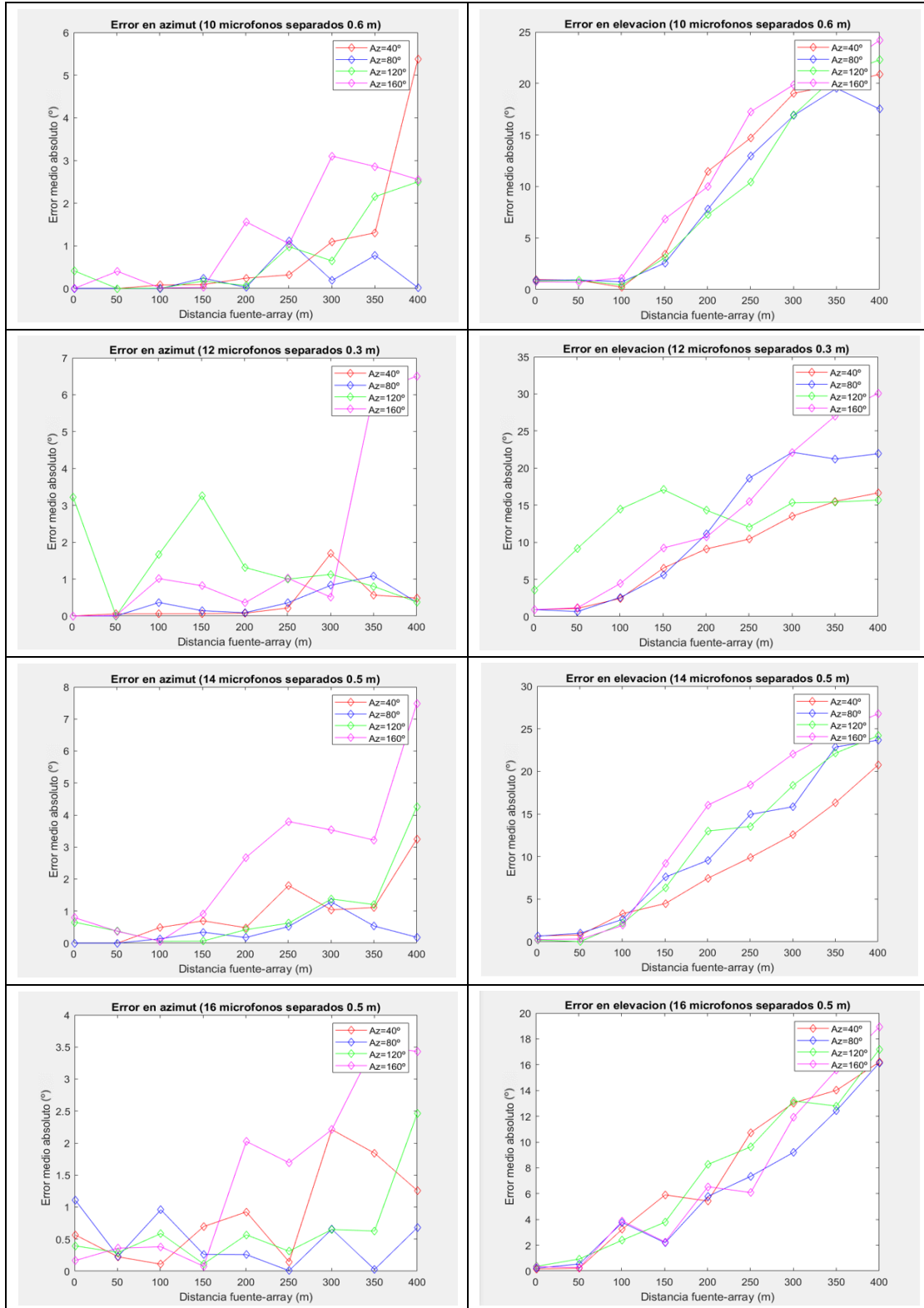
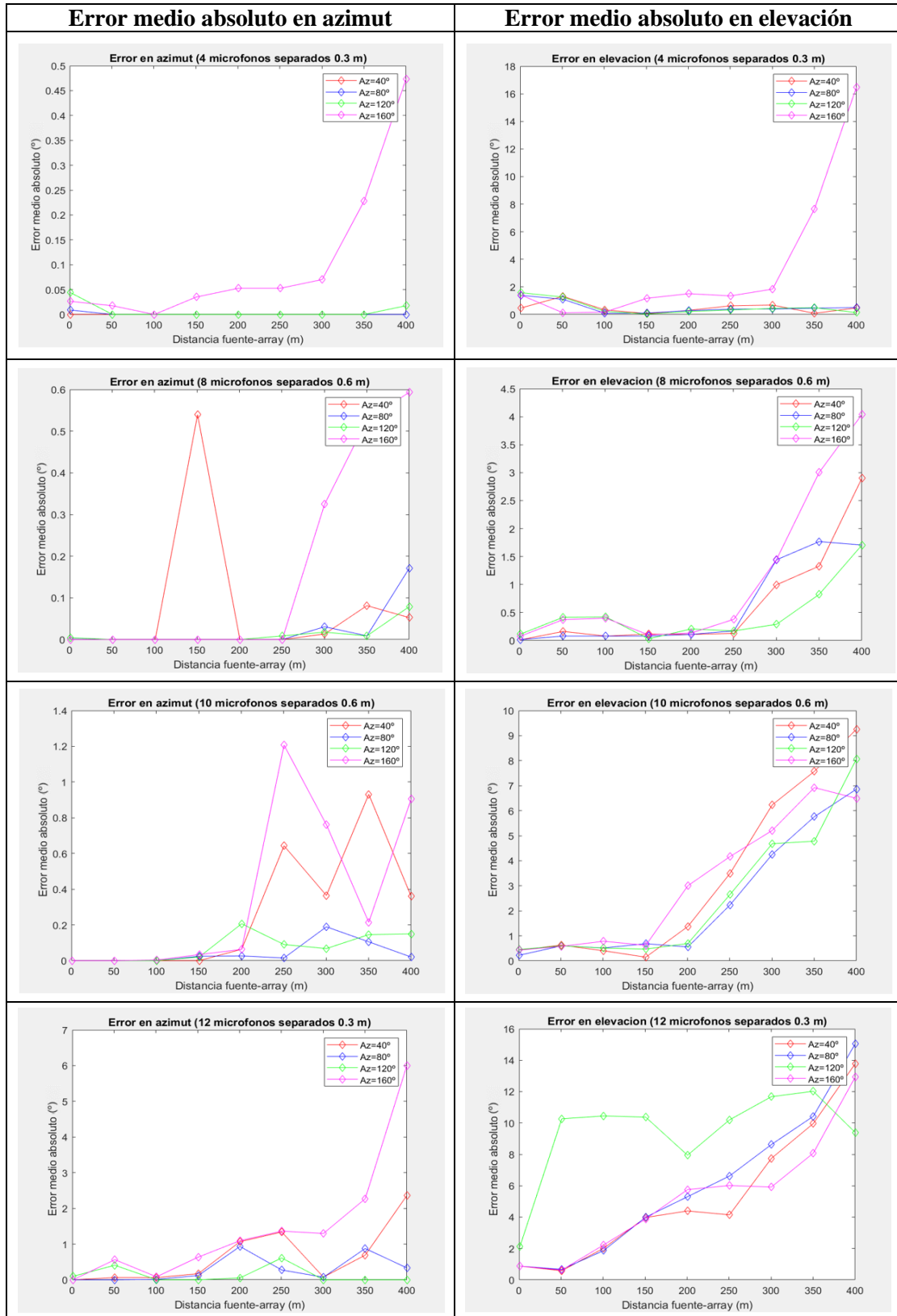


Figura 63 – Representación del error medio absoluto del algoritmo MVDR con UCA en distintas configuraciones en función de la distancia entre el array y el UAV para distintos valores en azimut y a una elevación de 20 grados.

Capítulo 6 – Análisis y resultados

A una elevación de 40 grados se obtiene:



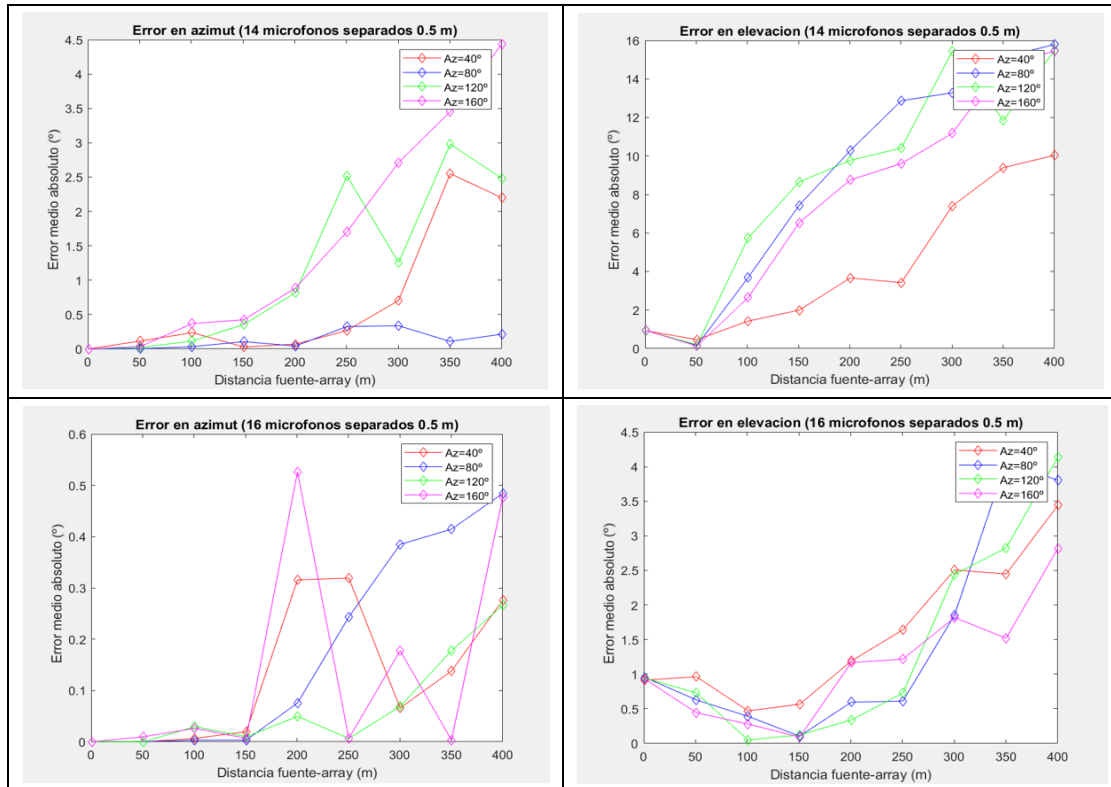


Figura 64 – Representación del error medio absoluto del algoritmo MVDR con UCA en distintas configuraciones en función de la distancia entre el array y el UAV para distintos valores en azimut y a una elevación de 40 grados.

