



PONTIFICIA **UNIVERSIDAD CATÓLICA** DEL PERÚ

Esta obra ha sido publicada bajo la licencia Creative Commons  
Reconocimiento-No comercial-Compartir bajo la misma licencia 2.5 Perú.

Para ver una copia de dicha licencia, visite  
<http://creativecommons.org/licenses/by-nc-sa/2.5/pe/>



**PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ**  
**FACULTAD DE CIENCIAS E INGENIERÍA**



**LOCALIZACIÓN DE UNA FUENTE DE SONIDO USANDO UN  
PEQUEÑO ARREGLO DE MICRÓFONOS**

Tesis para Optar el Título de:  
**INGENIERO ELECTRÓNICO**

Presentado por:  
**GUSTAVO ARRIETA VERA**

Lima – Perú  
2005

## RESUMEN

La localización de fuentes de sonido forma parte del procesamiento digital con arreglos de sensores empleando conceptos importantes como solapamiento espacial, teoría de señales aleatorias y estimación espectral de potencia, entre otros. Los métodos escogidos en este proyecto son la correlación cruzada en el dominio del tiempo, Beamforming (filtraje espacial) y MUSIC (Multiple Signal Classification) en el dominio de la frecuencia. Presentando un mayor grado de exactitud y complejidad según el mismo orden en que se los menciona.

El sistema de localización implementado en el presente trabajo de Tesis, consta de dos etapas. La primera etapa de conversión analógica / digital, compuesta por un arreglo de cuatro micrófonos omnidireccionales, una consola de audio de cuatro canales. La segunda etapa de procesamiento digital de señales, mediante el uso de una computadora personal y el entorno MATLAB, donde se implementaran los algoritmos de estimación.

Se llevaron a cabo simulaciones con señales sintetizadas, respetando el modelo de propagación de ondas acústicas, con el fin de observar el comportamiento del sistema de localización en condiciones ideales ó controlables, antes de trabajar con señales reales. Posteriormente, se realizaron experimentos con fuentes reales en dos ambientes. Los exteriores de la facultad de Ingeniería Electrónica como un ambiente casi sin reverberación y el estudio de grabación de la facultad de Ciencias y Artes de la Comunicación, donde existe cierto nivel de reverberación.

En el caso de localización de un tono de 500 hz, el error de estimación es pequeño en ausencia de la reverberación, pero al aumentar este el nivel de la reverberación el error de estimación se incrementa más de 3 veces. En el caso de localización de voces y ruido blanco (generado por un parlante), los resultados son alentadores para distancias cortas, incluso aumentando el nivel de reverberación, mostrando así un error medio de 10 grados aproximadamente.



*A mis padres.*

*Agradezco a mi familia por la comprensión, confianza y apoyo que tienen para conmigo en todo momento.*

*A Andrés Flores y Craig MacInnes por haberme dado la oportunidad y la asesoría para realizar este trabajo de Tesis.*

*A los profesores Jorge Moreno y Alberto Nakano por la asesoría y el apoyo brindado desde el inicio de este proyecto.*

*A mis amigos, a todos aquellos que me apoyaron, y de especial manera a Ursula García, Manuel García, Jordan Vitella, Akio Yoshimoto, Jorge Cumbreñas, Julio Paredes y a Raul Medina.*



*"Si te caes siete veces, levántate ocho."*

***Proverbio Chino***

*"Buscando el bien de nuestros semejantes encontramos el nuestro."*

***Platon***

*"Después de escalar una montaña, descubrimos que hay muchas otras por escalar."*

***Nelson Mandela***

## Índice General

<b>1</b>	<b>Introducción</b>	<b>1</b>
<b>2</b>	<b>Análisis y descripción del problema.</b>	<b>4</b>
2.1	El Sonido	6
2.1.1	El Sonido como onda acústica	6
2.1.2	Perturbaciones	8
2.1.2.1	Cálculo del tiempo de reverberación	10
2.1.2.2	Coefficiente de Absorción ( $\alpha$ )	12
2.2	Arreglos de sensores	12
2.2.1	Arreglo lineal uniforme	13
2.2.2	Ambigüedades	13
2.3	Muestreo Espacial	16
<b>3</b>	<b>Descripción del sistema.</b>	<b>18</b>
3.1	Sistema de localización	19
3.2	Metodología	21
3.3	Estimación en el dominio del Tiempo	24
3.3.1	Modelo de datos	25
3.3.2	Estimación de DOA usando Correlación Cruzada	25
3.3.3	Aproximación al problema real	28
3.4	Estimación en el dominio de la frecuencia	30
3.4.1	Modelo de datos	30
3.4.2	Estimación de DOA usando Beamforming	33
3.4.3	Estimación de DOA usando Music	34
3.4.4	Aproximación al problema real	36
<b>4</b>	<b>Simulaciones.</b>	<b>37</b>
4.1	Generación de señales sintéticas	39
4.2	Determinación del campo lejano	42
4.3	Error de estimación para diferentes direcciones de arribo	44
4.4	Influencia de la relación señal a ruido	45

4.5	Simulación de casos específicos	46
4.5.1	Error de Estimación Promedio	47
4.5.2	Variación del Error de Estimación	50
4.5.3	Tiempo de procesamiento	53
<b>5</b>	<b>Pruebas y resultados.</b>	<b>55</b>
5.1	Relación señal a ruido de los datos experimentales	58
5.2	Error de Estimación Promedio	60
5.3	Variación del Error de Estimación	63
	<b>Conclusiones</b>	<b>66</b>
	<b>Recomendaciones</b>	<b>69</b>
	<b>Bibliografía</b>	<b>72</b>
	<b>Anexos (Ver CD Adjunto)</b>	
A	Correlación Cruzada	
B	Descripción de Equipos	
B.1	Micrófonos Omnidireccionales AKG D-230	
B.2	Cables Belden 1813A	
B.3	Conectores Neutrik Nc 3MX	
B.4	Parlantes JBL	
B.5	Consola de audio MONA	
C	Códigos en MATLAB	

## Índice de Figuras

Figura 2.1	Características principales de una onda.	7
Figura 2.2	Oscilación de las moléculas al propagarse en una onda.	7
Figura 2.3	Frente de onda plano y rayos acústicos	8
Figura 2.4	Comportamiento del sonido en un cuarto.	9
Figura 2.5	Configuración para medición de Tiempo de reverberación.	11
Figura 2.6	Señales usadas para medición del tiempo de reverberación	11
Figura 2.7	Aproximación de curvas de decaimiento logarítmico para tiempos de reverberación T10, T20 y T30	11
Figura 2.8	Geometría de un arreglo lineal uniforme.	13
Figura 2.9	Caso hipotético de la posición de la fuente respecto a un arreglo lineal de 4 micrófonos	14
Figura 2.10	Cono de ambigüedad para un caso tridimensional.	14
Figura 2.11	Plano horizontal paralelo al piso	15
Figura 2.12	Restricción al plano horizontal.	15
Figura 2.13	Restricción al semiplano horizontal.	16
Figura 2.14	Muestreo espacial de una onda con sensores separados una distancia $d$ .	17
Figura 3.1	Diagrama General del Sistema de Localización	20
Figura 3.2	Intervalo de observación segmentado en cuadros de duración $m$ muestras.	21
Figura 3.3	Ventana Hamming de 128 puntos	22
Figura 3.4	Espectro de frecuencias de una señal con Ventana Rectangular y Ventana Hamming.	23
Figura 3.5	Arreglo de sensores frente a una onda plana.	24
Figura 3.6	Diferencia de caminos recorrida por una onda acústica en base a las mediciones de dos sensores separados una distancia $d$ .	25

Figura 3.7	Correlación cruzada de dos secuencias discretas reales.	26
Figura 3.8	Diferencia de caminos recorrida por una onda plana en base a la medición de dos sensores separados una distancia $d$ .	27
Figura 3.9	Resultados diferentes al aplicar correlación a dos secuencias discretas tomando referencias de inicio distintas	29
Figura 3.10	Estimación de un DOA de $60^\circ$ usando Beamforming	34
Figura 3.11	Estimación de un DOA de $60^\circ$ usando MUSIC.	35
Figura 3.12	Densidad Espectral de Potencia para un segmento de una señal sinusoidal real.	36
Figura 4.1	Formas de onda usadas para generar las señales sintéticas.	41
Figura 4.2	Patrón de ruido de fondo de ambos ambientes.	41
Figura 4.3	Variación del Error Promedio en función de la distancia entre la fuente y el arreglo de sensores.	43
Figura 4.4	Variación del Error Promedio en función del ángulo de arribo para una fuente situada a 2m del arreglo de sensores, sin interferencias.	44
Figura 4.5	Variación del Error Promedio en función de la relación señal a ruido.	45
Figura 4.6	Ubicación gráfica de los puntos de prueba.	46
Figura 4.7	Error Promedio al usar Correlación para la estimación del DOA.	47
Figura 4.8	Error Promedio al usar Beamforming para la estimación del DOA.	48
Figura 4.9	Error Promedio al usar MUSIC para la estimación del DOA.	49
Figura 4.10	Varianza del Error de estimación del DOA al usar Correlación.	50
Figura 4.11	Varianza del Error de estimación del DOA al usar Beamforming.	51

Figura 4.12	Varianza del Error de estimación del DOA al usar MUSIC.	52
Figura 4.13	Tiempo de procesamiento de un experimento en el dominio del tiempo.	53
Figura 4.14	Tiempo de procesamiento de un experimento en el dominio de la frecuencia.	54
Figura 5.1	Exteriores de la sección de Ing. Electrónica	56
Figura 5.2	Estudio de Grabación de la facultad de Ciencias y Artes de la Comunicación.	57
Figura 5.3	Patrón omnidireccional de los micrófonos	57
Figura 5.4	Relación señal a ruido (SNR)	59
Figura 5.5	Error Promedio al usar Correlación para la estimación del DOA.	60
Figura 5.6	Error Promedio al usar Beamforming para la estimación del DOA.	61
Figura 5.7	Error Promedio al usar MUSIC para la estimación del DOA.	62
Figura 5.8	Varianza del Error de estimación del DOA al usar Correlación.	63
Figura 5.9	Varianza del Error de estimación del DOA al usar Beamforming.	64
Figura 5.10	Varianza del Error de estimación del DOA al usar MUSIC.	65

## Índice de Tablas

Tabla 2.1	Coeficientes de absorción	12
Tabla 3.1	Equivalencia entre el retardo y el ángulo de incidencia para una onda plana muestreada a 8khz y a 44.1 kHz.	28
Tabla 4.1	Relación de ambientes de los experimentos.	39
Tabla 4.2	Coeficientes de absorción de los ambientes.	40
Tabla 5.1	Tiempo de reverberación T30 (en segundos).	58





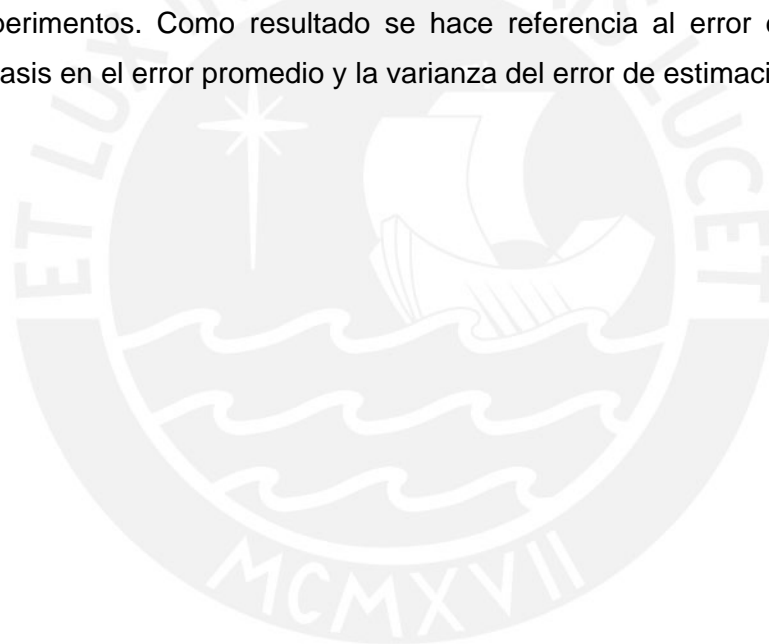
**Capítulo 1**  
**Introducción.**

La localización de fuentes de sonido [1- 4] se ha vuelto un tema de interés creciente en los últimos 20 años para científicos, investigadores y hombres de negocios, particularmente, en el área de procesamiento de señales con arreglos de sensores [5, 6, 7]. Campo que se centra en la habilidad de fusionar información temporal y espacial recogida por varios sensores para resolver una misma tarea. Este marco, usa como ventaja el priorizar información del sistema de adquisición (i.e. la geometría del arreglo [8, 9], algoritmos de estimación, etc.).

Las técnicas desarrolladas para dicha tarea han mostrado ser útiles en resolver problemas del mundo real, como por ejemplo, en sistemas de videoconferencia [10], mejora de voz [11], reconocimiento de voz [12], identificación de locutor [13] y equipos de ayuda de escucha [14]. Muchas tecnologías actuales podrían beneficiarse de una localización de un tipo específico de sonido. Nótese además que estas técnicas son usadas con distintos tipos de señales, y por ende en diversos campos tales como, la comunicación inalámbrica [15], la medicina [16] o inclusive la aeronáutica [17].

Haciendo uso de principios de procesamiento digital de señales [18, 19, 20], álgebra lineal [21], y conceptos de procesos estocásticos [22], este trabajo pretende analizar métodos de estimación en el dominio del tiempo y de la frecuencia de la dirección de arribo de una onda plana con un pequeño arreglo de micrófonos a fin de establecer una base para su aplicación en sistemas más complejos. En el dominio del tiempo se analizará un método que utiliza la correlación cruzada [23, 24] para obtener la localización de fuentes. Método ampliamente usado debido a su simplicidad. En el dominio de la frecuencia se analizarán dos métodos. El primer método en el dominio de la frecuencia es conocido como filtraje espacial o Beamforming [25], el cual data de la segunda guerra mundial y es una aplicación de un análisis espectral de Fourier de datos espacio-temporales muestreados. El segundo método llamado MUSIC (Clasificación Múltiple de Señales) [26], forma parte de un grupo de técnicas basadas en sub-espacios [27, 28], las cuales marcaron el inicio de una nueva era en la literatura sobre procesamiento de señales con arreglos de sensores. El acercamiento basado en sub-espacios recae en ciertas propiedades geométricas del modelo de datos asumido, resultando en una mayor capacidad de resolución [27, 28]. En comparación con el método en el dominio del tiempo, estos algoritmos tienen como ventaja una mayor precisión y un mayor grado de robustez ante interferencias como ecos y reverberación [29, 30, 31, 32].

A continuación, en el Capítulo 2 *Análisis y descripción del problema*, se hace una revisión de los conceptos necesarios para enfrentar el problema propuesto de localización de fuentes de sonido con un pequeño arreglo de sensores. Entre éstos se ve la definición del sonido como una onda acústica y las perturbaciones a la que es sometido normalmente. También se verá una definición de un arreglo de sensores y las restricciones necesarias para evitar ambigüedades al trabajar con el mismo. En el Capítulo 3 *Descripción del sistema*, se hace una explicación general del sistema utilizado, así como una descripción de las técnicas mencionadas en el dominio del tiempo y de la frecuencia. En el Capítulo 4 *Simulaciones*, se hace un estudio previo a los experimentos con señales sintetizadas con el objetivo de probar los algoritmos considerando una situación ideal y otra situación que considere un cierto nivel de reverberación. En el Capítulo 5 *Experimentos y Resultados*, se muestran los experimentos realizados con fuentes reales en diferentes condiciones y los resultados de estos experimentos. Como resultado se hace referencia al error de estimación. Poniendo énfasis en el error promedio y la varianza del error de estimación.





**Capítulo 2**  
**Análisis y Descripción del Problema.**

En el presente capítulo se hace una revisión de los conceptos necesarios para enfrentar el problema propuesto de localización de fuentes de sonido con un pequeño arreglo de sensores. Se debe analizar el tipo de señal con el cual se trabajará; así como su comportamiento, características y las perturbaciones del medio de propagación. Es de gran importancia conocer estos conceptos con el fin de establecer un modelo que los contemple. Con ello se analizarán herramientas que se cuentan para resolver el problema, particularmente, el arreglo de sensores y sus limitaciones en el procesamiento digital de señales.

Se ve la definición del sonido como una onda acústica, además de la forma en la cual se propaga y su comportamiento ante las reflexiones de éste ante el ambiente que lo rodea. Por lo tanto se explica la forma en que se puede modelar un ambiente considerando el tipo de material presente.

Se muestra una definición de un arreglo de sensores, así como la de un pequeño arreglo lineal uniforme de sensores y las restricciones necesarias para evitar ambigüedades [8, 9]. Estas restricciones son la causa de que la fuente de sonido y el arreglo de sensores se mantengan en la mitad de un plano dividido por una línea que pasa por la posición de los micrófonos.

Además, es necesario conocer los requisitos que se deben tomar en cuenta para que las señales que luego serán sensadas por el arreglo contengan información útil, es decir, que tengan un significado físico: para lo cual se hace una revisión del teorema de Nyquist en el caso temporal y en el espacial.

## 2.1 El Sonido

El sonido es una perturbación que se propaga en un medio elástico, ya sea gaseoso, líquido o sólido; la cual puede ser percibida por el oído humano. Cuando nos referimos al sonido audible por el oído humano, estamos hablando de la sensación detectada por nuestro oído, es decir, rápidas variaciones de presión en el aire por encima y por debajo de un valor estático. Éste valor estático nos lo da la presión atmosférica (alrededor de 100.000 pascales). La cual tiene unas variaciones pequeñas y de forma muy lenta, tal y como se puede comprobar en un barómetro.

Cuando las rápidas variaciones de presión se centran entre 20 y 20.000 veces por segundo (igual a una frecuencia de 20 Hz a 20 kHz) el sonido es potencialmente audible aunque las variaciones de presión puedan ser a veces tan pequeñas como la millonésima parte de un pascal. Los sonidos muy intensos son causados por grandes variaciones de presión, por ejemplo, una variación de 1 pascal se oiría como un sonido de muy alto nivel siempre y cuando la mayoría de la energía de dicho sonido estuviera contenida en las frecuencias medias (1kHz - 4 kHz) que es donde el oído humano es más sensitivo.

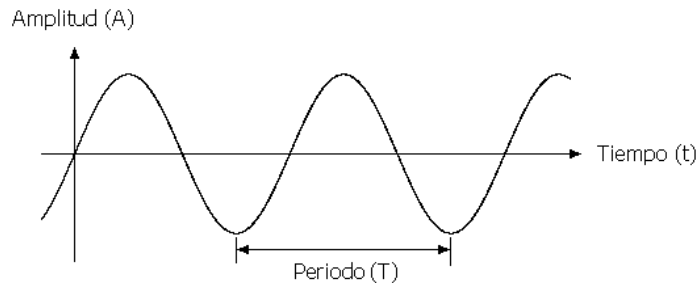
### 2.1.1 El Sonido Como Onda Acústica

Es debido al comportamiento descrito líneas arriba que se puede clasificar al sonido como una onda mecánica. Mecánica porque necesita un medio físico por donde propagarse. Con respecto a la *forma de propagación*, reconocemos dos, conforme la oscilación se realice transversal o paralela a la dirección de propagación, las primeras reciben el nombre de transversales y las segundas el nombre de longitudinales [1],

A continuación se explicarán los parámetros que caracterizan a una onda:

- Periodo (T): Es el tiempo que debe transcurrir para una oscilación completa.
- Frecuencia (f): Es el número de oscilaciones que se dan en una unidad de tiempo. Es la inversa del periodo, es decir, que  $f = 1/T$ .
- Longitud de Onda ( $\lambda$ ): Es la distancia que recorre la onda en una oscilación completa.

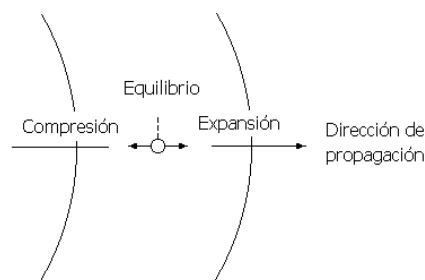
- Velocidad de propagación (C): Es la velocidad con la que se transmite la señal. Se cumple que  $c = \lambda/T = \lambda \cdot f$ .



**Figura 2.1** Características principales de una onda.

Cabe resaltar que los dos primeros parámetros (periodo y frecuencia) dependen exclusivamente de la fuente generadora de la onda, de la longitud de onda y velocidad de propagación; y del medio en el cual se esté propagando la onda. Ambos pueden cambiar si es que las condiciones del medio son variables (por ejemplo, temperatura, profundidad, densidad, etc.).

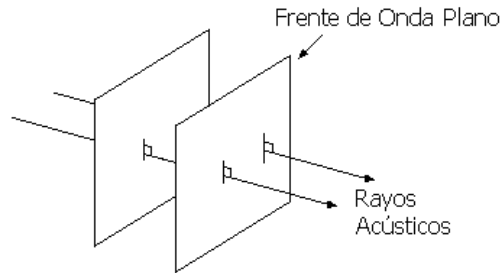
Como se mencionó, el sonido se produce debido a las compresiones y expansiones de un medio elástico, o sea, de las vibraciones que se generan en él.



**Figura 2.2** Oscilación de las moléculas al propagarse en una onda.

Al unir con una superficie imaginaria la zona de compresión, se definen superficies llamadas frentes de onda, donde los vectores perpendiculares a dichos frentes de onda serán conocidos como rayos acústicos. Se puede hacer una clasificación en función de la forma del frente de onda que presentan:

- Onda Plana: Se da cuando los frentes de onda son planos paralelos.
- Onda Esférica: Se da cuando los frentes de onda son esferas concéntricas.
- Onda Cilíndrica: Se da cuando los frentes de onda son cilíndricos esféricos.



**Figura 2.3** Frente de onda plano y rayos acústicos

En esta investigación se pretende trabajar con ondas planas debido al modelo matemático empleado. Pues al no tener ondas planas se tendrá un error de estimación grande y consecuentemente no se podrá estimar la dirección de arribo. Por lo tanto, es importante determinar la distancia crítica, que separa el campo cercano del campo lejano y será de este campo lejano donde se generen las ondas planas y la correcta estimación de la dirección de arribo (Direction of Arrival – DOA).

### 2.1.2 Perturbaciones

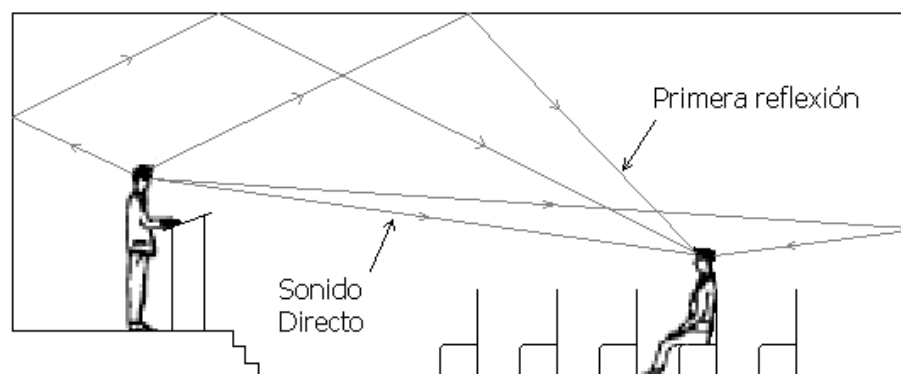
Cuando se genera un sonido en el interior de un local, las superficies que componen el mismo, ocasionan una serie de diferentes efectos dependiendo de las características de dichas superficies. Esto ocurre porque las ondas sonoras inciden en las diferentes superficies y éstas las reflejan de diferente forma según su coeficiente de reflexión acústica.

En situaciones naturales hablamos de sonido directo para referirnos al sonido que se transmite directamente desde la fuente sonora hasta nosotros (o hasta el mecanismo de captación que tengamos). Por otra parte, el sonido reflejado es el que percibimos después de que haya rebotado en las superficies que delimitan el recinto acústico, o en los objetos que se encuentren en su trayectoria. Evidentemente, la trayectoria del sonido reflejado siempre será más larga que la del sonido directo, de manera que -temporalmente- escuchamos primero el sonido seco, y unos instantes

más tarde escucharemos las primeras reflexiones; a medida que transcurre el tiempo las reflexiones que nos llegan son cada vez de menor intensidad, hasta que desaparecen. Nuestra sensación, no obstante, no es la de escuchar sonidos separados, ya que el cerebro los integra en un único precepto, siempre que las reflexiones lleguen con una separación menor de unos 50 milisegundos.

La reverberación es la suma total de las reflexiones del sonido que llegan al lugar del oyente en diferentes momentos del tiempo. Auditivamente se caracteriza por una prolongación, a modo de "cola sonora", que se añade al sonido original.

La reverberación de un local se mide según su tiempo de reverberación ( $T_{60}$ ) en segundos, definido como el tiempo que toma al sonido en caer 60 dB después de que la fuente deja de emitir el sonido. Lo que equivale a una disminución de 1000 veces la presión sonora. El tiempo de reverberación varía según la frecuencia de análisis que se utilice debido a que los diferentes materiales que componen las superficies del local no se comportan de igual modo en todo el espectro sonoro, y por tanto los coeficientes de absorción de cada superficie de un mismo material varían según la frecuencia. Conociendo el tiempo de reverberación de un local podemos saber cómo se comportará el mismo en diferentes aplicaciones. Cuando la energía reverberante alcanza valores muy altos con respecto al sonido directo, puede ocurrir un enmascaramiento de éste y se puede perder la capacidad de entender la información contenida en el mensaje que se percibe. Se conoce que el tiempo óptimo de reverberación para una habitación está entre 1 o 2 segundos como máximo.



**Figura 2.4** Comportamiento del sonido en un cuarto.

### 2.1.2.1 Cálculo del Tiempo de Reverberación

En una habitación cualquiera hay varios factores que se deben conocer para calcular el tiempo de reverberación de dicha habitación [2]. Éstos son:

- Dimensiones del cuarto
- Tipo de materiales del cuarto
- Área de las superficies de cada material del cuarto.

En la actualidad la ecuación más utilizada para calcular el tiempo de reverberación en ambientes poco reverberantes con distribución no uniforme es la ecuación de Fitzroy:

$$T_{60} = \frac{0.161 \times V}{S^2} [T_x + T_y + T_z] \quad (2.01)$$

$$T_x = \frac{S_x}{-\ln(1 - \alpha_x)}, \quad T_y = \frac{S_y}{-\ln(1 - \alpha_y)}, \quad T_z = \frac{S_z}{-\ln(1 - \alpha_z)} \quad (2.02)$$

Donde  $S_x$ : Superficie de las paredes laterales;

$S_y$ : Superficie de las paredes posterior y anterior.

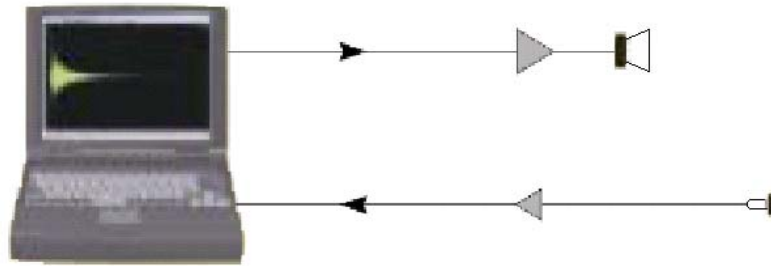
$S_z$ : Superficie del piso y del techo

$\bar{\alpha}_x$ : Absorción promedio de las paredes laterales

$\bar{\alpha}_y$ : Absorción promedio de las paredes posterior y anterior.

$\bar{\alpha}_z$ : Absorción promedio del piso y del techo.