Como se puede observar, aunque en condiciones ideales el análisis en el apartado 6.1 sugiriera que la mejor configuración para el algoritmo MVDR con una geometría UCA fuera la utilización de 4 micrófonos con una distancia entre ellos de 0,3 metros, cuando se enfrenta la señal acústica de un dron a un ambiente real, estas conclusiones cambian.

En las Figuras 63 y 64, a elevaciones de 20 y 40 grados, la estimación de la dirección en cuanto al ángulo de azimut no supone ningún problema, ya que, en todos los casos, el error frente a la dirección real del dron no sobrepasa los 8°, incluso a distancias de 400 metros entre la fuente sonora y el array de micrófonos. Por lo tanto, la selección de la configuración del array vendrá limitada por la detección del ángulo en elevación. A las elevaciones de 20 y 40 grados, la configuración que menor error presenta es la de 8 micrófonos, alcanzando un valor de error medio absoluto máximo en elevación de 10 y 4,5 grados, respectivamente.

Antes de seleccionar definitivamente dicha configuración, se evalúa la respuesta del algoritmo a elevaciones altas, por ejemplo, de 60 grados. En la Figura 65 se visualiza dicha estimación, donde se obtienen errores muy bajos para la mayoría de las direcciones. Por lo tanto, se va a adoptar un array circular uniforme con una configuración de 8 micrófonos separados 0,6 metros y el algoritmo MVDR para la estimación de la dirección de llegada de un dron.

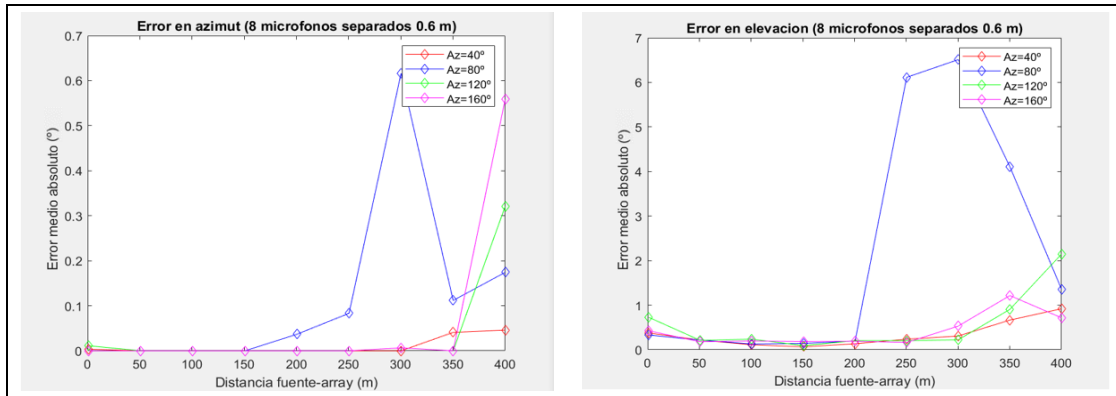


Figura 65 – Representación del error medio absoluto del algoritmo MVDR con UCA en configuración de 8 micrófonos separados 0,6 metros en función de la distancia entre el array y el UAV para distintos valores en azimut y a una elevación de 60 grados.

El estudio llevado a cabo en este capítulo se ha realizado con el dron QAV400, no obstante, se comprueba en la Figura 66 la configuración seleccionada para la detección del ángulo de llegada de un UAV con el dron Parrot AR y Harop Harpy, donde se puede comprobar que la respuesta es similar para los tres drones bajo análisis, debido a que tanto las características, como las componentes frecuenciales, son similares.

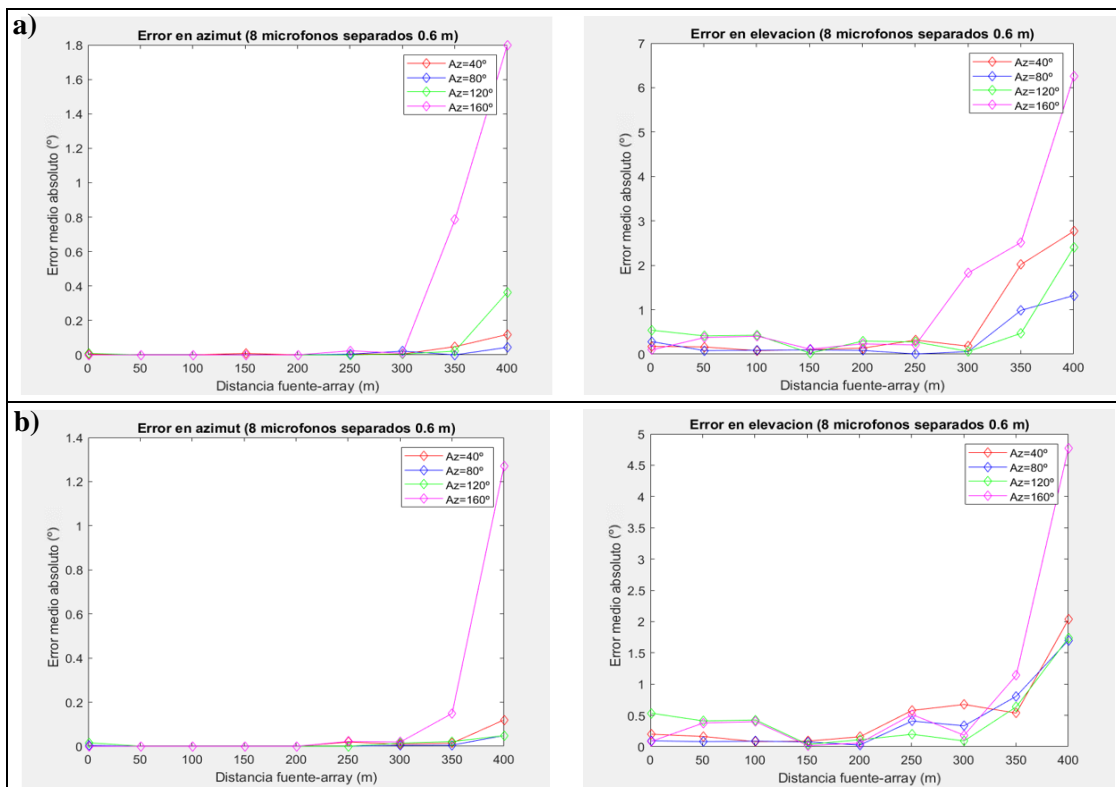


Figura 66 – Representación error del MVDR con UCA de 8 micrófonos y distancia de 0,6 metros en función de la distancia entre el array y el UAV para los drones: a) Parrot AR; b) Harop Harpy.

7 CONCLUSIÓN Y LÍNEAS FUTURAS

7.1 CONCLUSIONES

A lo largo de este trabajo se ha caracterizado el sonido proveniente de un UAV de pequeñas dimensiones, siendo éste el producido por la rotación de las hélices de los motores de un dron. Este movimiento rotacional de las hélices provoca un sonido no estacionario, ya que la velocidad de las hélices del dron va variando en cada momento, aunque con características pseudoaleatorias. Por otra parte, debido a su gran cantidad de componentes frecuenciales, se hace necesaria la obtención de la frecuencia fundamental y sus armónicos, además, de obtener las frecuencias donde el dron tiene mayor amplitud sonora.

En este trabajo se ha conseguido extraer la frecuencia fundamental de cada uno de los drones que se han tenido en cuenta. Esta extracción se ha obtenido mediante el análisis cepstral, una vez estudiada la señal en el dominio temporal, frecuencial y tiempo-frecuencial.

Después, en el siguiente capítulo, se ha estimado la propagación del sonido en ambientes reales, gracias a la norma de la ISO 9613. De esta forma, ha sido posible obtener la atenuación que sufrirá la onda de sonido en función, tanto de la frecuencia, como de la distancia entre el emisor y el receptor acústico. Así, se observa un comportamiento de la atenuación exponencial a medida que aumenta la distancia, y, además, según incrementa la frecuencia, la atenuación también será mayor, por lo tanto, las bandas de frecuencia que deben utilizarse para detectar al dron a larga distancia será las de menor frecuencia posible, es decir, la de la frecuencia fundamental.

Tras analizar la fuente y el canal de propagación, el trabajo se centra en la recepción del sonido, realizando un estudio de los receptores acústicos y la detección de ángulos de llegada. En cuanto a los receptores acústicos, se estudia el elemento principal, el micrófono, con todas sus características, y, se presentan los arrays de micrófonos, en especial los arrays de geometría uniforme (ULA, UCA y URA).

Por otro lado, en la detección de ángulos de llegada, se han presentado los tipos de algoritmos y métodos de detección, explicando, en primer lugar, la conversión entre banda ancha y estrecha necesaria para poder utilizar los algoritmos DOA. Después, se introducen cuatro algoritmos (Beamscan, MVDR, MUSIC y GCC-PHAT), que son los algoritmos con los que se trabajará en la sección siguiente.

Por último, se ha realizado un análisis para seleccionar el array de micrófonos y algoritmo DOA más conveniente en la detección de un dron. Para ello, se utilizan las características obtenidas respecto a la caracterización espectral de la señal y la propagación del sonido. El análisis consta de tres pasos: Primero, se van alternando configuraciones del array, tanto de su geometría como del número de micrófonos y distancia entre ellos, y se calcula el error obtenido por los algoritmos DOA, de forma que se consigue una configuración óptima para cada geometría de array y DOA. Después, con las configuraciones obtenidas, se realiza un barrido en azimut y elevación, y se selecciona una única geometría para cada algoritmo, de forma que sea la que mejor funcione en la mayoría de las posiciones espaciales. Finalmente, se calcula la respuesta de cada configuración en condiciones reales, de atenuación y ruido, y se selecciona un único algoritmo con la configuración del array que mejor resultados obtenga.

De esta forma, se han evaluado todas las posibles configuraciones del sistema. En primer lugar, en cuanto al array de micrófonos, el ULA se descarta directamente, debido a que no ofrece ninguna información acerca del ángulo en elevación. Después, el array rectangular uniforme estima mejores resultados que el UCA hasta los 90° en azimut, sin embargo, a partir de dicho valor, el error crece vigorosamente y, es por ello, por lo que se ha escogido una configuración circular, ya que ofrece errores bajos y constantes para la mayoría de los escenarios. En cuanto al algoritmo DOA, Beamscan se descarta debido a su falta de información en un escenario con atenuación y ruido, más parecido a las condiciones reales. Luego, GCC-PHAT y MUSIC también se descartan ya que éstos estimaban la posición del dron con menor precisión que MVDR.

Tras elegir el sistema array-DOA, se vuelve a examinar todas las configuraciones de números de micrófonos y distancia entre ellos, de forma que finalmente, la configuración que obtiene los mejores resultados es el array circular uniforme con una configuración de 8 micrófonos separados 0,6 metros entre ellos y el algoritmo de estimación de la dirección de llegada MVDR. En definitiva, se ha logrado obtener la mejor configuración array-DOA para las características frecuenciales de un UAV, teniendo en cuenta la propagación acústica por un ambiente real.

7.2 LÍNEAS FUTURAS

En cuanto a las líneas futuras de este estudio, en primer lugar, en vez de limitarse a los arrays de estructuras geométricas uniformes, como los que se han estudiado en este trabajo (array circular uniforme, array lineal uniforme y array rectangular uniforme) se propone la investigación de estructuras emergentes de arrays de micrófonos, como los arrays co-primos y los arrays anidados. Este tipo de arrays dispersos proporcionan una apertura de matriz más grande que el ULA con el mismo número de sensores para mejorar la resolución. Más importante aún, las matrices dispersas permiten romper la limitación de los grados de libertad [49].

Después, debido a las propiedades de las señales de audio tomadas de un ambiente real, donde las fuentes emiten sonidos simultáneamente y los obstáculos en el camino producen ecos y reverberación [50], se deberá completar el sistema con algoritmos de estimación del número de fuentes sonoras.

Por otro lado, se propone el estudio de la ganancia que se obtiene si cada micrófono es sustituido por un array que aumente la directividad y la ganancia, con el objetivo de poder determinar la dirección de la fuente con mayor fiabilidad.

Por último, se plantea el estudio de la combinación de arrays para así, no sólo determinar la dirección de la fuente, sino también su posición, de forma que sea posible la estimación de la fuente sonora en 3D.

ANEXOS

I. PLIEGO DE CONDICIONES

En esta sección se van a introducir las condiciones generales, de materiales y equipos y de ejecución para la elaboración de este proyecto.

El objetivo principal de este trabajo consiste en diseñar un array de micrófonos que use un algoritmo básico de estimación de llegada, para poder conocer la posición de un UAV gracias a su huella acústica. Para llevar a cabo la elaboración de este trabajo, se utiliza la herramienta *MATLAB*.

Un dron o UAV de tamaño reducido es un vehículo aéreo no tripulado que se utiliza tanto en el ámbito militar como civil. La detección de un UAV se hace necesaria debido al rápido desarrollo de estas aeronaves, y la utilización de éstas para realizar actos ilícitos.

Para llevar a cabo la detección en este proyecto, en primer lugar, se obtiene vía Internet ejemplos del sonido que genera un dron en movimiento. Dicho sonido se analiza para poder obtener las características espectrales de un UAV, de forma que se pueda obtener la frecuencia fundamental, y, así poder tener caracterizadas las componentes frecuenciales del UAV, ya que es una señal rica en armónicos. Con este objetivo, el capítulo 2 comienza analizando la señal en el dominio del tiempo, después, la señal se transforma al dominio de la frecuencia, y, al dominio tiempo-frecuencial. Finalmente, se comparan dos métodos de obtención de la frecuencia fundamental, la autocorrelación y el análisis cepstral, y se utiliza este último para poder caracterizar finalmente las componentes frecuenciales del dron.

Después, se simula un escenario aeroportuario para poder caracterizar la propagación del sonido en este ambiente, y, así, poder obtener la atenuación que sufre el sonido debido a varios efectos como la divergencia geométrica, la absorción atmosférica, el efecto del suelo, el apantallamiento por efectos de barreras y otros efectos similares. Esta parte del trabajo se realiza bajo las normas de la Organización Internacional de Normalización (ISO), en concreto, la norma 9613, que hace referencia a la atenuación del sonido durante la propagación en exteriores.

Más adelante, en el capítulo 4, se definen los micrófonos como receptores acústicos, y se tratan aspectos referentes a sus características y a los arrays de micrófonos de geometrías uniformes. Posteriormente, en el capítulo 5, se definen los algoritmos de detección de llegada. En este punto se estudia un algoritmo de banda ancha, GCC-PHAT (*Generalized Cross Correlation with Phase Transform*), y, algoritmos que reciben como entrada una señal de banda estrecha. Estos últimos se basan en la estimación del espectro espacial. Dicha estimación se divide en dos ramas, la primera, beamforming, donde se estudian el beamformer convencional y MVDR (*Minimum Variance Distorsionless Response*), y, la otra rama consiste en técnicas basadas en el subespacio, donde se estudia el algoritmo MUSIC (*Multiple Signal Classification*).

Por último, se realiza un análisis en el que, mediante el audio obtenido de tres drones distintos, se selecciona el array de micrófonos que mejor se ajuste para cada algoritmo DOA y, finalmente, se selecciona la configuración array y algoritmo DOA que mejores resultados obtenga. Todo ello teniendo en cuenta las características espectrales y atenuantes calculadas a lo largo del proyecto.

Para abordar la realización de este trabajo, se han logrado una serie de objetivos:

- 1) Estudio espectral de los sonidos generados por drones. En dicho estudio se ha analizado la no estacionariedad del sonido en el dominio del tiempo, la distribución frecuencial en el dominio de la frecuencia, las líneas HERM en el dominio tiempo-frecuencia, y la frecuencia fundamental gracias al análisis cepstral.
- 2) Estudio de la propagación del sonido generado por drones en ambientes reales y selección de las bandas de frecuencia más favorables para la detección. Para ello, se ha escogido un escenario aeroportuario donde se consigue caracterizar la atenuación en función de la distancia para distintas bandas de frecuencia. Siendo las bandas de menor frecuencia, las que menor atenuación poseen.
- 3) Diseño de un array de micrófonos para la localización del dron a partir de la señal en la banda seleccionada. Para cumplir este objetivo se han estudiado los receptores acústicos usando la toolbox ‘*Phase Array*’ de *MATLAB*, y se ha seleccionado un array circular uniforme (UCA) con un número de 8 micrófonos, como resultado final.
- 4) Estudio de algoritmos de localización de la fuente de sonido a partir de las señales captadas por los micrófonos que conforman el array. En este punto, se han estudiado los algoritmos clásicos Bartlett beamforming, MVDR, MUSIC y GCC-PHAT.
- 5) Obtención del mejor resultado de detección teniendo en cuenta todos los aspectos comentados en los cinco puntos anteriores. Para ello, se ha realizado un estudio que consta de tres partes, en el que finalmente, tras analizar todas las configuraciones distintas, se ha seleccionado un array circular uniforme de 8 micrófonos separados una distancia de 0,6 metros, junto con el algoritmo MVDR para la detección acústica de un UAV.

a) Requisitos de hardware y software

Este proyecto se realiza mediante el software *MATLAB*. Por lo tanto, para poder ejecutar dicho programa en su última versión se necesita un ordenador que disponga de ciertos elementos según los requerimientos que aparecen en la página web de dicho programa [24]. Estos requisitos dependen de si el software del ordenador es de Windows, MacOS o Linux y se definen en la siguiente tabla.

	WINDOWS	MACOS	LINUX
Sistema operativo	Windows 10 (versión 1803 o superior)	macOS Big Sur (11)	Ubuntu 20.04 LTS Ubuntu 18.04 LTS
	Windows 7 Service Pack 1	macOS Catalina (10.15)	Ubuntu 16.04 LTS Debian 10
	Windows Server 2019	macOS Mojave (10.14)	Red Hat Enterprise Linux 8 (mínimo 8.1) Red Hat Enterprise Linux 7 (mínimo 7.6)
	Windows Server 2016		SUSE Linux Enterprise Desktop 12 (mínimo SP2)
			SUSE Linux Enterprise 15
			SUSE Linux Enterprise Server 12 (mínimo SP2)
SUSE Linux Enterprise Server 15			
Procesador	Mínimo: Cualquier procesador de Intel o AMD x86-64	Mínimo: Cualquier procesador Intel x86-64	Mínimo: Cualquier procesador de Intel o AMD x86-64
	Recomendado: Cualquier procesador de Intel o AMD x86-64 con 4 núcleos lógicos y soporte AVX2	Recomendado: Cualquier procesador Intel x86-64 con 4 núcleos lógicos y soporte AVX2	Recomendado: Cualquier procesador de Intel o AMD x86-64 con 4 núcleos lógicos y soporte AVX2
Disco	Mínimo: 3.4 GB de espacio HDD para <i>MATLAB</i> únicamente (5-8 GB para una instalación típica)	Mínimo: 3 GB de espacio SDD para <i>MATLAB</i> únicamente (5-8 GB para una instalación típica)	Mínimo: 3.3 GB de espacio HDD para <i>MATLAB</i> únicamente (5-8 GB para una instalación típica)
	Recomendado: SDD	Recomendado: SDD	Recomendado: SDD
RAM	Mínimo: 4 GB	Mínimo: 4 GB	Mínimo: 4 GB
	Recomendado: 8 GB	Recomendado: 8 GB	Recomendado: 8 GB
Gráfica	No necesita tarjeta específica	No necesita tarjeta específica	No necesita tarjeta específica

 Figura 67 – Tabla de requisitos para la ejecución de *MATLAB*.