Otra forma de medir el tiempo de reverberación es de una manera experimental. Se utilizó el software DIRAC Room Acoustics de Brüel & Kjær, un amplificador SAE modelo A14, un arreglo de parlantes en forma de dodecaedro (omnidireccional) y un micrófono omnidireccional, usando el método de integración inversa [3].

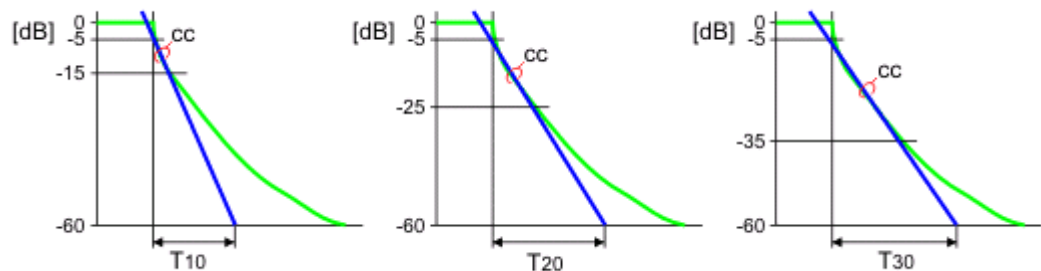


**Figura 2.5** Configuración para medición de Tiempo de reverberación.



**Figura 2.6** Señales usadas para medición del tiempo de reverberación

El procedimiento para la medición del tiempo de reverberación consiste en generar una señal MLS (Ruido pseudoaleatorio) ó una señal Sweep (Barrido senoidal periódico) para luego calcular los tiempos en que el nivel de presión sonora decae 10, 20 o 30 dBs.



**Figura 2.7** Aproximación de curvas de decaimiento logarítmico para tiempos de reverberación T10, T20 y T30

El T60 puede estimarse a partir de las pendientes mostradas en la figura 2.7 debido a que los niveles de presión sonora no son suficientes para medir correctamente el tiempo en que el sonido decae 60 dBs.

### 2.1.2.2 Coeficiente de Absorción ( $\alpha$ )

El coeficiente de absorción de un material es la relación entre la energía absorbida por el material y la energía reflejada por el mismo. Dada esta formulación su valor siempre está comprendido entre 0 y 1. El máximo coeficiente de absorción está determinado por un valor de 1 donde toda la energía que incide en el material es absorbida por el mismo, y el mínimo es 0 donde toda la energía es reflejada.

Material	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz
Piso de Concreto	0.02	0.02	0.02	0.04	0.05	0.05
Piso de Madera	0.15	0.2	0.1	0.1	0.1	0.1
Piso de alfombra	0.1	0.15	0.25	0.3	0.3	0.3
Pared de ladrillos	0.05	0.04	0.02	0.04	0.05	0.05
Cortinas	0.05	0.12	0.15	0.27	0.37	0.5
Vidrio	0.04	0.04	0.03	0.03	0.02	0.02

**Tabla 2.1** Coeficientes de absorción

El coeficiente de absorción varía con la frecuencia y por lo tanto los fabricantes de materiales acústicos dan los coeficientes de absorción por lo menos en resolución de una octava.

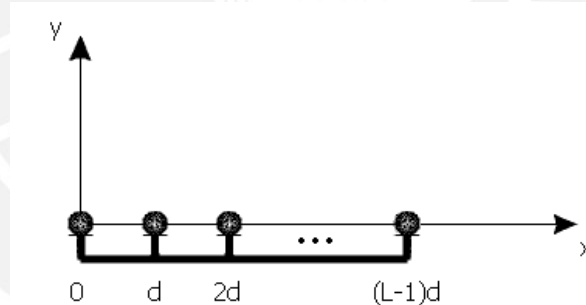
## 2.2 Arreglos de Sensores

Un arreglo de sensores es simplemente un grupo de sensores distribuidos ya sea de manera geométrica (definida) o aleatoria, tal que se obtengan datos de cada uno de ellos y se procesen con un mismo fin. Una de las ventajas de usar arreglos de sensores es evitar la complejidad en la parte física o hardware de un sistema y que el procesamiento digital (ó analógico) de las señales sea la parte principal del sistema; permitiendo una mayor flexibilidad del sistema debido a que es más sencillo modificar el código o programa (software) que modificar la disposición física de los equipos (hardware).

Los pequeños arreglos de micrófonos son de dos a seis micrófonos próximos unos a otros. El arreglo más simple (el arreglo lineal de 2 micrófonos) proporciona la base por sobre la cual los otros arreglos se derivan. En el arreglo, cada micrófono tiene una relación dependiente de tiempo de retardo con los otros micrófonos del arreglo establecida por la situación de la fuente de sonido. De estos retardos observados, se puede localizar la fuente de sonido como se explicará en el siguiente capítulo.

### 2.2.1 Arreglo Lineal Uniforme

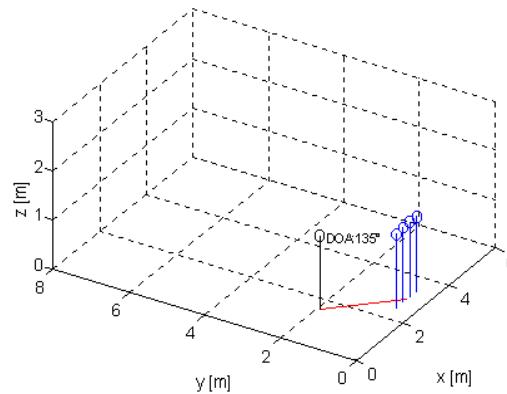
Un arreglo lineal uniforme consiste en un grupo de sensores en línea de tal forma que el espacio entre sensores vecinos es constante. También puede verse como arreglos de 2 micrófonos uno al lado del otro como se muestra en la figura 2.8. Este tipo de arreglo de sensores es ampliamente usado como se puede apreciar en diversos trabajos [5, 6, 7].



**Figura 2.8** Geometría de un arreglo lineal uniforme.

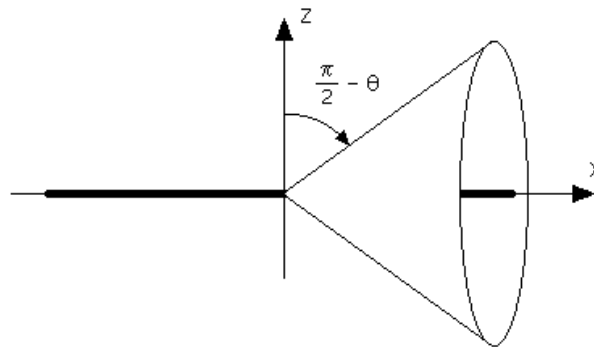
### 2.2.2 Ambigüedades

Debido al modelo de datos y al arreglo lineal uniforme utilizado es necesario restringir la fuente de sonido a ciertas posiciones en el espacio [8, 9].



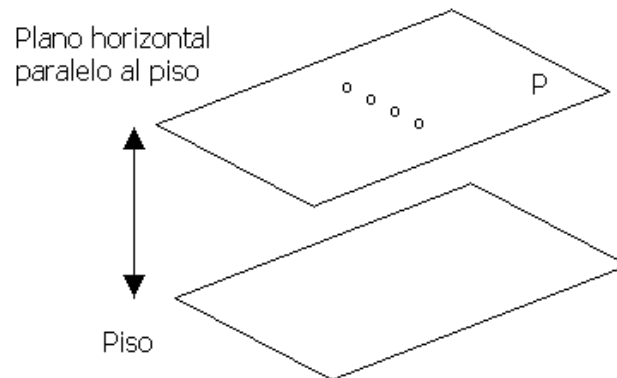
**Figura 2.9** Caso hipotético de la posición de la fuente respecto a un arreglo lineal de 4 micrófonos

Supongamos que la dirección de arribo (DOA, direction of arrival) de una señal sea  $135^\circ$  como se muestra en la figura 2.9, si bien la ubicación de la fuente a la misma altura que el arreglo de micrófonos implica el DOA  $135^\circ$ , existen otros lugares que también implican el mismo DOA. Estos lugares donde existe ambigüedad respecto al DOA  $135^\circ$ , se pueden representar como la superficie de un cono como se muestra en la figura 2.10.

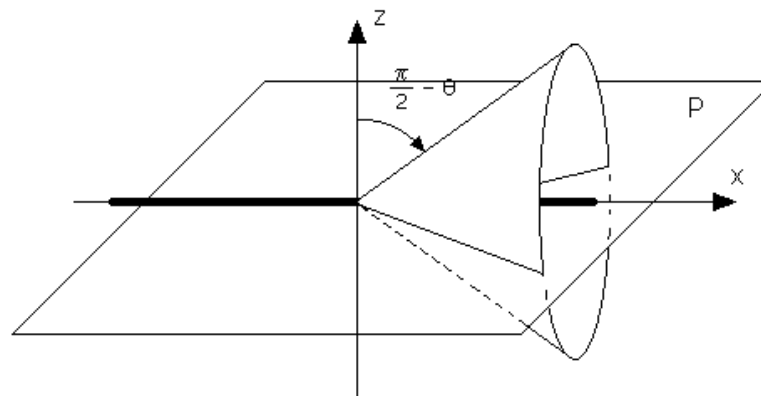


**Figura 2.10** Cono de ambigüedad para un caso tridimensional.

La primera restricción es ubicar la fuente en un plano horizontal paralelo al suelo a la altura del arreglo de micrófonos, como si se cortara el cono con un plano que pasa por la ubicación de los sensores.

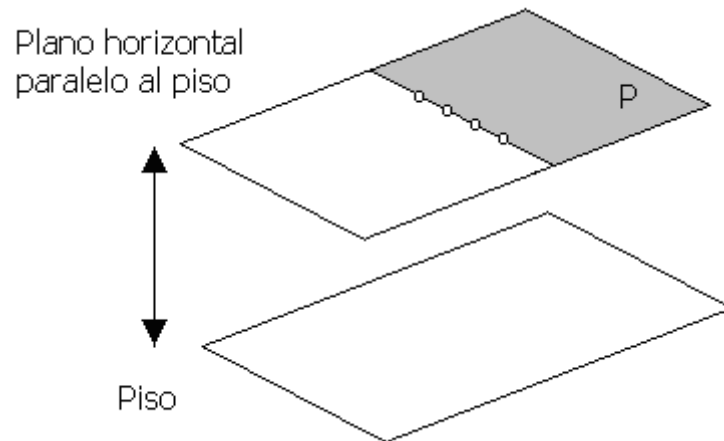


**Figura 2.11** Plano horizontal paralelo al piso



**Figura 2.12** Restricción al plano horizontal.

La segunda restricción es ubicar la fuente en uno de los semiplanos que resultan de cortar el plano con la línea imaginaria donde se ubican los micrófonos.



**Figura 2.13** Restricción al semiplano horizontal.

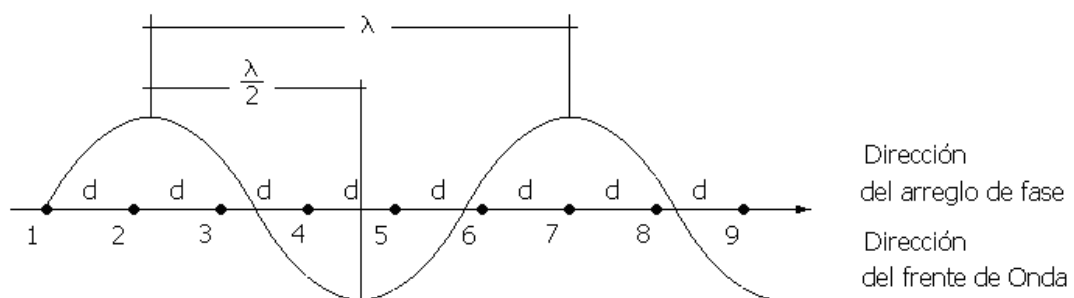
### 2.3 Muestreo Espacial

Para todos es muy conocido el Teorema de Nyquist, el cual establece que para poder digitalizar una señal de naturaleza continua y limitada en banda, cuyo contenido frecuencial máximo viene dado por la frecuencia  $f_n$ , se debe muestrear a dicha señal con una frecuencia  $f_s$  mayor o igual que el doble de  $f_n$ . En general, es recomendable muestrear con una frecuencia mucho mayor que la mínima ( $f_s \gg f_n$ ).

Se sabe que cuando se muestrea una señal en el tiempo se tienen que cumplir con ciertas restricciones (Teorema de Nyquist) para que la información que se obtiene de la señal muestreada sea equivalente a la información contenida en su versión continua. Es lógico suponer que cuando se muestrea en el espacio se deba cumplir también con algunas restricciones para que la información que registramos tenga un significado físico real. Estas restricciones existen y son conocidas como el Teorema de Nyquist en el caso espacial, el cual se enuncia a continuación:

“Si se tiene una onda que viaja en cierto medio, a una velocidad  $c$ , y cuyo contenido frecuencial está limitado en banda, y viene dado por  $f_n$ , se cumple que  $c = f_n \cdot \lambda_n$ ; de esto se afirma que la mínima longitud de onda contenida en dicha onda viene dada por  $\lambda_n$ . Luego para muestrear espacialmente a dicha señal, sin que la misma pierda su significado físico, la separación máxima entre los elementos que se empleen para dicho fin deberá cumplir:  $d_{max} = \lambda_n / 2$ ”

Para una mejor explicación de este teorema se hace referencia a la figura 2.14, aquí se presenta una señal cuyo frente de onda incide sobre el arreglo de sensores en la misma dirección.



**Figura 2.14** Muestreo espacial de una onda con sensores separados una distancia  $d$ .

En esta figura se representa el punto que estaría muestreando cada elemento del arreglo. Tal como está dispuesto el arreglo de 9 elementos, se realiza el muestreo espacial de manera correcta, ya que se está cumpliendo con la condición expresada en el teorema 2.1. Se puede entender esta condición como que en todo momento se conoce cuál es el signo de la primera derivada (pendiente) de la onda que se está muestreando, ya que si por ejemplo el arreglo de fase estaría constituido por los elementos 3, 6 y 9 no se podría afirmar si el signo de la pendiente de la señal en el elemento 6 es positiva. Si luego se observa los elementos 6 y 9, se podría afirmar que la pendiente del elemento 9 es negativa, afirmación que es falsa; se estaría frente a un caso de solapamiento espacial.