Los requisitos anteriores son los necesarios para la ejecución adecuada de la herramienta *MATLAB*. Además, para que funcione todo adecuadamente, la versión de instalación debe ser la versión 2016b o posterior, y contener la caja de herramientas ‘*Phased Array System Toolbox*’, para poder tener acceso a los arrays de micrófonos y los algoritmos de detección de llegada. Asimismo, la herramienta *MATLAB* debe tener descargados todos los paquetes iniciales, que vienen ya seleccionados en la descarga del programa, para poder analizar la señal espectralmente, específicamente es necesaria la herramienta ‘*Signal Processing Toolbox*’.

Por último, para poder trabajar con las señales acústicas de drones se debe realizar la descarga de varios archivos de audio mediante internet. En el desarrollo de este trabajo se han descargado los audios de los UAVs: QAV400, Harop Harpy y Parrot AR gracias a la investigación desarrollada por Gonçalo F. Cardoso y Ruben Pater en “A Study into 21st Century Drone Acoustics” [47].

II. PRESUPUESTO

El presupuesto total del proyecto se calcula teniendo en cuenta varios aspectos: primero, se estiman el inmovilizado material, el inmovilizado intangible y los servicios exteriores para calcular el coste total del material necesario. Después, a dicho coste se le suma la mano de obra, obteniendo el presupuesto de ejecución material. Gracias a este presupuesto, junto a los gastos generales y el beneficio industrial, se calcula el presupuesto de contrata. Finalmente, se obtiene el presupuesto total, el cual viene determinado por el de contrata, además de los honorarios técnicos correspondientes.

a) Presupuesto de ejecución material

El presupuesto de ejecución material se calcula como el coste total del material necesario más la mano de obra.

a1) Coste total del material

Dentro del coste total del material necesario se encuentran: el inmovilizado material, el cual se refiere a los dispositivos físicos que se han utilizado para la elaboración del trabajo; el inmovilizado intangible, es decir, los programas software necesarios para la ejecución, y; una serie de servicios exteriores, como son la luz y la red telefónica, para el uso de Internet.

Inmovilizado material

En cuanto al inmovilizado material, han sido utilizados los siguientes materiales físicos:

- Una pantalla adquirida para la realización del trabajo, que posteriormente se utilizará en futuros ejercicios, valorada a 129 €. Ésta se amortizará por el método de cuotas constantes con una vida útil de 2 años (24 meses).

$$\text{Cuota de amortización mensual (pantalla)} = \frac{129 \text{ €}}{24 \text{ meses}} = 5,375 \text{ €/mes} \quad (78)$$

Debido a que la durabilidad del trabajo fue de 4 meses, se incorpora para el presupuesto de éste un total de 21,5 €.

- Un portátil por valor de 899 € para el cual se usará el método de amortización por trabajo realizado. En este caso, se usó el dispositivo un total de 552 horas, y, se estima una vida útil igual a 10.000 horas. Por tanto, su cuota de amortización corresponde:

$$\text{Cuota de amortización 552 horas (portátil)} = \frac{899 \text{ €} \cdot 552 \text{ horas}}{10000 \text{ horas}} = 49,6248 \text{ €} \quad (79)$$

- Se compró mobiliario por un valor de 300 €, que se amortizarán en este caso por el método de amortización lineal durante los próximos 3 años (36 meses).

$$\text{Cuota mensual de amortización (mobiliario)} = \frac{300 \text{ €}}{36 \text{ meses}} = 8,33 \text{ €/mes} \quad (80)$$

Por lo tanto, para los 4 meses de trabajo, la amortización acumulada del mismo será de 33,33 €.

Elemento	Método de amortización	Precio de adquisición	Vida útil	Cuota mensual	Coste del trabajo
Portátil	Por trabajo realizado	899 €	10.000 horas	-	49,62 €
Pantalla	Lineal	129 €	2 años	5,375 €	21,5 €
Mobiliario	Lineal	300 €	3 años	8,33 €	33,33 €
TOTAL					104,45€

*Los valores han sido redondeados a 2 decimales

Figura 68 – Tabla de costes del inmovilizado material.

Inmovilizado intangible

En el apartado de aplicaciones informáticos, se han utilizado dos programas *Windows 10 professional* y *MATLAB*, adquiridos a un valor de 50 € y 3000 €, respectivamente. Estos programas se van a amortizar al plazo máximo permitido por las normas contables españolas de 5 años (60 meses). En ambos casos, la amortización será lineal debido a que se usarán estas licencias en futuros proyectos.

$$\text{Cuota de amortización mensual (MATLAB)} = \frac{3000 \text{ €}}{60 \text{ meses}} = 50 \text{ €/mes} \quad (81)$$

$$\text{Cuota de amortización mensual (Windows)} = \frac{50 \text{ €}}{60 \text{ meses}} = 0,83 \text{ €/mes} \quad (82)$$

Con lo cual, *MATLAB* tendrá una amortización total de 200 € y *Windows* de 3,32 € durante los 4 meses del presupuesto del trabajo.

Elemento	Método de amortización	Precio de adquisición	Vida útil	Cuota mensual	Coste para el trabajo
Windows 10 professional	lineal	50 €	5 años	0,83 €	3,32 €
Matlab	Lineal	3000 €	5 años	50 €	200 €
TOTAL					203,32 €

*Redondeado a 2 decimales

Figura 69 – Tabla de costes del inmovilizado intangible.

Servicios exteriores

En cuanto a los servicios exteriores, se van a incluir gastos por suministros correspondientes a la luz y la red telefónica utilizada.

En primer lugar, la luz necesaria para la ejecución del trabajo se paga en facturas bimensuales, donde la primera cuota es de 59,62 € y la segunda de 63,24 €. Por lo tanto, el total del coste de la luz en estos 4 meses de trabajo asciende a un total de 122,86 €.

Por otro lado, la red telefónica es mensual fija de 60 €, por lo que se incorpora un total de 240 € equivalente a la duración del trabajo.

Elemento	Cuota	Coste para el trabajo
Luz	Bimensual (59,62 € ; 63,24 €)	122,86 €
Red telefónica	Mensual (60 €)	240 €
TOTAL		362,86 €

Figura 70 – Tabla de costes de suministros

Anexos: Presupuesto

Finalmente, la suma del total de las Figuras 68-70, que corresponden con el inmovilizado material, el inmovilizado intangible y los suministros, corresponde con el coste total del material, representado en la tabla de la Figura 71.

Inmovilizado material	104,45 €
Inmovilizado intangible	203,32 €
Servicios exteriores	362,86 €
COSTE TOTAL DEL MATERIAL	670,63 €

Figura 71 – Coste total del material

a2) Mano de obra

En cuanto a la mano de obra necesaria para la elaboración de este trabajo, se recoge en la siguiente figura en base a las horas trabajadas.

Actividad desempeñada	Coste por hora	Número de horas trabajadas	Coste
Dirección	25 €	300	7500 €
Desarrollo	20 €	552	11.040 €
Redacción	8 €	100	8.000 €
TOTAL			26.540€

Figura 72 – Tabla de costes de la mano de obra.

Finalmente, el presupuesto de ejecución material será:

Coste total del material	670,63 €
Mano de obra	26.540 €
PRESUPUESTO DE EJECUCIÓN MATERIAL	27.210,63 €

Figura 73 – Presupuesto de ejecución material

b) Presupuesto por contrata

El presupuesto por contrata se determina sumando al presupuesto de ejecución material, los gastos generales y el beneficio industrial. En este caso, se va a considerar que el beneficio y los gastos industriales suponen un 19% del presupuesto de ejecución material, correspondiendo un 6% al beneficio industrial y un 13% a los gastos generales.

Presupuesto de ejecución material	27.210,63 €
6% del presupuesto de ejecución material (beneficio industrial)	1.632,64 €
13% del presupuesto de ejecución material (gastos generales)	3.537,381 €
Presupuesto de contrata del proyecto	32.380,65 €

*Redondeado a 2 decimales

Figura 74 – Presupuesto de contrata del proyecto.

c) Presupuesto total del proyecto

Por último, el presupuesto total del proyecto se calcula como la suma del presupuesto por contrata, añadiendo a este el 10 % del IVA, y los honorarios técnicos correspondientes. En este caso, se han estimado unos honorarios técnicos de 1500 € a los cuales se les debe incorporar el 21 % de IVA.

	Coste sin incluir el IVA	IVA incorporado	Coste con el IVA incorporado
Presupuesto de contrata del proyecto	32.380,65 €	10%	35.618,72 €
Honorarios técnicos	1500 €	21%	1.815 €
PRESUPUESTO TOTAL DEL PROYECTO			37.433,72 €

Figura 75 – Presupuesto total del proyecto.

III. MANUAL DE USUARIO

En este apartado se incluyen los programas y códigos necesarios para poder ejecutar el sistema de detección de un UAV mediante señales acústicas. Para ello, primero se va a introducir la forma en la que se han conseguido los audios de los drones, y, después, la ejecución de ciertos *scripts* mediante la herramienta *MATLAB* para lograr caracterizar espectralmente el sonido del dron, la propagación del sonido en ambientes reales y el algoritmo de detección de la señal sonora en conjunto con los receptores acústicos necesarios para poder detectar dicho UAV.

a) Señal acústica UAVs

Para la elaboración de este trabajo se han seleccionado varios audios de drones obtenidos vía Internet. Mediante esta vía, se ha utilizado la investigación ‘A Study into 21st Century Drone Acoustics’ del compositor Gonçalo F. Cardoso y el diseñador Ruben Pater en [47]. En esta investigación auditiva, se grabaron hasta 17 tipos de drones para poder escuchar el tipo de motores que tienen los drones, el sonido que producen y los efectos psicológicos de estos sonidos en áreas de conflicto.

Por lo tanto, para poder realizar este estudio, se han seleccionado tres tipos de drones, en los que se encuentran QAV400 y Parrot AR como drones cuadricópteros de uso comercial, y, Harop Harpy, como un dron de uso militar.

b) *MATLAB*

Con el objetivo de poder obtener todos los cálculos, algoritmos, gráficos e imágenes se ha utilizado la herramienta *MATLAB*.

MATLAB combina un entorno de escritorio ajustado para procesos de diseño y análisis iterativos con un lenguaje de programación que expresa matrices y arrays matemáticos directamente. Las herramientas de *MATLAB* se desarrollan profesionalmente, y se prueban rigurosamente [24]. Asimismo, proporciona diferentes algoritmos los cuales se pueden ajustar para lograr el resultado deseado.