**Capítulo 3**  
**Descripción del Sistema.**

En este capítulo se presenta una descripción general de todo el sistema de localización. Primero se mostrará de manera global el sistema en sí, es decir, los equipos, algunas características y las funciones que realizan cada uno de ellos. Se podrán distinguir esencialmente dos etapas, una es sobre la adquisición de señales con el arreglo de micrófonos y la consola de audio MONA. Debido a esta consola los datos son digitalizados y transmitidos al computador personal para que luego se realice la estimación del ángulo de incidencia, lo que conformaría la segunda etapa.

En segundo lugar, se muestra la metodología con la que se trata a las señales para obtener una buena estimación. La metodología consistirá básicamente en la segmentación de las señales en pequeños cuadros con suficiente información para luego (mediante un promediado de ciertos valores involucrados en la estimación) lograr una mayor precisión en los resultados.

Y por último, una explicación de cada uno de los algoritmos de estimación en los dominios del tiempo y de la frecuencia. Se podrá ver cómo es que todas las técnicas de estimación utilizan como base los retardos de las señales recibidas entre los micrófonos del arreglo. En el dominio del tiempo, como se ha mencionado antes, la correlación cruzada es la esencia de la estima al obtener el retardo o desfase entre dos secuencias discretas. En cambio, en el dominio de la frecuencia es necesario aplicar más herramientas, tales como: la Transformada de Fourier, la densidad espectral de potencia y la matriz de covarianza, entre otros. Las técnicas en este último dominio se pueden describir como la búsqueda del ángulo de incidencia al “apuntar” con el arreglo de sensores en las direcciones donde pueda estar la fuente; formando así un espectro espacial y cuyo máximo indica el ángulo buscado.

### 3.1 Sistema de Localización

En esta sección se procede a describir al sistema en su integridad y a los métodos de estimación del ángulo de arribo de las señales. En la figura 3.1 se puede ver la estructura básica del sistema. Este sistema está compuesto por una etapa de conversión analógica / digital y una etapa de procesamiento digital de señales donde se realizará la estimación.



**Figura 3.1** Diagrama General del Sistema de Localización

A continuación se hará una descripción a grandes rasgos de las partes del sistema, empezando por los equipos. Si se desea más información de las especificaciones técnicas de los equipos se sugiere revisar el anexo B en el CD adjunto.

Como se mencionó anteriormente, el sistema puede dividirse en dos partes, las cuales describiré a continuación:

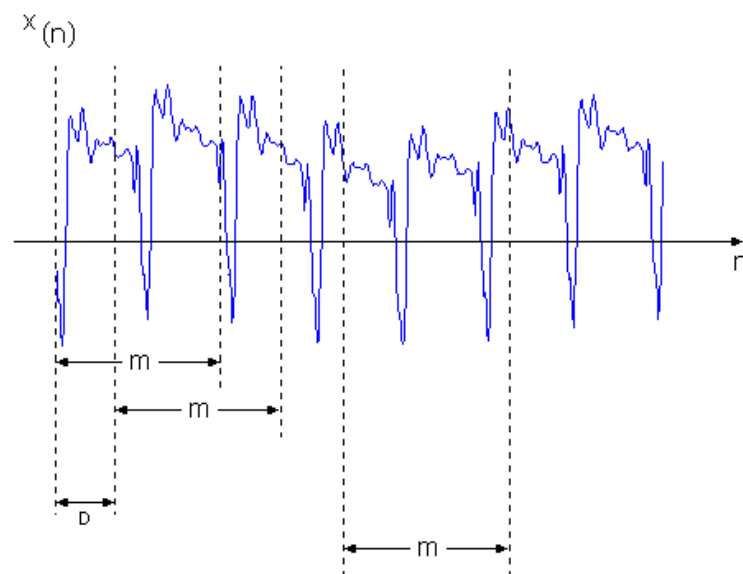
La primera etapa consiste en la conversión analógica/digital por la consola de Audio MONA, la cual tiene entre sus características principales 4 canales analógicos de entrada con frecuencia de muestreo de hasta 96 khz por canal en simultáneo; es decir, que se puede muestrear los 4 canales a la vez a más de 8khz. Ésta es una frecuencia de muestreo muy utilizada en aplicaciones con voz.

Al usar micrófonos omni-direccionales, el arreglo de micrófonos es sencillo ya que si fueran direccionales habría que hacer un barrido por cada dirección de arribo posible con ayuda de un sistema electromecánico muy complejo. Una vez que se tiene acceso a los datos digitalizados, estos son almacenados para un análisis posterior para aplicar distintos algoritmos de estimación a los mismos datos.

La segunda etapa consiste en analizar estas señales en busca de la dirección de arribo y de la posición de la fuente de sonido, en los dominios del tiempo y de la frecuencia. Esto se realizó en un computador personal como se puede apreciar en la gráfica y como característica principal es contar con un microprocesador Pentium IV de 1GHz y 256 MB de RAM. Se hará hincapié en la utilización de herramientas matemáticas que trabajen con vectores, en este caso, el software seleccionado es MATLAB versión 6.0 o Release 12.

### 3.2 Metodología

Uno de los aspectos prácticos del procesamiento de señales es el promediado de las señales recibidas en el tiempo para reducir la variabilidad de las estimaciones [7, 18]. Como se muestra en la figura 3.2, se puede segmentar la señal en  $N$  cuadros, cada uno de duración  $m$  muestras y un solapamiento  $D$ , medido en porcentaje de la cantidad de muestras por cuadro. Vale decir que un solapamiento del 100% resultaría en el análisis del primer cuadro solamente.

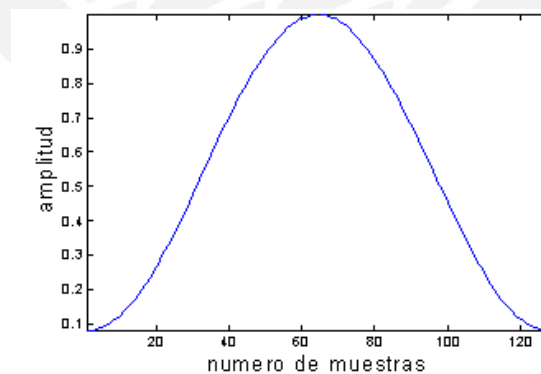


**Figura 3.2** Intervalo de observación segmentado en cuadros de duración  $m$  muestras.

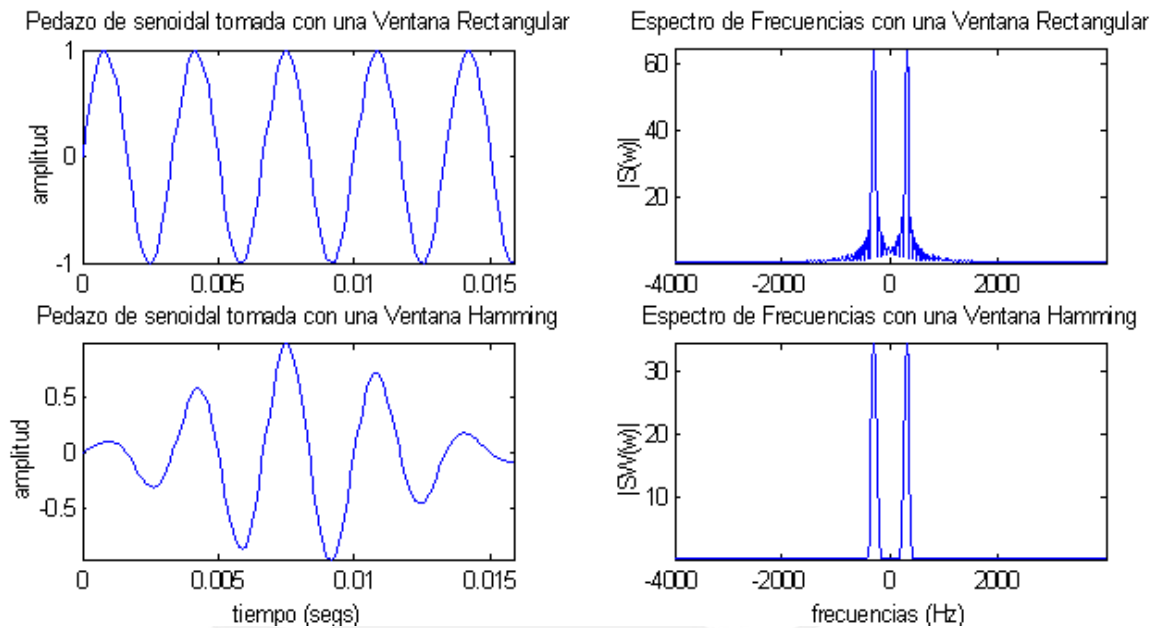
Se debe hacer notar que para el cálculo de tal espectro de frecuencias, se utiliza un enventanamiento Hamming [19] con el fin de reducir los lóbulos laterales que se producen normalmente al obtener la transformada de Fourier de un cuadro. Como se sabe, la transformada de Fourier se define para señales periódicas y la señal que puede contener un cuadro no lo será necesariamente.

La cantidad de puntos de esta ventana Hamming debe ser 4 o más veces la inversa de la frecuencia de interés multiplicado por la frecuencia de muestreo. Estas frecuencias de interés estarán en el rango de 300 hz a 566 hz. El límite inferior de 300hz se escogió para evitar la frecuencia de la señal de alimentación de los equipos, 60 hz y su primer armónico a 120 hz. En cambio, el límite superior fue seleccionado porque es la máxima frecuencia de una señal que se puede muestrear respetando el teorema de Nyquist para un arreglo donde sus elementos están separados una distancia  $d = 0.30m$ .

Si consideramos la frecuencia de interés mínima, entonces se podrá calcular la longitud de la ventana Hamming máxima. Para una frecuencia de muestreo de 8 khz, la cantidad de puntos de la ventana Hamming está alrededor de 128 puntos, valores que se utilizarán para realizar las simulaciones y los experimentos.



**Figura 3.3** Ventana Hamming de 128 puntos



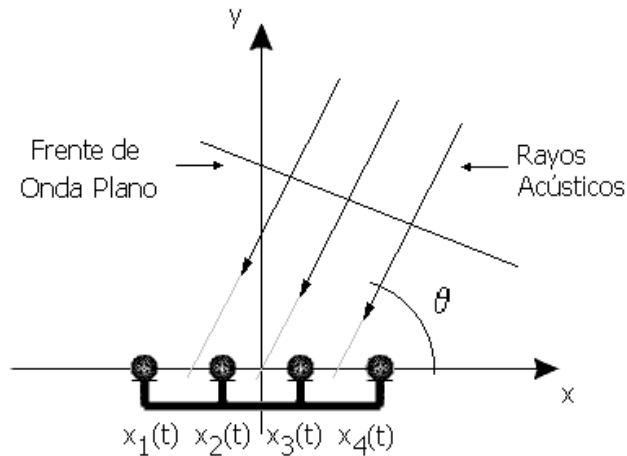
**Figura 3.4** Espectro de frecuencias de una señal con Ventana Rectangular y Ventana Hamming.

La longitud  $m$  de cada cuadro debe ser tal que contenga información suficiente para una buena estimación. En el caso de la estimación en el dominio del tiempo, el valor limitante de dicha longitud es el desfase o retardo máximo entre dos secuencias discretas. Este concepto se mostrará más adelante. En el caso de las técnicas en el dominio de la frecuencia, el valor limitante es la cantidad de puntos de la ventana Hamming. Al comparar estos dos valores, se puede demostrar que la cantidad de puntos de la ventana Hamming es el mayor y por lo tanto, será el que definirá la longitud  $m$  de cada cuadro en este trabajo.

De esta manera se define un experimento como la toma de  $N$  cuadros de duración  $m$  (en muestras) y con un solapamiento  $D$ ; experimento que al repetirse  $n$  veces proporcionará  $n$  estimaciones del ángulo de arribo o  $n$  errores de estimación; donde los cuadros de cada experimento serán procesados para obtener una dirección de arribo.

A causa de una mayor facilidad de entendimiento, se asume que  $\theta$  es un valor real referido como el ángulo de incidencia o dirección de arribo (DOA, *Direction of*

Arrival) de un frente de onda plano dirigido hacia el arreglo de sensores como se muestra en la siguiente figura:



**Figura 3.5** Arreglo de sensores frente a una onda plana.

Nótese la onda plana que recibe el arreglo de sensores. Ésta será la forma correcta de tratar a las señales captadas si es que la fuente está lo suficientemente lejos del centro del arreglo de sensores.

Este error de estimación se define como el valor absoluto de la diferencia entre el ángulo de arribo real y el estimado:

$$\text{Error de estimación} = \left| DOA_{real} - DOA_{estimado} \right| \quad (3.01)$$

Una vez obtenidos los  $n$  errores de estimación; se podrá obtener el error promedio, el error máximo y la varianza de este error como resultados finales [20, 21, 22]. En las siguientes secciones se muestran los algoritmos de estimación que serán aplicados a cada cuadro del experimento.

### 3.3 Estimación en el Dominio del Tiempo

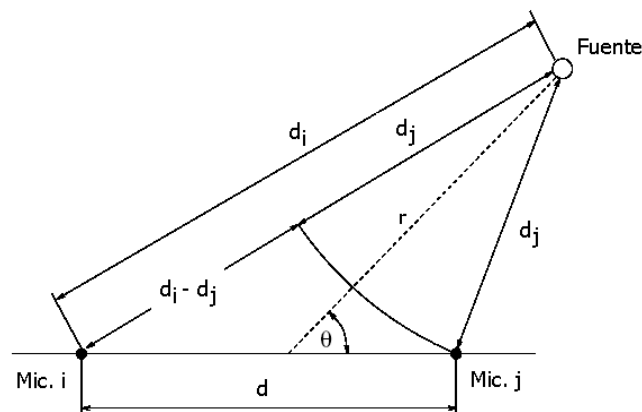
La estimación en el dominio del tiempo se basa en la correlación cruzada de dos señales para encontrar el desfase o retardo relativo entre las señales captadas por los micrófonos [23, 24]; y así determinar la dirección de donde proviene la señal. Si

extendemos este concepto y utilizamos dos o más pares de sensores, se puede incluso encontrar la posición al intersectar dos direcciones de arribo distintas.

### 3.3.1 Modelo de Datos

Hay que percatarse que el cálculo del ángulo de incidencia equivale al cálculo del retardo (en número de muestras) entre las señales captadas por los elementos del arreglo de sensores; en otras palabras, el cálculo o estima de uno de ellos implica a los otros dos. Por esta razón, la referencia a uno de ellos en particular es indistinta.

Estos retardos varían desde cero cuando la fuente genera señales con dirección de arribo de  $90^\circ$  hasta un máximo cuando la fuente genera señales con dirección de arribo de  $0^\circ$ .



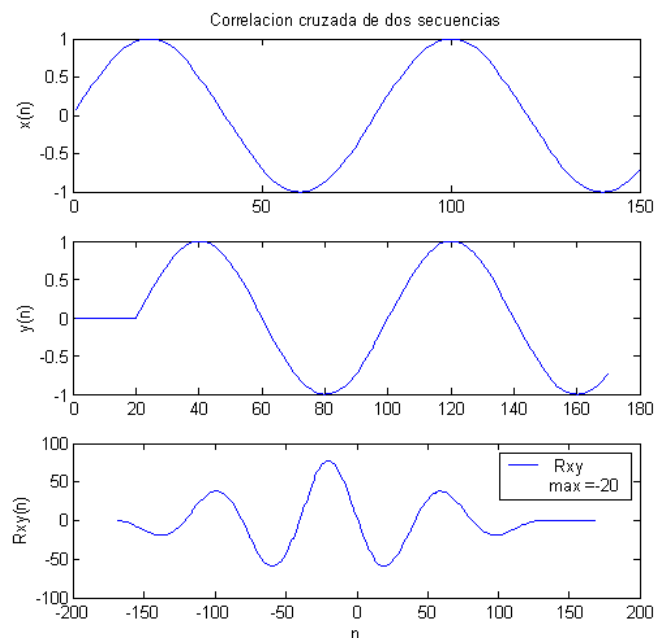
**Figura 3.6** Diferencia de caminos recorrida por una onda acústica en base a las mediciones de dos sensores separados una distancia  $d$ .

### 3.3.2 Estimación de DOA usando Correlación Cruzada

El cálculo de este retardo se halla usando la correlación cruzada como se mencionó antes. Si consideramos dos secuencias discretas reales  $x(n)$  e  $y(n)$ , ambas de energía finita, la correlación cruzada de las secuencias  $x(n)$  e  $y(n)$  es la secuencia  $r_{xy}(l)$ ; que se define como:

$$r_{xy}(l) = \sum_{n=-\infty}^{\infty} x(n)y(n-l), \quad l = 0, \pm 1, \pm 2, \dots \quad (3.02)$$

Si estas dos secuencias  $x(n)$  e  $y(n)$  son iguales, pero desfasadas en un número de muestras  $l$ , entonces se puede predecir que ocurrirá un máximo en la secuencia de correlación cruzada en el índice  $l$ . El método que se plantea es detectar el índice  $l$  del punto máximo en la secuencia de correlación. Esto se da entre dos secuencias de datos obtenidas por los micrófonos y así se obtiene el desfase. En la figura 3.7 se puede apreciar dos señales similares pero una retardada de otra. Además, la secuencia de correlación tiene un máximo en la muestra -20, lo cual quiere decir que la señal  $y(n)$  tiene un retardo de 20 muestras respecto a  $x(n)$ .



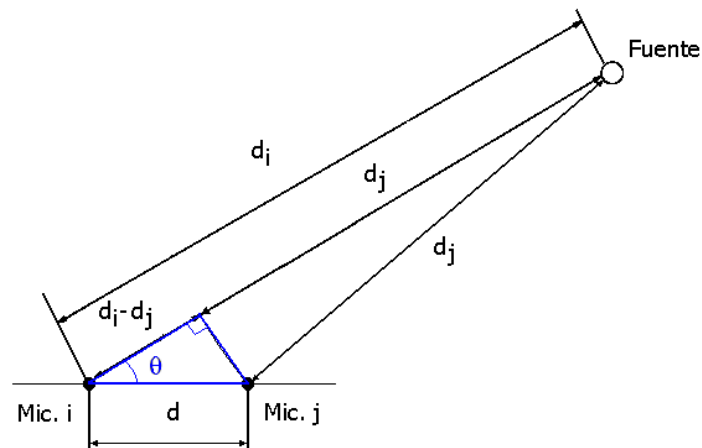
**Figura 3.7** Correlación cruzada de dos secuencias discretas reales.

Como se puede apreciar, la correlación cruzada necesita dos juegos de datos para obtener un retardo; por consiguiente, es necesario contar por lo menos con dos micrófonos para obtener cualquier información significativa. En un arreglo de dos micrófonos, un micrófono es más cercano a la fuente que el otro o en caso de que ambos sean equidistantes de la fuente, no habrá retardo.

Una vez obtenido el retardo del número de muestras entre señales captadas por 2 micrófonos, se puede obtener el ángulo de incidencia  $\theta$  haciendo uso de la siguiente relación:

$$d_i - d_j = c \cdot \left( \frac{\text{retardo}}{f_s} \right), (m), \quad (3.03)$$

donde el  $\text{retardo} / f_s$  es la diferencia de tiempos que al multiplicarse por la velocidad del sonido  $c$  (340 m/s), nos da la diferencia de caminos  $d_i - d_j$  en metros.



**Figura 3.8** Diferencia de caminos recorrida por una onda plana en base a la medición de dos sensores separados una distancia  $d$ .

Al trabajar con ondas planas, el cálculo de incidencia, se puede determinar con la ecuación 3.04 si se conoce las longitudes de los lados adyacentes al ángulo de incidencia  $\theta$  del triángulo rectángulo mostrado en la figura 3.8.

$$\theta = \cos^{-1} \left( \frac{d_i - d_j}{d} \right) \quad (3.04)$$

Nótese que a medida que se tienen más pares de micrófonos “independientes” se obtendrá una mejor estimación del ángulo de arribo; pero a la vez la distancia a la cual se sitúe la fuente de sonido para generar una onda plana será mayor. Micrófonos independientes se refiere a que si por ejemplo se tienen 4 micrófonos, sólo habrá que considerar 3 pares de micrófonos ya que otras parejas (aparte de las 3 iniciales) sólo nos darían resultados que se pueden calcular con los resultados que se obtienen de las 3 parejas iniciales.

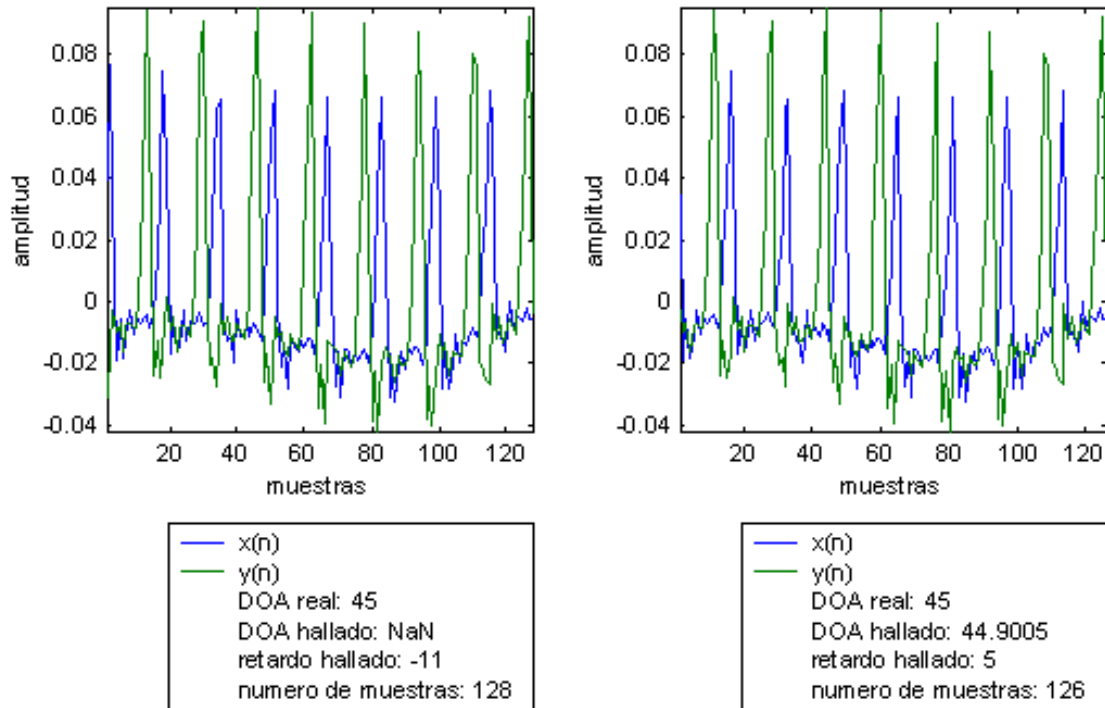
### 3.3.3 Aproximación al Problema Real

En un caso práctico puede que la correlación de dos señales indique un retardo mayor al que se puede esperar, y en esta oportunidad el retardo debe ser tal que la diferencia de caminos sea menor a la distancia entre sensores, es decir,  $d_i - d_j < d$ . En la tabla 3.1 se puede apreciar la relación entre los retardos y la dirección de arribo, la cual dependerá de la frecuencia de muestreo utilizada.

retardo (muestras)	DOA @ fs = 8 khz (grados)	DOA @ fs = 44.1 khz (grados)
0	90.00	90.00
1	81.86	88.53
2	73.54	87.05
3	64.85	85.58
4	55.48	84.10
5	44.90	82.62
6	31.79	81.13
7	7.40	79.64
8	Incoherente	78.14
...	...	...
38	Incoherente	12.43

**Tabla 3.1** Equivalencia entre el retardo y el ángulo de incidencia para una onda plana muestreada a 8khz y a 44.1 kHz.

Estas incoherencias se dan debido a una mala interpretación de los datos al aplicar la correlación. Como se muestra en la figura 3.9, al considerar una distancia entre sensores  $d = 30cm$ , una frecuencia de muestreo  $f_s = 8khz$  y dos secuencias discretas  $x(n)$  e  $y(n)$ , se pueden obtener resultados de la correlación distintos al considerar referencias de inicio distintas, es decir, que al considerar los 128 puntos de cada una de las secuencias se obtiene un retardo de -11, lo cual es cierto si se observa la gráfica de estas secuencias de 128 puntos. En esta gráfica  $x(n)$  parece llegar primero que  $y(n)$  si asumimos que son señales captadas por 2 sensores.



**Figura 3.9** Resultados diferentes al aplicar correlación a dos secuencias discretas tomando referencias de inicio distintas

Sin embargo, si consideramos que el retardo máximo es 7 muestras, un retardo de 11 muestras es absurdo. Por esta razón se debe hacer una corrección. Si se considera los 126 puntos finales de ambas secuencias se obtiene un retardo de 5 muestras, lo cual indica que  $y(n)$  llega antes que  $x(n)$ . Este retardo está dentro del rango de posibles retardos por lo que será la respuesta final de esta estimación.

Como se puede apreciar, dependerá del retardo máximo entre las señales captadas por dos sensores para decidir cual es el resultado de la estimación en el dominio del tiempo que se deba considerar.

Es debido a esta consideración que las parejas escogidas para la estimación en el dominio del tiempo serán los pares de micrófonos consecutivos, es decir, si alineamos 4 micrófonos como en la figura 3.5 y consideramos las secuencias obtenidas por los micrófonos  $x_1(t)$ ,  $x_2(t)$ ,  $x_3(t)$  y  $x_4(t)$ ; la correlación se hará entre los pares de secuencias  $x_1(t) - x_2(t)$ ,  $x_2(t) - x_3(t)$  y  $x_3(t) - x_4(t)$ .

### 3.4 Estimación en el Dominio de la Frecuencia

En esta sección se explicarán dos algoritmos comúnmente usados en la localización de fuentes de diferentes tipos, como lo son las fuentes de sonido.

El primero es el llamado *Beamforming* o *filtraje espacial* [25], el cual debe su nombre a una analogía con una linterna, donde la búsqueda se hace con un rayo (*beam*) de luz, usando los vectores de dirección mostrados en la siguiente sección.

El segundo método conocido como MUSIC (Clasificación múltiple de señales) [26] utiliza conceptos de álgebra lineal, en especial, de subespacios. Pero antes de explicar las técnicas; es necesario establecer el modelo de datos, revisar la definición de matriz de covarianza y su descomposición en subespacios.

#### 3.4.1 Modelo de Datos

Muchos acercamientos modernos al procesamiento de señales están basados en modelos, en el sentido que recaen en ciertas suposiciones hechas sobre los datos observados. El interés principal está en señales de banda angosta (lo cual no es realmente una restricción, sino que se hace uso de la propiedad de superposición, donde una señal puede ser considerada como una combinación lineal de señales de banda angosta).

Siendo  $L = 4$  el número de sensores y  $x_1(t)$ ,  $x_2(t)$ ,  $x_3(t)$  y  $x_4(t)$  las señales recibidas por los micrófonos como se muestra en la figura 3.5; se define el vector salida del arreglo de sensores como:

$$\mathbf{x}(t) = [x_1(t) \ x_2(t) \ x_3(t) \ x_4(t)]^T, \quad (3.05)$$

donde  $^T$  denota la transpuesta. Como se muestra en la figura 3.5, es fácil ver que para una señal de banda angosta:  $s_1(t) \approx \alpha e^{j2\pi f_c t}$ , las señales recibidas por los micrófonos son iguales a las que emite la fuente pero con diferentes retardos

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} = \begin{bmatrix} s_1(t) \\ s_1(t + \tau_{12}) \\ s_1(t + \tau_{13}) \\ s_1(t + \tau_{14}) \end{bmatrix} = \begin{bmatrix} 1 \\ e^{j2\pi f_c \tau_{12}} \\ \vdots \\ e^{j2\pi f_c \tau_{14}} \end{bmatrix} s_1(t) = \mathbf{a}(\theta) s_1(t), \quad (3.06)$$

donde  $\tau_{ij}$  es el retardo entre la señal captada por el mic.  $i$  y el mic  $j$ .