Para poder hacer uso de este software se necesita de una licencia, cuyo precio se indica en el capítulo de ‘Presupuesto’. Una vez se tiene la licencia, se debe instalar el programa seleccionando una versión a instalar. En la elaboración de este trabajo se han utilizado las versiones R2019b y R2020b, aunque será posible ejecutar los documentos con versiones a partir de la R2016b. Durante dicha instalación se pedirá la selección de cajas de herramientas (*toolbox*) que se deseen descargar, en este caso, además de las básicas que ya vienen incluidas, se ha utilizado las herramientas de la *toolbox* ‘*Phased Array System Toolbox*’, por lo que será necesaria su descarga.

Una vez se tiene instalado *MATLAB*, para poder ejecutar los *scripts* de la siguiente sección, es necesario que los audios utilizados se incluyan en la misma carpeta en la que se encuentren los códigos. Además, la metodología se realiza mediante un código principal ‘*mainUAVs.m*’. Dicho *script* contiene la llamada a todas las funciones que irán realizando todas las diferentes partes del

trabajo. Entonces, para poder obtener los resultados de este trabajo se debe abrir dicho código y pulsar el botón de ejecutar de *MATLAB*.

Códigos de *MATLAB*

En este apartado se van a mostrar los códigos correspondientes a la función principal y los *scripts* de las funciones a las que llama dicho código para poder obtener los resultados. No obstante, en el análisis de los arrays de micrófonos y DOA, se va a mostrar un ejemplo de dicho código, ya que, entre el uso de un algoritmo u otro, varía el colector de la señal y los elementos del array de micrófonos, pero la base de dicho código es la misma.

Función principal

```
close all;clc;clear;

uavs=["QAV400", "Parrot AR", "Harop Harpy"];
frecuencia_fund=zeros(1,3);

for ii=1:3
    if uavs(ii) == "QAV400"
        [audio_doble,fs]=audioread('QAV400.wav');
    elseif uavs(ii) == "Parrot AR"
        [audio_doble,fs]=audioread('Parrot_AR.wav');
    else
        [audio_doble,fs]=audioread('Harop_Harpy.wav');
    end

    % Caracterizacion espectral de la señal acústica
    [frecuencia_fund(ii),fn]=FreqDron(audio_doble,fs,ii);
    disp(['Frecuencia fundamental del ' num2str(uavs(ii)) ': '...
        num2str(frecuencia_fund(ii)) ' Hz']);

    % Atenuación en ambiente real
    [A]=Atenuacion(fs,frecuencia_fund(ii),ii,fn);

    % Eleccion del array para cada DOA
    [BeamsCan_UCA, BeamsCan_URA, MVDR_UCA, MVDR_URA, MUSIC_UCA,MUSIC_URA,...
        GCCPHAT_UCA, GCCPHAT_URA]=EleccionArray_DOA(audio_doble);

    % Seleccion configuracion mediante barrido de angulos

    for e1=0:20:80
        [Error_BEAMSCAN_uca,Error_BEAMSCAN_ura,Error_MVDR_uca,Error_MVDR_ura,...
            Error_MUSIC_uca,Error_MUSIC_ura,Error_GCC_ura]=Eleccion_DOA(audio_doble,e1);
    end

    % Seleccion array y DOA en ambiente real

    for e1=[20,40,60]
        [error_BeamScan,error_MVDR,error_MUSIC,error_GCCPHAT]= ...
            EleccionArrayDOA(audio_doble,e1);
    end
end
```


Caracterización espectral de los UAVs

```

function [Freq_fund,fn]=FreqDron(audio,fs,ii)

uavs=["QAV400", "Parrot AR", "Harop Harpy"];
audio= (audio(:,1)+audio(:,2))/2; % Señal obtenida con 2 micrófonos.
M = length(audio); % Muestras que tiene el audio

% Tiempo
tiempo = (0:M-1)/fs; % Eje de tiempos

figure;
subplot(2,1,1);
stem(tiempo,audio);
axis([0 tiempo(end) -1 1]); xlabel('Tiempo(s)'); ylabel('Amplitud');
title(['Sonido del ' num2str(uavs(ii)) ' en función del tiempo']);

subplot(2,1,2);
plot(tiempo,audio);
axis([8 8.05 -1 1]); xlabel('Tiempo(s)'); ylabel('Amplitud');
title([num2str(uavs(ii)) ' ampliado a trama de 50ms']);
sgtitle('Análisis en el dominio del tiempo');

% Frecuencia (|fft|)

L = 2^nextpow2(M); % Número de trozos en los que se divide la señal.
f = fs*(0:(L/2))/L; % Eje de frecuencias

audio_fft = fft(audio,L,1);
audio_freq = abs(audio_fft/L); % FFT de la señal

figure;
plot(f,audio_freq(1:L/2+1)); % Se representa solo la mitad de la señal, ya que
    % la otra parte sería la compleja conjugada de la de la izquierda.
xlim([0 3000]);
xlabel('Frecuencia (Hz)'); ylabel('|UAV|');
sgtitle(['Análisis en el dominio de la frecuencia de ' num2str(uavs(ii))]);

% Espectrograma (STFT)

L_espectrograma=fs*0.02; % 20ms cada segmento
overlap=0.01*fs; % Solapamiento

figure;
spectrogram(audio,hann(L_espectrograma),overlap,2048,fs,'yaxis','power');
colormap('jet');
title(['Espectrograma de la señal del ' num2str(uavs(ii))]);

% Cepstrum (Frecuencia fundamental)

% Cepstrum
cepstrum=rceps(audio);

figure;
plot(tiempo,cepstrum);
title(['Análisis cepstral del dron ' num2str(uavs(ii))]);
  
```

```

    xlabel('Cuefrecencia');ylabel('Gamnitud');

% Cepstrum aplicando el liftado
fpb_inic=fs*(1/0.5e3); % Inicio del liftado
fpb_fin=fs*(1/50); % Final del liftado

cepstrum_lift=cepstrum(fpb_inic:fpb_fin);
eje_cuefrecencia=(0:length(cepstrum_lift)-1)/fs;

[peaks,Cuefrecs]=findpeaks(cepstrum_lift,eje_cuefrecencia);
valor_max=find(peaks==max(peaks));
Freq_fund=fs/(Cuefrecs(valor_max)*fs+fpb_inic);

figure;
plot(eje_cuefrecencia,cepstrum_lift);
hold on;
plot(Cuefrecs(valor_max),peaks(valor_max),'md','Linewidth',2);
xlabel('Cuefrecencia');ylabel('Gamnitud');
suptitle(['Cepstrum del ' num2str(uavs(ii))]);
title(['Frecuencia fundamental = ' num2str(Freq_fund) ' Hz']);

% Armonicos de la frecuencia fundamental

N=7; % 6 armonicos (más la frecuencia fundamental)
fudge_factor=25;
ventana(1,:)=linspace(100,300);
fn=zeros(1,7);
fn(1)=Freq_fund;

for n=2:N
    ventana(n,:)=n*fn(n-1)/(n-1)+[-fudge_factor fudge_factor];
    ventana2=linspace(ventana(n,1),ventana(n,2),length(ventana(n,1):ventana(n,2)));

    [pks_armonicos,locs_armonicos]=findpeaks(audio_freq(round(ventana(n,1)):
    round(ventana(n,2))),ventana2');
    amplitud=max(pks_armonicos);
    ind_armonicos=find(pks_armonicos==amplitud);
    fn(n)=locs_armonicos(ind_armonicos);
end

figure;
plot(f,audio_freq(1:length(audio_freq)/2+1));xlim([0 2000]);
hold on;
plot(fn,zeros(1,7),'dr','Linewidth',2);
xlabel('Frecuencia (Hz)');ylabel('Amplitud');
title(['Frecuencia fundamental y armónicos del ' num2str(uavs(ii))]);
end

```

Atenuación del sonido en ambientes reales

```

function [A]=Atenuacion(fs,frecuencia_fundamental,k,freqs)

uavs=["QAV400", "Parrot AR", "Harop Harpy"];

d=1:400; % Distancia desde la fuente al receptor. (Distancias hasta 400 metros).

```

```

figure;

for ii=1:length(freqs)

% A_div: ATENUACION POR DIVERGENCIA GEOMETRICA
d0=1; % 1m (distancia de referencia)
A_div(ii,:)=20*log10(d/d0)+11;

% A_atm: ATENUACION POR ABSORCION ATMOSFERICA
pr=101.325; % kPa reference ambient atmospheric pressure
pa=101.325; % kPa Ambient atmospheric pressure
T0=293.15; % K
T_grados=20;
T=T_grados+273.15; % Temperatura ambiente (kelvin)
Rv=461.5; % Constante de vapor de agua (Pa*m^3/°K*Kg)

Ra=286.9871; % Constante del aire seco (Pa*m^3/°K*Kg)
pvs=6.11*exp(17.27*T_grados/T)*0.1; % kPa Aproximacion para presion del vapor de
                                agua a saturacion

HR=40; %humedad relativa
w=(pvs*HR*Ra)/(Rv*(100*pa-HR*pvs)); % Humedad absoluta parcial
% v=(Ra+Rv*w*T)/pa; % Volumen del vapor de agua (aproximación)

% Sin embargo, si se hace mediante el diagrama psicrometrico directamente,
% sale el valor exacto del volumen.
v=838; % volumen del vapor de agua a 20°C y 40% de humedad (mediante diagrama
        psicrometrico(m^3/Kg))

Pv=Rv*T/v; % Presion parcial del vapor de agua
h=(Pv/pa); % Concentracion molar de vapor de agua

fr0=pa/pr*(24+(4.04*10^4*h*((0.02+h)/(0.391+h))));
frN=pa/pr*(T/T0)^(-1/2)*(9+(280*h*exp(-4.170*((T/T0)^(-1/3)-1))));

alfa(ii)=8.686.*freqs(ii).^2.*((1.84*10^(-11)*(pa/pr)^(-1)*(T/T0)^(1/2))...
    +(T/T0)^(-5/2)*(0.01275*(exp(-2239.1/T)).*(fr0+(freqs(ii).^2/fr0))^(-1)...
    +0.1068*(exp(-3352.0/T)).*(frN+(freqs(ii).^2/frN)).^(-1)));

alfa(ii)=alfa(ii)*10^3;

A_atm(ii,:)=alfa(ii)*d/1000;

% A_gr: ATENUACION POR EFECTO DEL SUELO

hs=d./2;% Altura de la fuente
hr=0;% Altura del receptor
dp=sqrt(d.^2-hs.^2);% Distancia fuente-receptor proyectada en el suelo (Tª
                    Pitagoras)

% Si suponemos suelo duro como el de un aeropuerto, Gs=Gr=Gm=0, por lo que As=Ar=-
1,5 sin tener que considerar más factores.
As(ii,:)= -1.5*ones(1,400);
Ar(ii,:)= -1.5*ones(1,400);