Para una mejor comprensión de la estimación del ángulo de arribo con Music, se presentará la extensión de este modelo para varias señales. Por lo tanto, si  $M$  señales llegan al arreglo de dimensión  $L$  de distintas direcciones de arribo  $\theta_1, \dots, \theta_M$ , el vector de salida toma la forma:

$$\mathbf{x}(t) = \mathbf{a}(\theta_1) s_1(t) + \mathbf{a}(\theta_2) s_2(t) + \dots + \mathbf{a}(\theta_M) s_M(t) \quad (3.07)$$

$$\mathbf{x}(t) = \sum_{k=1}^M \mathbf{a}(\theta_k) s_k(t) \quad (3.08)$$

La ecuación de salida puede ser puesta de una forma mas compacta definiendo una matriz de dirección y un vector de señales:

$$\mathbf{A}(\theta) = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_M)], \quad (L \times M) \quad (3.09)$$

$$s(t) = [s_1(t), \dots, s_M(t)]^T, \quad (M \times N) \quad (3.10)$$

En presencia de un ruido aditivo  $\mathbf{n}(t)$  se puede ahora obtener el modelo comúnmente usado:

$$\mathbf{x}(t) = \mathbf{A}(\theta) s(t) + \mathbf{n}(t) \quad (3.11)$$

Los métodos a ser presentados requieren que  $M < L$ , condición que se cumple pues se tiene 4 sensores y una fuente. Es interesante notar que en caso de ausencia de ruido, la salida del arreglo es confinada a un subespacio complejo de

dimensión  $L$ ; el cual es creado por combinaciones lineales de los vectores de dirección  $\mathbf{a}(\theta)$ . Este es el subespacio de la señal y su observación forma la base para los métodos basados en subespacios.

El cálculo de la matriz de covarianza es un paso necesario para la estimación en el dominio de la frecuencia y esta se define como:

$$\mathbf{R} = E\{x(t)x^H(t)\}, \quad (3.12)$$

donde  $E\{\cdot\}$  denota la esperanza estadística y  $^H$  la transpuesta conjugada.

Sean los autovalores  $\{\lambda_k\}$  ordenados en valores decrecientes con  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L > 0$  y los correspondientes autovectores como  $\{v_i, i=1, \dots, M\}$ . En ausencia de ruido, los autovalores  $\lambda_i, i=1, \dots, M$  son distintos de cero mientras  $\lambda_{M+1} = \lambda_{M+2} = \dots = \lambda_M = 0$ . Además, la matriz de covarianza se puede expresar como:

$$\mathbf{R} = \sum_{k=1}^M \lambda_k v_k v_k^H \quad (3.13)$$

Así, los autovectores  $v_i, i=1, \dots, M$  despliegan el mismo subespacio de señales que los vectores directores  $a_i(\theta), i=1, \dots, M$ . Estos  $M$  autovectores para el subespacio de señales se denominan *autovectores principales*; y los correspondientes autovalores se denominan *autovalores principales*.

En presencia de ruido, los autovalores son desplazados por la varianza del ruido i.e.  $\{\lambda_1 + \sigma^2, \dots, \lambda_M + \sigma^2, \sigma^2, \dots, \sigma^2\}$  y la matriz de correlación se puede escribir como:

$$\mathbf{R} = \sum_{k=1}^M \lambda_k v_k v_k^H + \sum_{k=M+1}^L \sigma^2 v_k v_k^H \quad (3.14)$$

La autodescomposición separa los autovectores en dos conjuntos. El conjunto  $\{v_i, i = 1, \dots, M\}$ , que son los autovectores principales, despliegan el subespacio de señales; mientras el subconjunto  $\{v_i, i = M + 1, \dots, L\}$ , que son ortogonales a los autovectores principales, se dice que pertenecen al subespacio de ruido. Como los vectores directores  $a_i(\theta), i = 1, \dots, M$  están en el subespacio de señales, se desprende que los  $\{a_i(\theta)\}$  son simplemente combinaciones lineales de los autovectores principales y que también son ortogonales a los vectores del subespacio de ruido.

Asimismo, se pueden definir  $\hat{\Pi}$  y  $\hat{\Pi}^\perp$  como las matrices de proyección a los subespacios de la señal y de ruido respectivamente

$$\hat{\Pi} = \sum_{i=1}^M v_i v_i^H \quad (3.15)$$

$$\hat{\Pi}^\perp = \sum_{i=M+1}^L v_i v_i^H = \mathbf{I} - \hat{\Pi} \quad (3.16)$$

### 3.4.2 Estimación de DOA usando Beamforming

La idea principal aquí es “dirigir” el arreglo en una dirección a la vez y medir la potencia de salida en busca de un máximo. Al usar el modelo paramétrico de datos mostrado en la sección 3.4.1 y el producto de dos vectores se define el estimador:

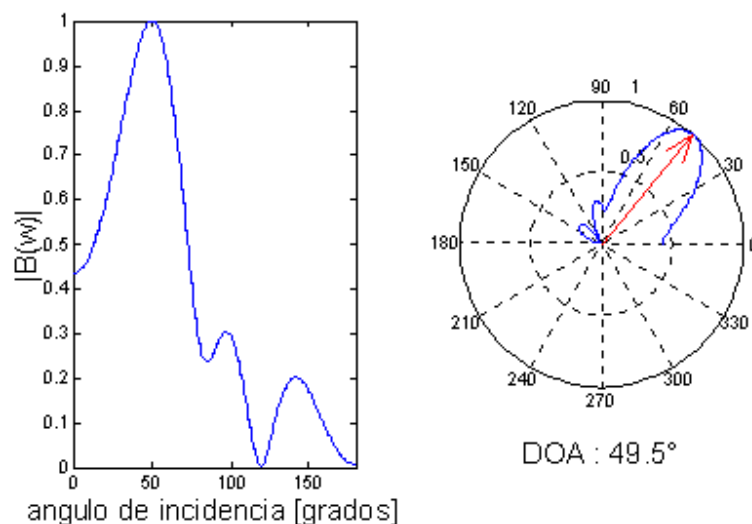
$$\mathbf{B}(\theta) = \frac{1}{N} \sum_{t=1}^N |\mathbf{b}(\theta)^H \mathbf{x}(t)|^2, \quad (3.19)$$

donde  $N$  indica la cantidad de puntos del vector  $\mathbf{x}(t)$  y  $\mathbf{b}(\theta)$  tiene la forma de *vector director*. De esta manera, al variar la dirección a donde apunta  $\mathbf{b}(\theta)$ , se formará un espectro espacial  $\mathbf{B}(\theta)$  con máximo en la dirección del ángulo de arribo.

Sin embargo, esta expresión no es la más indicada ya que requiere el uso de  $\mathbf{x}(t)$ , por lo que se opta por utilizar la matriz de covarianza estimada  $\hat{\mathbf{R}}$  [7, 25] como se muestra a continuación:

$$\mathbf{B}(\theta) = \frac{1}{N} \sum_{t=1}^N \mathbf{b}(\theta)^H \mathbf{x}(t) \mathbf{x}^H(t) \mathbf{b}(\theta) = \mathbf{b}(\theta)^H \hat{\mathbf{R}} \mathbf{b}(\theta) \quad (3.20)$$

En la figura 3.10 se muestra el resultado de una simulación con un ángulo de incidencia de  $45^\circ$ , una distancia entre sensores de 0.3 m y una relación señal a ruido de 20 dBs. En la gráfica de la derecha se puede apreciar fácilmente el resultado de la estimación señalado con una flecha roja.



**Figura 3.10** Estimación de un DOA de  $60^\circ$  usando Beamforming.

### 3.4.3 Estimación de DOA usando MUSIC

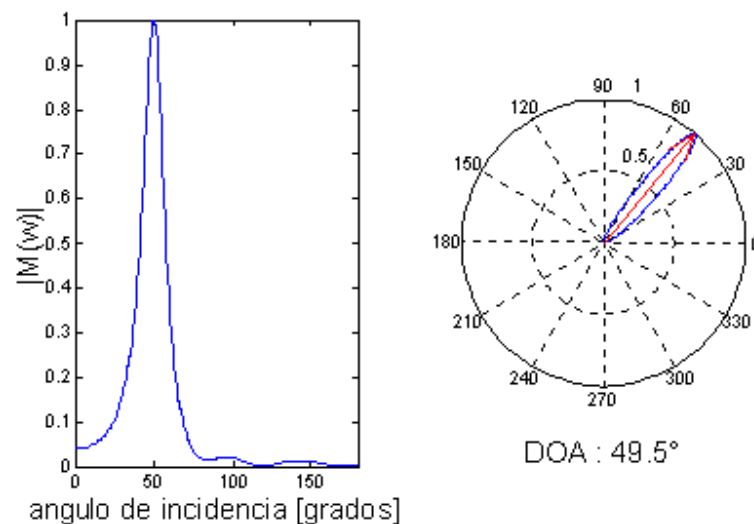
Una de las contribuciones más significantes sobre procesamiento de señales con arreglos fue la aplicación de la autodescomposición de la matriz de covarianza y sus propiedades para proveer una solución al problema de estimación en procesos observables [20, 21]. Dicha autodescomposición permite utilizar la matriz del subespacio de ruido para formar el espectro espacial  $M(\theta)$ , de la siguiente manera:

$$\mathbf{M}(\theta) = \frac{1}{\mathbf{b}(\theta)^H \hat{\Pi}^\perp \mathbf{b}(\theta)} \quad (3.21)$$

De igual forma que en el Beamforming, el algoritmo MUSIC plantea variar el vector director de búsqueda  $\mathbf{b}(\theta)$  para luego encontrar un máximo en el espectro espacial cuando el vector director apunte en la dirección de arribo de las señales [26].

Al apuntar  $\mathbf{b}(\theta)$  en la dirección de arribo de las señales este vector estará comprendido en el subespacio de señal; y como todo vector que pertenece al subespacio de señal será perpendicular al subespacio de ruido; de esta forma, la proyección  $\hat{\Pi}^\perp \mathbf{b}(\theta)$  valdrá cero y el estimador indicará un máximo.

En la figura 3.11 se muestra el resultado de una simulación de una señal con ángulo de incidencia de  $45^\circ$ , una distancia entre sensores de 0.3 m y una relación señal a ruido de 20 dBs. En la gráfica de la derecha se puede apreciar fácilmente el resultado de la estimación señalado con una flecha roja.



**Figura 3.11** Estimación de un DOA de  $60^\circ$  usando MUSIC.

### 3.4.4 Aproximación al Problema Real

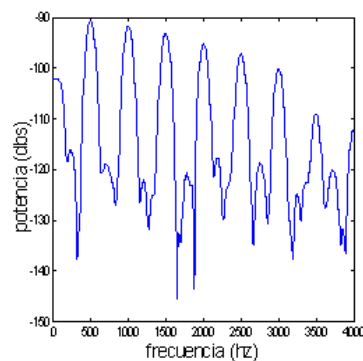
Toda la formulación necesaria para el modelo de datos suponía la existencia de cantidades exactas (i.e. tiempo de observación infinito); sin embargo, es obvio que en la práctica están disponibles sólo algunas estimaciones, las cuales se denotan con un sombrero (^). Una estimación natural de  $\mathbf{R}$  es la *matriz de covarianza estimada* [3]:

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{t=1}^N \mathbf{a}(\theta) \mathbf{a}^H(\theta) \quad (3.17)$$

Por lo que el problema reside en calcular el vector director  $\mathbf{a}(\theta)$  de los  $N$  cuadros en que se segmentó la señal. Dicho vector que tiene la forma mostrada en la ecuación 3.06, y en la práctica se obtiene mediante divisiones de los coeficientes de los espectros de frecuencia  $S_l(w_c)$ ; donde  $w_c = 2\pi f_c$  representa la frecuencia de la señal  $s(t)$ , por lo que se puede escribir:

$$\mathbf{a}_l(\theta) = \frac{S_l(w_c)}{S_1(w_c)} \quad (3.18)$$

Para determinar esta frecuencia  $f_c$  se debe calcular la densidad espectral de potencia (Power Spectral Density – PSD), y obtener la frecuencia máxima en el rango de frecuencias de interés; es decir, entre 300 hz y 566 hz como se mencionó anteriormente.



**Figura 3.12** Densidad Espectral de Potencia para un segmento de una señal sinusoidal real.



**Capítulo 4**  
**Simulaciones.**

Este capítulo se centra en simular respuestas de los algoritmos ante diferentes valores de los parámetros relevantes, los que intervienen en la estimación de la dirección de arribo de la fuente de sonido con el fin de mostrar el comportamiento del error de estimación ante variaciones específicas de cada parámetro.

El desarrollo de las simulaciones fue realizado en el entorno de MATLAB. Éste ofrece una gran facilidad al estar orientado a operaciones matriciales, así como el diseño y prueba de prototipos de sistemas complejos. Dicho programa permite el almacenamiento de las variables en archivos, los que se incluyen en el CD adjunto a este trabajo de tesis.

El desarrollo de las simulaciones considera una fuente de sonido, cuyas señales serán creadas con ayuda del entorno de MATLAB y que se basa en el modelo explicado en el capítulo 1. Para modelar el comportamiento del sonido en un cuarto se utilizó un método muy reconocido en su ámbito.

Las variaciones de los parámetros contemplan distintas formas de onda, distancia entre la fuente de sonido y el centro del arreglo de sensores, la dirección de arribo, la relación señal a ruido e incluso el nivel de reverberación de los ambientes [30, 31, 32].

Simulaciones posteriores fueron orientadas al caso real que será visto en el siguiente capítulo. Además, se mostrará el tiempo de procesamiento para cada una de estas últimas simulaciones, donde se podrá apreciar que este tiempo es muy pequeño para cualquiera de los métodos utilizados. Estos valores de tiempo de procesamiento al ser pequeños indican la alta posibilidad de una aplicación en tiempo real de este sistema.

#### 4.1 Generación de Señales Sintéticas

Para llevar a cabo las simulaciones presentadas en este capítulo, se utilizó el método de las imágenes [32] para hallar de la función de transferencia que atraviesa la señal generada por la fuente hasta la posición del micrófono; es decir, que se calculará 4 de estas funciones de transferencias o sus respuestas al impulso.

Para poder crear estas funciones de transferencias es necesario un modelo del ambiente donde se realizará el experimento virtual. El modelo debe contar con las dimensiones del ambiente y los coeficientes de absorción promedio para cada pared, piso y techo.

Los ambientes seleccionados para los experimentos se muestran en la tabla 4.1 y se pueden ordenar por el grado de reverberación:

Ambiente	Descripción	Perturbación
1	Exteriores de la sección de Ing. Electrónica.	Baja
2	Estudio de Grabación de la Facultad de Artes y Ciencias de la Comunicación.	Media

**Tabla 4.1** Relación de ambientes de los experimentos.

Según estos ambientes, se escogieron los coeficientes de absorción mostrados en la tabla 4.2, los cuales dan valores de tiempos de reverberación  $T_{60}$  muy similares a los medidos en la realidad. En el ambiente 1 el tiempo de reverberación es nulo y por tanto se asignan coeficientes de absorción ideales. En cambio, en el ambiente 2, se determina mediante los valores típicos de los diversos materiales. Según estos últimos valores se calculó un tiempo de reverberación de 0.8 segundos lo cual se aproxima mucho a la realidad. Como se verá en el siguiente capítulo, el tiempo de reverberación  $T_{30}$  fue aproximadamente de 0.5 segundos (valor que indica un ambiente de poca reverberación).

Superficie	Material	Largo (mts)	Ancho (mts)	$\alpha_{\text{ambiente1}}$ @500hz	$\alpha_{\text{ambiente2}}$ @500hz
Pared derecha	Vidrio	8	3	1	0.03
Pared izquierda	Cortinas	8	3	1	0.15
Pared frontal	Cortinas	6	3	1	0.15
Pared trasera	Cortinas	6	3	1	0.15
Techo	Concreto	6	8	1	0.1
Piso	Alfombra	6	8	1	0.25

**Tabla 4.2** Coeficientes de absorción de los ambientes.

A partir de estos coeficientes se generan las 4 funciones de transferencia para una posición dada de la fuente de sonido y del arreglo de sensores, mediante el método de las ocho imágenes [32]. El siguiente paso consiste en convolucionar una forma de onda determinada con las respuestas impulsivas de cada una de las funciones de transferencia para obtener las señales que cada micrófono captaría en ausencia de ruido.

Para las simulaciones realizadas se utilizaron las formas de onda mostradas en la figura 4.1, las cuales se diferencian principalmente en sus componentes de frecuencia. La primera señal es un tono de 500 hz generado por el parlante JBL (ver anexo B). La segunda señal es voz humana (específicamente la vocal 'a') y la tercera, es ruido blanco generado por parlante.

El último paso en la generación de las señales sintéticas consiste en añadir ruido blanco según el nivel de SNR deseado. Aunque se sabe que el ruido blanco de fondo en los ambientes no es totalmente plano, se utilizó este tipo de ruido tomando en consideración que solo se analizará una banda de frecuencia (de 300 a 500 Hz) y que en esta banda el ruido de fondo registrado por los micrófonos es casi plano como se muestra en la figura 4.2. Tómese en cuenta también que la SNR observable tanto en aplicaciones de voz como en los experimentos realizados tiene valores entre 15 y 20 dBs, donde el ruido del ambiente será muy pequeño.

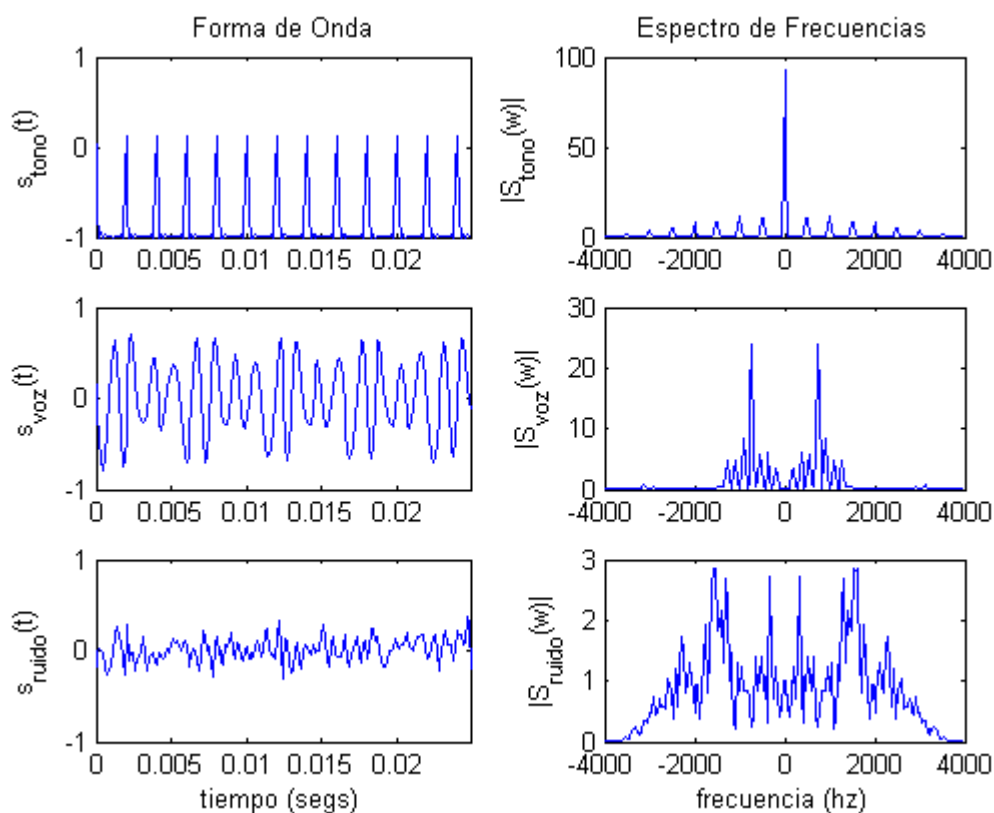


Figura 4.1 Formas de onda usadas para generar las señales sintéticas.

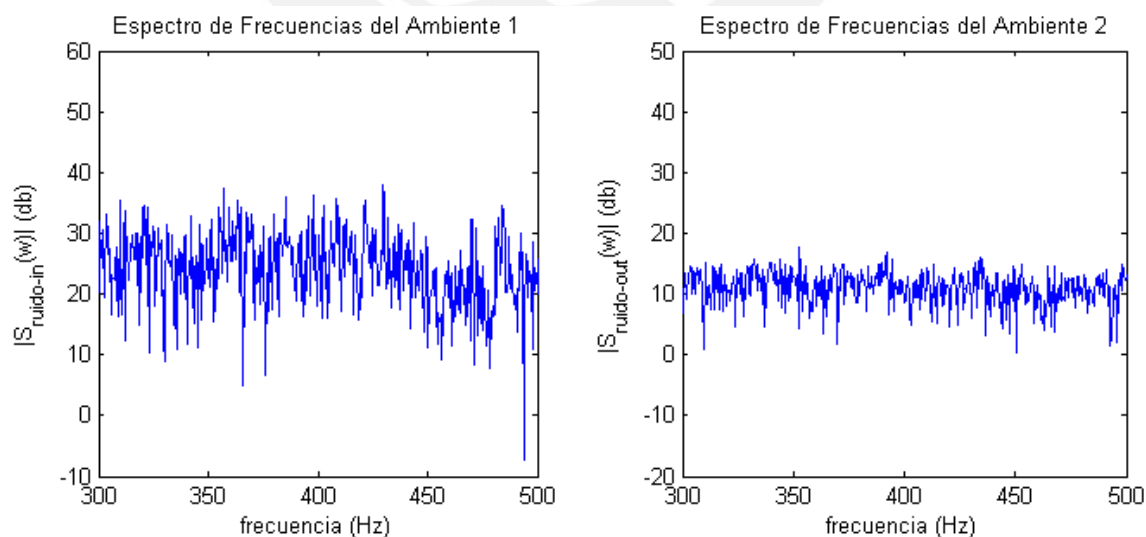


Figura 4.2 Patrón de ruido de fondo de ambos ambientes.

Se sabe que la relación señal a ruido (SNR) se obtiene con la ecuación 4.1, por lo que para obtener una señal con cierta SNR a partir de una potencia de señal dada, habrá que añadir un cierto nivel de ruido.

$$SNR = 10 \log \left( \frac{\text{Potencia}_{\text{señal}}}{\text{Potencia}_{\text{ruido}}} \right) \quad (4.1)$$

La potencia de ruido tiene relación directa con la varianza  $\sigma^2$  [22] y ésta se muestra en la siguiente ecuación:

$$\sigma^2 = \frac{1}{N} \sum_{n=1}^N y^2(n) \quad (4.2)$$

Donde  $y(n)$  es la secuencia de longitud  $N$  de ruido aleatorio blanco gaussiano no correlacionado. Una vez añadido el ruido, se podrá aplicar los métodos de estimación explicados en el capítulo anterior.

#### 4.2 Determinación del Campo Lejano

Para determinar la distancia entre la fuente y el centro del arreglo de sensores en la cual se puede tener una onda plana, se debe escoger una distancia mayor a la longitud del arreglo. En la figura 4.3, se muestra una simulación sin reverberación y en ausencia de ruido para obtener la variación del error de estimación en función de la distancia entre la fuente y el centro del arreglo de 4 micrófonos separados uniformemente 30 cm. Esto quiere decir que la longitud del arreglo será de 0.9 metros y, según el error de estimación obtenido, es razonable afirmar que al colocar una fuente a más de un metro del arreglo de sensores para generar ondas planas.

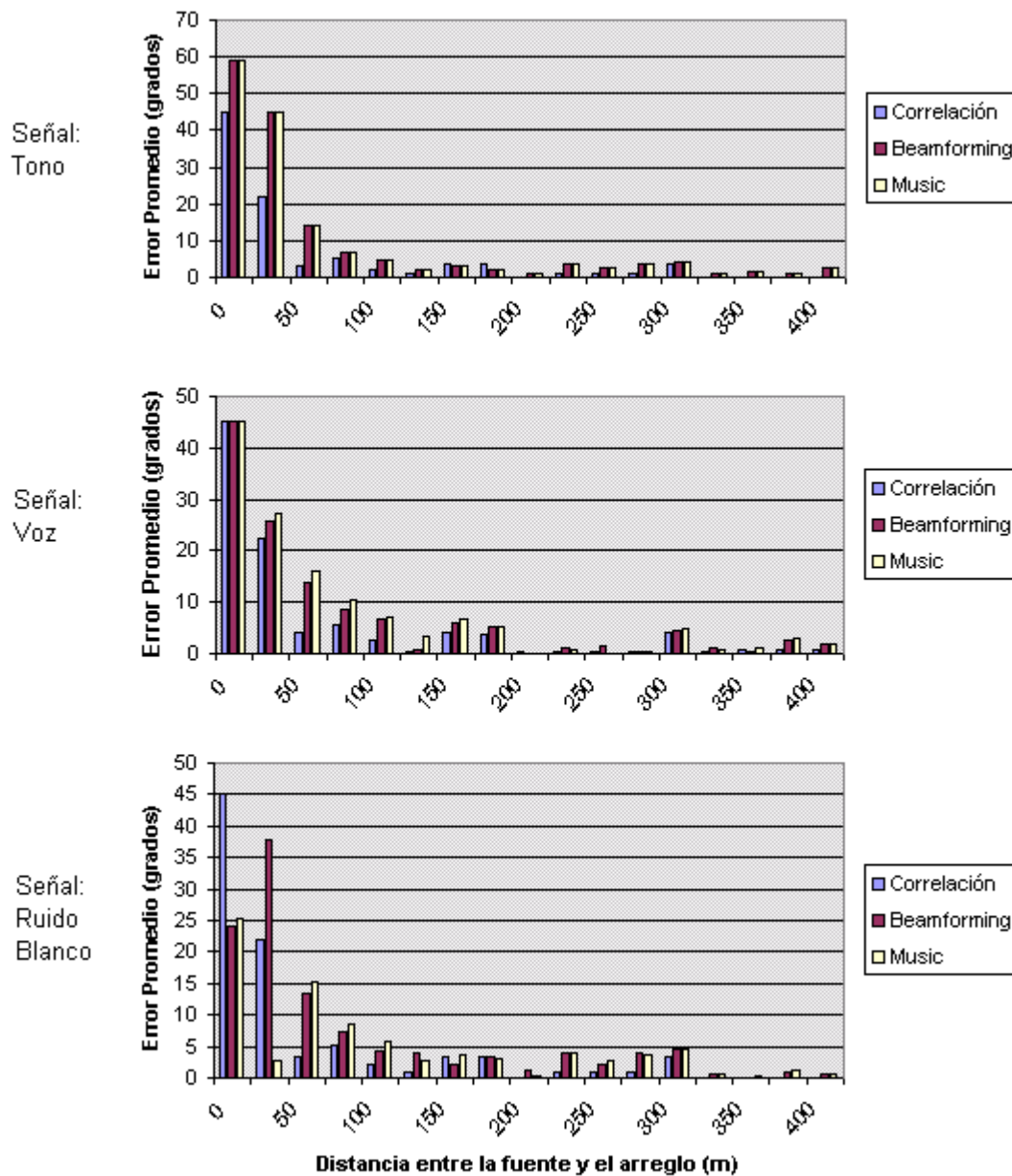
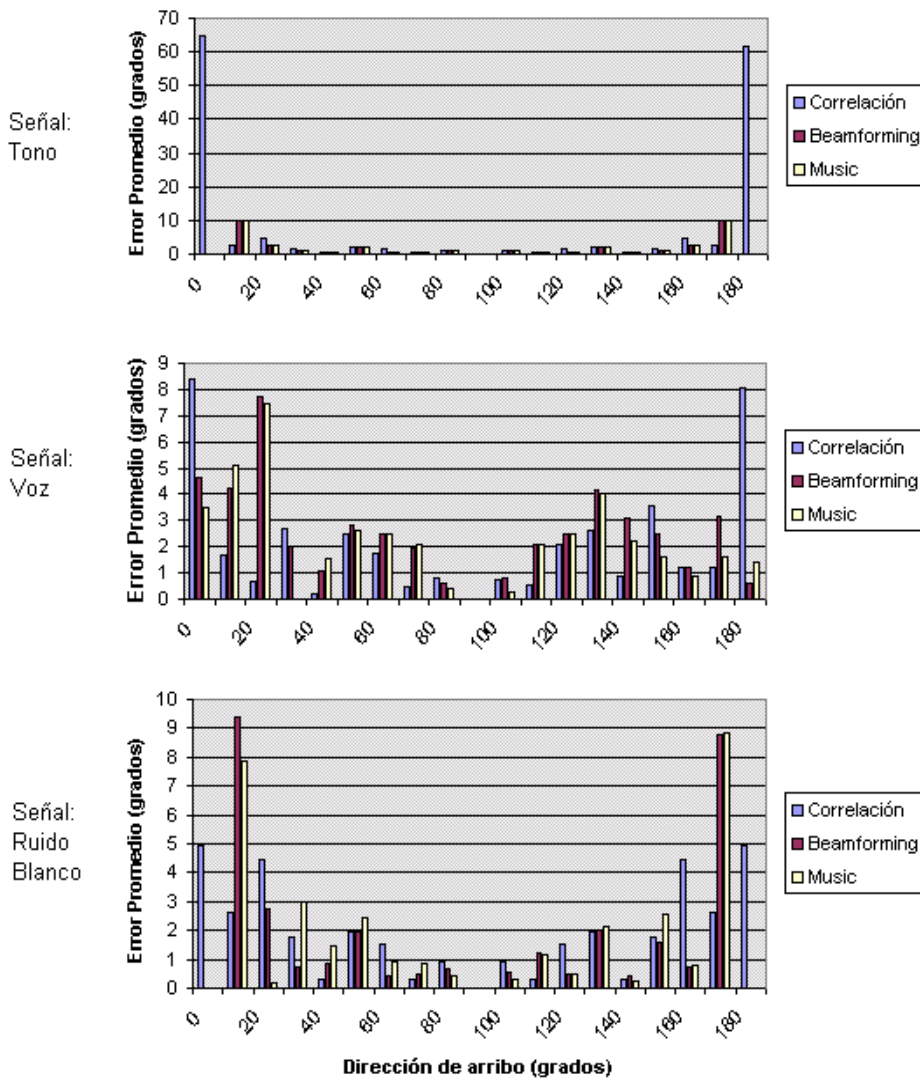