```

```

for jj=1:400
    if dp(jj)<30*(hs(jj)+hr)
        q(jj)=0;
        Am(ii,jj)=-3*q(jj);
    else
        q(jj)=1-30*(hs(jj)+hr)/dp(jj);
        Am(ii,jj)=-3*q(jj);% Porque Gm=0;
    end
end

A_gr(ii,:)=As(ii,:)+Ar(ii,:)+Am(ii,:);

for pp=1:length(d)
    if A_gr(pp,:)<0
        A_gr(pp,:)=0;
    end
end

% A_bar: APANTALLAMIENTO POR OBSTACULOS
obstaculo=0;
if obstaculo~=0
    lambda(ii)=3e8/freqs(ii);
    C2=20;
    C3=1;
    dss % Distancia desde la fuente al primer borde de difraccion
    dsr % Distancia desde borde difraccion hasta receptor
    z(ii,:)=((dss+dsr)^2+a^2)-d;
    for jj=1:length(d)
        if z(ii,jj)<0
            kmet(ii,jj)=1;
        end
    end
    Dz=10*log10(3+(C2/lambda(ii))*C3*z(ii,:)*kmet(ii,jj));
    A_bar=Dz-A_gr;
else
    % En el caso abordado no hay obstaculos, por lo que dicha atenuacion es 0.
    A_bar(ii,:)=zeros(1,400);
end

% A_misc: ATENUACION POR EFECTOS DIVERSOS
% A_misc=A_fol+A_site+A_hous;
% A_fol % Atenuacion al pasar por un bosque.
% A_site % Atenuacion por sitios industriales
% A_hous % Atenuacion por regiones de casas.

% Se supone un escenario aeroportuario
A_misc(ii,:)=zeros(1,400);

% A: ATENUACION TOTAL
A(ii,:)=A_div(ii,:)+A_atm(ii,:)+A_gr(ii,:)+A_bar(ii,:)+A_misc(ii,:);

plot(d,A(ii,:));hold on;
xlabel('Distancia(m)');ylabel('Atenuación (dB)');

```

```

    legend('Freq fundamental','1º armónico','2º armónico','3º armónico',...
          '4º armónico','5º armónico','6º armónico','7º armónico','8º armónico',...
          '9º armónico','Location','southeast');

end

    title(['Atenuacion del sonido del ' num2str(uavs(k))]);

end
  
```

Elección del array para cada DOA

Una vez se han calculado las características de la señal de entrada y del canal de propagación, se realiza el análisis comentado en el capítulo 6, para poder obtener la mejor combinación del array de micrófonos y algoritmo DOA. Para poder mostrar un ejemplo de *script* que permita calcular el apartado 6.1 se va a exponer el caso del algoritmo MUSIC con la geometría UCA. Para ello, primero se introduce la función y los parámetros de entrada, que será común para todos los algoritmos a analizar.

```

function [Beamscan_UCA,Beamscan_URA,MVDR_UCA,MVDR_URA,MUSIC_UCA,MUSIC_URA,...
        GCCPHAT_UCA, GCCPHAT_URA]=EleccionArray_DOA(audio_doble)

% Parametros de entrada
audio_48= (audio_doble(:,1)+audio_doble(:,2))/2; %Señal obtenida con 2 microfonos.
audio_8=resample(audio_48,8000,48000); % Se remuestrea de la señal a 8 kHz.

t=0:1/8000:5;
audio=audio_8(round(length(audio_8)/2)+1:...
              round(length(audio_8)/2+length(t))); %5 segundos del audio

microphone=phased.CustomMicrophoneElement('FrequencyVector', [0 20e3]);
c=343; %m/s velocidad sonido
d=0.3:0.05:0.6;
ang_UAV = [40; 30]; % Angulo donde se encuentra el UAV (azimut, elevacion)

err_elevacion=[];err_azimuth=[];
error_azimuth=[];error_elevacion=[];

for M=4:2:16
    for ii=1:length(d)

        % Array de micrófonos
        r=(d(ii)/2)/asin(pi/M);
        uca = phased.UCA(M,r,'Element',microphone);

        % Colector de la señal
        collector=phased.widebandCollector('Sensor',uca,'PropagationSpeed',c,...
            'SampleRate',8000,'ModulatedInput',false,'NumSubbands',1024);

        x=collector(audio,ang_UAV);

        % Division en subbandas (señal banda estrecha)
        for nmic=1:M
            S=stft(x(:,nmic),8000,'window',hann(512),'OverlapLength',511,...
                'FFTLength',512,'Centered',true);
        end
    end
end
  
```

```

    for sub=1:16
        subband(nmic, sub, :)=(S(256+4+sub, :)+S(256-4-sub, :))/2;
    end
    end
    L=length(subband(nmic, sub, :));

    azuca=[];
    eluca=[];

    % Estimación del ángulo
    for sub=1:16
        estimator=phased.MUSICEstimator2D('SensorArray',uca,...
            'OperatingFrequency',(sub+4)*8000/512,'PropagationSpeed',c,...
            'NumSignalsSource','Property','DOAOutputPort',true,'NumSignals',1,...
            'AzimuthScanAngles',0:180,'ElevationScanAngles',0:90);

        azm=[];
        elm=[];
        for m=1:fix(L/2000)
            input=squeeze(subband(:, sub, ((m-1)*2000+1:m*2000)));
            [~,doas]=estimator(input');%MUSIC
            azm=[azm doas(1)];
            elm=[elm doas(2)];
        end
        azuca=[azuca; azm];
        eluca=[eluca; elm];
    end

    % Error en azimuth y elevación del ángulo respecto el teórico
    media_filas_az=mean(azuca,2);
    media_az=mean(media_filas_az,'omitnan');
    err_az=mean(abs(ang_UAV(1)-media_az));
    err_azimuth=[err_azimuth err_az];

    media_filas_el=mean(eluca,2);
    media_el=mean(media_filas_el,'omitnan');
    err_el=mean(abs(ang_UAV(2)-media_el));
    err_elevacion=[err_elevacion err_el];

    fprintf('Terminado')
end

error_azimuth=[error_azimuth ; err_azimuth];
error_elevacion=[error_elevacion ; err_elevacion];
err_azimuth=[];
err_elevacion=[];
end

% Combinación de azimuth-elevación que menor error obtenga para cada
% número de microfonos.

error_total=[];
for ii=1:length(d)
    error_total(ii,:)=error_azimuth(ii,:)+error_elevacion(ii,:);
    minimo_error(ii)=min(error_total(ii,:));
end

```

```

MejorComb_mic=find(minimo_error==min(minimo_error));
if length(MejorComb_mic)>1
    MejorComb_mic=MejorComb_mic(1);
end

MejorComb_dist=find(error_total(MejorComb_mic,:)==min(minimo_error));
if length(MejorComb_dist)>1
    MejorComb_dist=MejorComb_dist(1);

end

MUSIC_UCA=[2*MejorComb_mic+2, d(MejorComb_dist)];

disp(['La mejor combinacion es ' num2str(2*MejorComb_mic+2) ...
      ' microfonos con una distancia de ' num2str(d(MejorComb_dist)) ' m']);
  
```

Selección configuración mediante barrido de ángulos

En la segunda parte del análisis (apartado 6.2) se selecciona una configuración de array para cada uno de los algoritmos DOA. Al igual que pasaba en el anterior caso, para evitar la redundancia, se muestra un ejemplo, que en este caso será el algoritmo Beamscan con el array URA.

```

function [Error_BEAMSCAN_uca,Error_BEAMSCAN_ura,Error_MVDR_uca,Error_MVDR_ura,...
        Error_MUSIC_uca,Error_MUSIC_ura,Error_GCC_ura]=Eleccion_DOA(audio_doble,e1)

% Parametros de entrada
audio_48= (QAV400_doble(:,1)+QAV400_doble(:,2))/2; %Señal obtenida con 2 microfonos.
audio_8=resample(audio_48,8000,48000); % Se remuestrea de la señal a 8 kHz.

t=0:1/8000:5;
audio=audio_8(round(length(audio_8)/2)+1:...
    round(length(audio_8)/2+length(t))); %5 segundos del audio

microphone=phased.CustomMicrophoneElement('FrequencyVector', [0 20e3]);
c=343; % m/s velocidad sonido

% BEAMSCAN URA

% Array de microfonos
M=6;
d=0.55;

err_elevacion_BeamScan_ura=[];err_azimuth_BeamScan_ura=[];

ura = phased.URA([M/2,M/2],[d,d],'Element',microphone);

% Colector de la señal
collector=phased.widebandCollector('Sensor',ura,'PropagationSpeed',c,...
    'SampleRate',8000,'ModulatedInput',false,'NumSubbands',1024);

for az=0:10:180

    x=collector(audio,[az,e1]);
  
```

```

% Divicion en subbandas (señal banda estrecha)
for nmic=1:(M/2)^2
    S=stft(x(:,nmic),8000,'window',hann(512),'OverlapLength',511,...
        'FFTLenght',512,'Centered',true);
    for sub=1:16
        subband(nmic,sub,:)=(S(256+4+sub,:)+S(256-4-sub,:))/2;
    end
end
L=length(subband(nmic,sub,:));

azura_BeamScan=[];
elura_BeamScan=[];

% Estimacion del angulo
for sub=1:16

    estimator=phased.BeamscanEstimator2D('SensorArray',ura,...
        'PropagationSpeed',c,'OperatingFrequency',(sub+4)*8000/512,...
        'AzimuthScanAngles',0:180,'ElevationScanAngles',0:90,'DOAOutputPort',true,...
        'NumSignals',1);

    azm=[];
    elm=[];
    for m=1:fix(L/2000)
        input=squeeze(subband(:,sub,((m-1)*2000+1:m*2000)));
        [~,doas]=estimator(input');
        azm=[azm doas(1)];
        elm=[elm doas(2)];
    end
    azura_BeamScan=[azura_BeamScan; azm];
    elura_BeamScan=[elura_BeamScan; elm];
end

% Error en azimut y elevacion respecto al teorico
media_filas_az=mean(azura_BeamScan,2);
media_az=mean(media_filas_az,'omitnan');
err_az=mean(abs(az-media_az));
err_azimuth_BeamScan_ura=[err_azimuth_BeamScan_ura err_az];

media_filas_el=mean(elura_BeamScan,2);
media_el=mean(media_filas_el,'omitnan');
err_el=mean(abs(el-media_el));
err_elevacion_BeamScan_ura=[err_elevacion_BeamScan_ura err_el];

fprintf('Terminado')

end

Error_BEAMSCAN_ura=[err_azimuth_BeamScan_ura,err_elevacion_BeamScan_ura];

clear subband;

```

La representación de las gráficas del apartado 6.2 se realizó tras ejecutar toda la función anterior, que corresponde a los códigos de los diferentes algoritmos, de tal forma que se representan siempre teniendo en cuenta que el error medio absoluto en la dirección de azimut y de elevación se calculan por separado.


```

figure;
    plot(0:10:180, err_azimuth_BeamScan_uca, '-dm');
        hold on;
    plot(0:10:180, err_azimuth_BeamScan_ura, '-db');
    plot(0:10:180, err_azimuth_MVDR_uca, '-dr');
    plot(0:10:180, err_azimuth_MVDR_ura, '-dg');
    plot(0:10:180, err_azimuth_MUSIC_uca, '-dc');
    plot(0:10:180, err_azimuth_MUSIC_ura, '-dy');
    plot(0:10:180, err_azimuth_GCC_ura, '-dk');

    legend('BEAMSCAN UCA', 'BEAMSCAN URA', 'MVDR UCA', 'MVDR URA', 'MUSIC UCA', ...
        'MUSIC URA', 'GCC PHAT URA', 'Location', 'northwest');
    xlabel('Azimuth (°)');
    ylabel('Error absoluto medio (°)');
    title(['Error absoluto medio de azimut a una elevación de ' num2str(e1) '°']);

figure;
    plot(0:10:180, err_elevacion_BeamScan_uca, '-dm');
        hold on;
    plot(0:10:180, err_elevacion_BeamScan_ura, '-db');
    plot(0:10:180, err_elevacion_MVDR_uca, '-dr');
    plot(0:10:180, err_elevacion_MVDR_ura, '-dg');
    plot(0:10:180, err_elevacion_MUSIC_uca, '-dc');
    plot(0:10:180, err_elevacion_MUSIC_ura, '-dy');
    plot(0:10:180, err_elevacion_GCC_ura, '-dk');

    legend('BEAMSCAN UCA', 'BEAMSCAN URA', 'MVDR UCA', 'MVDR URA', 'MUSIC UCA', ...
        'MUSIC URA', 'GCC PHAT URA', 'Location', 'northwest');
    xlabel('Azimuth (°)');
    ylabel('Error absoluto medio (°)');
    title(['Error absoluto medio de elevacion a una elevación de ' num2str(e1) '°']);
  
```

Selección array y DOA en un ambiente real

Por último, para poder realizar el análisis de la configuración final, se evalúan los algoritmos con la configuración seleccionada en los pasos anteriores para condiciones no ideales. Con este fin, se muestra la siguiente función en la que, tras declarar los parámetros de entrada y la atenuación, se muestra el ejemplo de la detección mediante el algoritmo MVDR con el array de geometría circular.

```

function [error_BeamScan,error_MVDR,error_MUSIC,error_GCCPHAT]=...
    EleccionArrayDOA(audio_doble)

% Parametros de entrada
    audio_48= (audio_doble(:,1)+audio_doble(:,2))/2;
    audio_8=resample(audio_48,8000,48000);

    t=0:1/8000:5;
    audio=audio_8(round(length(audio_8)/2)+1:round(length(audio_8)/2+length(t)));

    microphone=phased.CustomMicrophoneElement('FrequencyVector', [0 20e3]);
    c=343; %m/s velocidad sonido
  
```

```

% Atenuacion
fs=8000;
L = length(audio);
f = fs*(0:(L/2))/L; %Eje de frecuencias

audio_fft = fft(audio,L,1);
audio_freq = 10*log10(abs(audio_fft/L));
% d=1:400;
% [atenuacion]=Atenuacion_v2(d,L)
load atenuacion.mat

atenuacion_nodB=10.^(atenuacion/10);

for d=1:400
    % En el dominio de Fourier
    audio_fft_at_dB(:,d)=audio_freq-atenuacion(:,d);

    % En el dominio del tiempo
    audio_fft_at(:,d)=audio_fft./atenuacion_nodB(:,d);
    audio_audio_at(:,d)=ifft(audio_fft_at(:,d),L,1);

end

audio_at_real=real(audio_audio_at);

% MVDR UCA

% Array de microfonos
M=4;
d=0.3;

r=(d/2)/asin(pi/M);
uca = phased.UCA(M,r,'Element',microphone);

% Colector de la señal
collector=phased.widebandCollector('Sensor',uca,'PropagationSpeed',c,...
    'SampleRate',8000,'ModulatedInput',false,'NumSubbands',1024);

% Se añade un 10% de ruido al canal
noise = 0.1*(randn(size(collector))+1i*randn(size(collector)));

err_elevacion_MVDR_uca=[];err_azimuth_MVDR_uca=[];
error_MVDR=[];

for az=40:40:180
    for dist=linspace(1,400,9)

        x=collector(audio_at_real(:,floor(dist)),[az; e1]);

        % Division en subbandas (Señal banda estrecha)
        for nmic=1:M
            S=stft(x(:,nmic),8000,'window',hann(512),'OverlapLength',511,...
                'FFTLenght',512,'Centered',true);
            for sub=1:16 %xq 16*512 (Que es la longitud de la stft, da 8192Hz, que es el
                rango que nos interesa, lo demas no lo necesitamos).
                subband(nmic,sub,:)=(S(256+4+sub,:)+S(256-4-sub,:))/2;
            end
        end
    end
end

```

```

end
L=length(subband(nmic,sub,:));

azuca_MVDR=[];
elUCA_MVDR=[];

% Estimacion del angulo
for sub=1:16
    estimator=phased.MVDEstimator2D('SensorArray',uca,...
    'OperatingFrequency',(sub+4)*8000/512,'PropagationSpeed',c,...
    'DOAOutputPort',true,'NumSignals',1,...
    'AzimuthScanAngles',0:180,'ElevationScanAngles',0:90);

    azm=[];
    elm=[];
    for m=1:fix(L/2000)
        input=squeeze(subband(:,sub,((m-1)*2000+1:m*2000)));
        [~,doas]=estimator(input+noise);
        azm=[azm doas(1)];
        elm=[elm doas(2)];
    end
    azuca_MVDR=[azuca_MVDR; azm];
    elUCA_MVDR=[elUCA_MVDR; elm];
end

% Error en azimut y elevacion respecto al teorico
media_filas_az=mean(azuca_MVDR,2);
media_az=mean(media_filas_az,'omitnan');
err_az=mean(abs(az-media_az));
err_azimuth_MVDR_uca=[err_azimuth_MVDR_uca err_az];

media_filas_el=mean(elUCA_MVDR,2);
media_el=mean(media_filas_el,'omitnan');
err_el=mean(abs(el-media_el));
err_elevacion_MVDR_uca=[err_elevacion_MVDR_uca err_el];

end

error_MVDR_az=[ error_MVDR_az; err_azimuth_MVDR_uca];
err_azimuth_MVDR_uca=[];
error_MVDR_el=[error_MVDR_el ; err_elevacion_MVDR_uca];
err_elevacion_MVDR_uca=[];

end

error_MVDR=[error_MVDR_az, error_MVDR_el];

figure;
plot(linspace(1,400,9),error_MVDR_az(1,:),'-rd');
hold on;
plot(linspace(1,400,9),error_MVDR(2,:)_az,'-bd');
plot(linspace(1,400,9),error_MVDR(3,:)_az,'-gd');
plot(linspace(1,400,9),error_MVDR(4,:)_az,'-md');

legend('Az=40°','Az=80°','Az=120°','Az=160°');
xlabel('Distancia fuente-array (m)');
ylabel('Error medio absoluto (°)');

```

```
title(['Error medio absoluto en azimut (4 microfonos separados 0,3 m)']);

figure;
plot(linspace(1,400,9),error_MVDR_e1(1,:), '-rd');
hold on;
plot(linspace(1,400,9),error_MVDR(2, :)_e1, '-bd');
plot(linspace(1,400,9),error_MVDR(3, :)_e1, '-gd');
plot(linspace(1,400,9),error_MVDR(4, :)_e1, '-md');

legend('Az=40°', 'Az=80°', 'Az=120°', 'Az=160°');
xlabel('Distancia fuente-array (m)');
ylabel('Error medio absouto (°)');
title(['Error medio absoluto en elevacion (4 microfonos separados 0,3 m)']);

clear subband;
```

Finalmente, tras elegir el algoritmo MVDR, para seleccionar la configuración definitiva, se varían en este último código los parámetros M y d del array de micrófonos.

BIBLIOGRAFÍA

- [1] Jesús Bobadilla, Pedro Gómez y Jesús Bernal, “La transformada de Fourier. Una visión Pedagógica,” *Estudios de fonética experimental*, 10, pp. 41-74, 1999
- [2] Diccionario Oxford Languages and Google [online], (<https://languages.oup.com/google-dictionary-es/>)
- [3] Ivan Djurek, Antonio Petosic, Sanja Grubesa and Mia Suhanek, “Analysis of a Quadcopter’s Acoustic Signature in Different Flight Regimes,” *IEEE Access (Volume: 8)*, 2020, pp. 10662-10670.
- [4] “Controles de un dron multicóptero,” [online] (<https://www.hispadrones.com/principiantes/aprendizaje-consejos/controles-dron-multicoptero/>)
- [5] Alan V. Oppenheim, Alan S. Willsky with Ian T. Young, “Signals and Systems”, Prentice-Hall International Editions, 1983
- [6] M. Blanco Velasco, F. Cruz Roldán, R. Jiménez Martínez, J. Sáez Landete, *Tratamiento Digital de Señales*, Universidad de Alcalá, 2013
- [7] Alan V. Oppenheim, Ronald W. Schafer, “Discrete-Time Signal Processing,” Prentice Hall Signal Processing Series, 1989
- [8] Giraldo C, Daniel Quintero M, Olga L., “Análisis de señales de audio utilizando la transformada de Gabor,” Universidad EAFIT, Grupo de Modelado Matemático y Departamento de Ciencias Básicas, 2014
- [9] Hohyub Jeon, Yongchul Jung, Seongjoo Lee and Yunho Jung, “Area-Efficient Short-Time Fourier Transform Processor for Time-Frequency Analysis of Non-Stationary Signals,” *Article applied sciences*, MDPI, 2020
- [10] Elkin Flórez, Salvador Cardona, Lluïsa Jordi, “Selección de la ventana temporal en la transformada de Fourier en tiempos cortos utilizada en el análisis de señales de vibración para determinar planos en las ruedas de un tren,” *Rev. Fac. Ing. Univ. Antioquia N.º 50* pp. 145-158, 2009
- [11] Juan Luis Navarro Mesa, “Procesador Acústico: El Bloque de Extracción de Características,” Departamento de Señales y Comunicaciones, *Tratamiento de la Señal de Audio*, Universidad de las Palmas de Gran Canaria
- [12] Peter Klaer, Andi Huang, Pascale Sévigny, Sreeraman Rajan, Shashank Pant, Prakash Patnaik and Bhashyam Balaji, “An Investigation of Rotary Drone HERM Line Spectrum under Manoeuvring Conditions,” *Article sensors*, MDPI, 2020
- [13] Shubha Kadambe and G. Faye Boudreaux-Bartels, “Application of the Wavelet Transform for Pitch Detection of Speech Signals,” *IEEE Transactions on Information Theory (Volume: 38, Issue: 2)*, pp. 917-924, 1992
- [14] Subha Kadambe and G. Faye Boudreaux-Bartels, “A Comparison of a Wavelet Transform Event Detection Pitch Detector with Classical Pitch Detectors,” *IEEE Article of Conference Record Twenty-Fourth Asimolar Confreence on Signals, Systems and Computers*, 1990

- [15] A. M. Noll, and M. R. Schroeder, "Short-Time "Cepstrum" Pitch Detection," Acoustic Society of America, 1964
- [16] R. I. A. Harmanny, J. J. M. de Wit, G. Prémel Cabic, "Radar Micro-Doppler Feature Extraction Using the Spectrogram and the Cepstrogram," IEEE, 2014 11th European Radar Conference
- [17] Andi Huang, Pascale Sévigny, Bhashyam Balaji and Sreeraman Rajan, "Fundamental Frequency Estimation of HERM Lines of Drones," 2020 IEEE International Radar Conference (RADAR)
- [18] Xianyu Chang, Chaoqun Yang, Xiufang Shi, Pengfeng Li, Zhigou Shi and Jiming Chen, "Feature Extracted DOA Estimation Algorithm Using Acoustic Array for Drone Surveillance," 2018 IEEE 87th Vehicular Technology Conference (VTC Spring)
- [19] Normas ISO, ISOTools Excellence [online], (<https://www.isotools.org/normas/>)
- [20] "Atenuación del Sonido Durante la Propagación en Exteriores," ISO 9613 [9613-1 (1993), 1963-2 (1996)]
- [21] Dr Isaac Pilatowsky Figueroa, "Psicrometría, Métodos de Humidificación y Deshumidificación y sus aplicaciones en el Diseño Arquitectónico," Centro de Investigación en Energía, Universidad Nacional Autónoma de México, 2002
- [22] Online Interactive Psychrometric Chart [online], (<http://www.flycarpet.net/en/PsyOnline>)
- [23] Khyati R. Zalawadia, Twinkle V. Doshi, U.D. Dalal, "Adaptive Beam Former Design Using RLS Algorithm for Smart Antenna System," IEEE, 2011 International Conference on Computational Intelligence and Communication Networks, India
- [24] Plataforma Online Oficial MathWorks [online] (https://www.mathworks.com/help/matlab/index.html?s_tid=CRUX_lftnav)
- [25] Jacob Benesty, Jingdong Chen, Yiteng Huang, "Microphone Array Signal Processing," Springer Science & Business Media, 2008
- [26] Harry L. Van Trees, "Optimum Array Processing" Part IV of Detection, Estimation, and Modulation Theory, A John Wiley & Sons, INC., 2002.
- [27] Jacob Benesty, Jingsong Chen, Israel Cohen, "Design of Circular Differential Microphone Arrays," Springer Topics in Signal Processing, Volume 12, 2015
- [28] Honghao Tang, "DOA estimation based on MUSIC algorithm," bachelor's degree, Institutionen för Fysik och Elektroteknik, 2014
- [29] Farida Akbari, Shahriar Shirvani Moghaddam, Vahid Tabataba Vakili, "MUSIC and MVDR DOA Estimation Algorithms with Higher Resolution and Accuracy," 2010 5th International Symposium on Telecommunications, IEEE, 2011
- [30] Hamid Krim and Mats Viberg, "Two Decades of Array Signal Processing Research," IEEE Signal Processing Magazine (Volume: 13, Issue: 4), pp. 67-94, 1996
- [31] Michael Chryssomallis, "Smart antennas," IEEE Antennas and Propagation Magazine (Volume: 42, Issue: 3), pp. 129-139, 2000

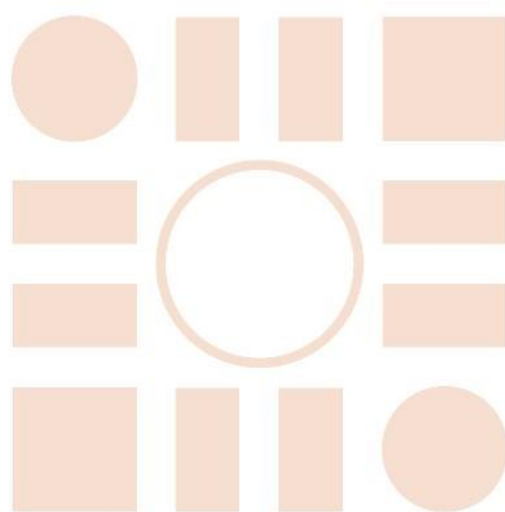
- [32] Nauman Anwar Baig and Mohammad Bilal Malik, "Comparison of Direction of Arrival (DOA) Estimation Techniques for Closely Spaced Targets," *International Journal of Future Computer and Communication*, Vol. 2, No. 6, December 2013
- [33] M. Devendra, Dr. K. Manjunathachari, "DOA Estimation of a System Using MUSIC Method," 2015 International Conference on Signal Processing and Communication Engineering Systems, IEEE, 205
- [34] D. Liu, Andy W. H. Khong, "A Wavelet-based GCC Prefiltering Algorithm for Speech DOA Estimation," *The Journal of the Acoustic Society of America*, November 2019
- [35] "Utilización de la autocorrelación para detectar señales enmascaradas por el ruido," [online]
(<http://www.ehu.eus/Procesadodesenales/tema8/corre1g.html#solucion>)
- [36] Andrea Bernardini, Federica Mangiatordi, Emiliano Palloti, Licia Capodiferro, "Drone detection by acoustic signature identification," *IS&T International Symposium on Electronic Imaging, Imaging and Multimedia Analysis in a Web and Mobile World*, 2017.
- [37] Xianyu Chang, Chaoqun Yang, Junfeng Wu, Xiufang Shi and Zhiguo Shi, "A Surveillance System for Drone Localization and Tracking Using Acoustic Arrays," 2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM), pp. 573-577.
- [38] Ismail Güvenç, Farshad Koochifar, Simran Singh, Mihail L. Sichitiu, and David Matolak, "Detection, Tracking, and Interdiction for Amateur Drones," *IEEE Communications Magazine* (Volume: 56, Issue: 4), pp. 75-81, 2018
- [39] Iain McCowan, "Microphone Arrays: A Tutorial," Doctoral dissertation, Queensland University of Technology, Brisbane, 2001
- [40] A.V Oppenheim, R.W. Schafer, "From frequency to quefrequency: a history of the cepstrum", *IEEE Signal Processing Magazine* (Volume: 21, Issue: 5), pp. 95-106, 2004.
- [41] Raúl Oscar González-Pacheco García, "Técnica para la Localización de Fuentes Basada en la Perspectiva de Reconstrucción de la Señal", Tesis doctoral, Universidad de Alcalá, 2013.
- [42] Jorge Combo Val, "Desarrollo de simulador para determinación de ángulo de llegada en RF," Trabajo de fin de grado, Universidad de Vigo, 2019-2020
- [43] "Standard Atmosphere", ISO 2533 (1975)
- [44] Weijie Zhu, "DOA Estimation for Broadband Signal That Use FFT Interpolation Method," 2013 IEEE 4th Internatoinal Conference on Software Engineering and Service Science
- [45] Jungfeng Guo, Ishtiaq Ahmad and KyungHi Chang, "Classification, positioning, and tracking of drones by HMM using acoustic circular microphones array beamforming," *EURASIP Jorunal on Wireless Communications and Networking*, Springer Open, 2020
- [46] Cherian P. Mathews and Michael D. Zoltowski, "Eigenstructure Techniques for 2-D Angle Estimation with Uniform Circular Arrays," *IEEE Transactions on Signal Processing*, Vol. 42, No.9, 1994
- [47] Gonçalo F. Cardoso y Ruben Pater, "A Study into 21st Century Drone Acosutics," [online]
(<http://www.droneacoustics.org/>)
- [48] SHANNON, Claude Elwood, "Communication in the presence of noide," *Proceedings of the IRE*, 1949, vol. 37, no 1, p. 10-21

Bibliografía

[49] Chengwei Zhou, Yujie Gu, Xing Fan, Zhigou Shi, Guoqiang Mao, Yimin D. Zhang, “Direction-of-Arrival Estimation for Coprime Array via Virtual Array Interpolation,” IEEE Transactions on Signal Processing, Volume: 66, Issue: 22, Nov.15, 2018

[50] P. Common and C. Jutten, “Handbook of Blind Source Separation,” Chapter 1: Introduction, pages 1–22. Elsevier Ltd, 1 edition, 2010

Universidad de Alcalá
Escuela Politécnica Superior



ESCUELA POLITECNICA
SUPERIOR



Universidad
de Alcalá