Figura 4.3 Variación del Error Promedio en función de la distancia entre la fuente y el arreglo de sensores.

### 4.3 Error de Estimación Para Diferentes Direcciones de Arribo

No está demás observar la respuesta de los algoritmos con señales y con diferentes direcciones de arribo. Al hacer una simulación sin reverberación y en ausencia de ruido, se puede ver que la variación del error es pequeña.



**Figura 4.4** Variación del Error Promedio en función del ángulo de arribo para una fuente situada a 2m del arreglo de sensores, sin interferencias.

#### 4.4 Influencia de la Relación Señal a Ruido

Un parámetro importante para una buena estimación es la relación señal a ruido. (SNR). Dependerá de esta relación que se tenga una buena estimación. En una simulación sin reverberación se puede apreciar que existe la tendencia a un incremento del error para SNR cada vez más bajas.

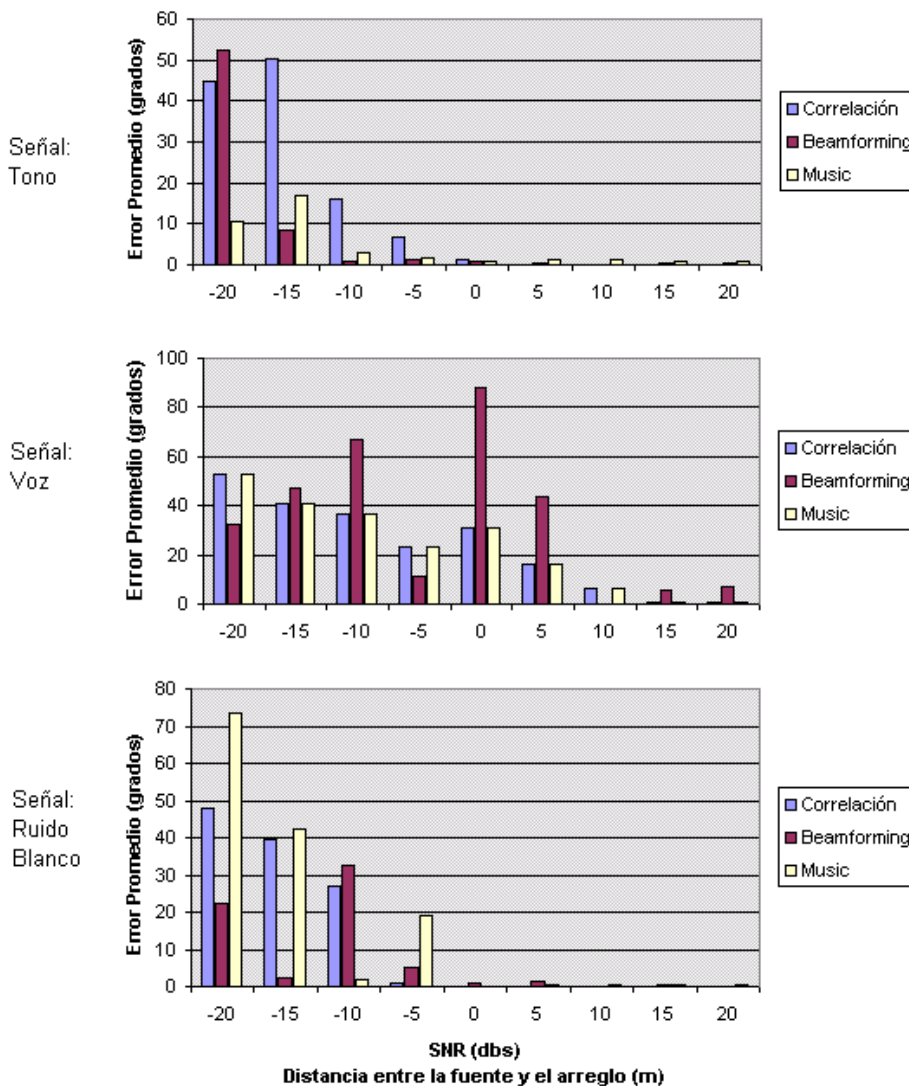
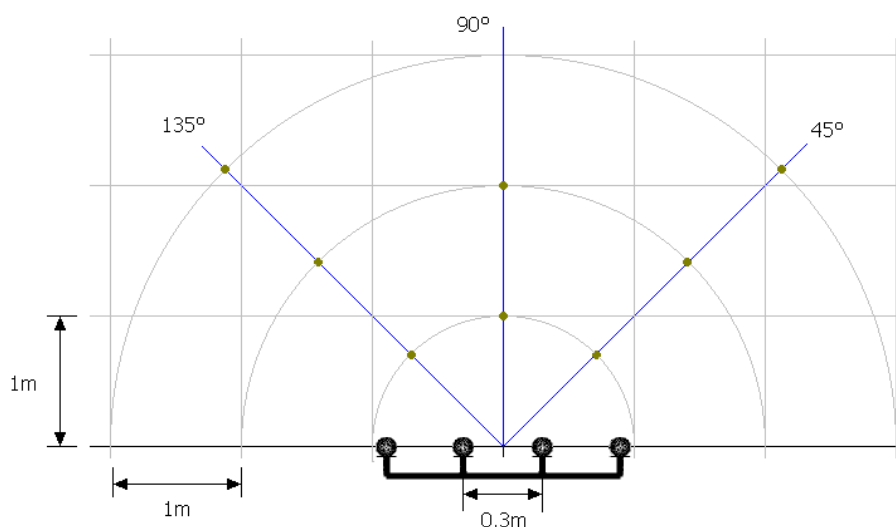


Figura 4.5 Variación del Error Promedio en función de la relación señal a ruido.

#### 4.5 Simulación de Casos Específicos

Las simulaciones se realizaron con una frecuencia de muestreo de 8 khz, una relación señal a ruido de 20 dBs, 10 cuadros por experimento y 10 experimentos en total. En las siguientes figuras se mostrará el promedio y la varianza de los 10 errores de estimación de la dirección de arriba.

En la figura 4.6 se puede apreciar las 8 diferentes posiciones donde se ubicó la fuente de sonido con puntos verdes. Cabe resaltar que estos puntos de prueba también se utilizaron al realizar los experimentos que se analizarán en el siguiente capítulo. De esta manera se podrá tener una idea del tipo de error de estimación que se encontrará.



**Figura 4.6** Ubicación gráfica de los puntos de prueba.

Cabe mencionar que estas posiciones fueron escogidas tales como se muestran debido a que en cierta medida las demás posiciones podrían generar resultados predecibles. Fuentes más alejadas generarían resultados muy similares a las posiciones con distancias máximas, pues que el alejar la fuente de sonido se produciría frentes de ondas más planas, lo cual disminuiría el error de estimación, pero el SNR de las señales, disminuiría y el error de estimación aumentaría; principalmente en el caso de la señal de voz.

### 4.5.1 Error de Estimación Promedio

Los resultados indican que con un nivel bajo de reverberación como es el que existe al aire libre es posible obtener una buena estimación con cualquiera de las señales analizadas. Al aumentar el nivel de reverberación el error de estimación aumenta considerablemente. Nótese que el error es menor cuando la dirección de arribo real es 90°, lo cual indica que el error crece conforme la dirección de arribo es más próxima a 0° o a 180°.

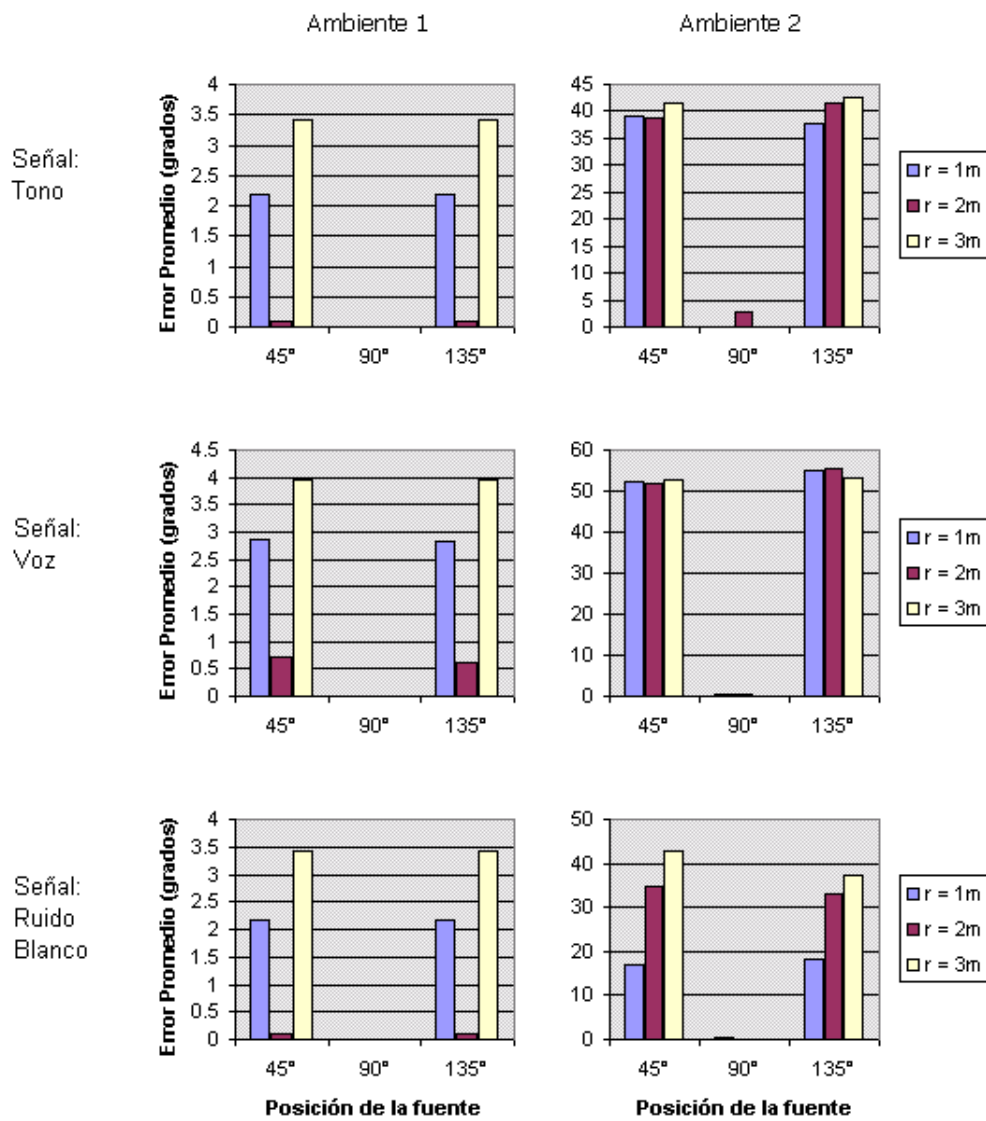


Figura 4.7 Error Promedio al usar Correlación para la estimación del DOA.

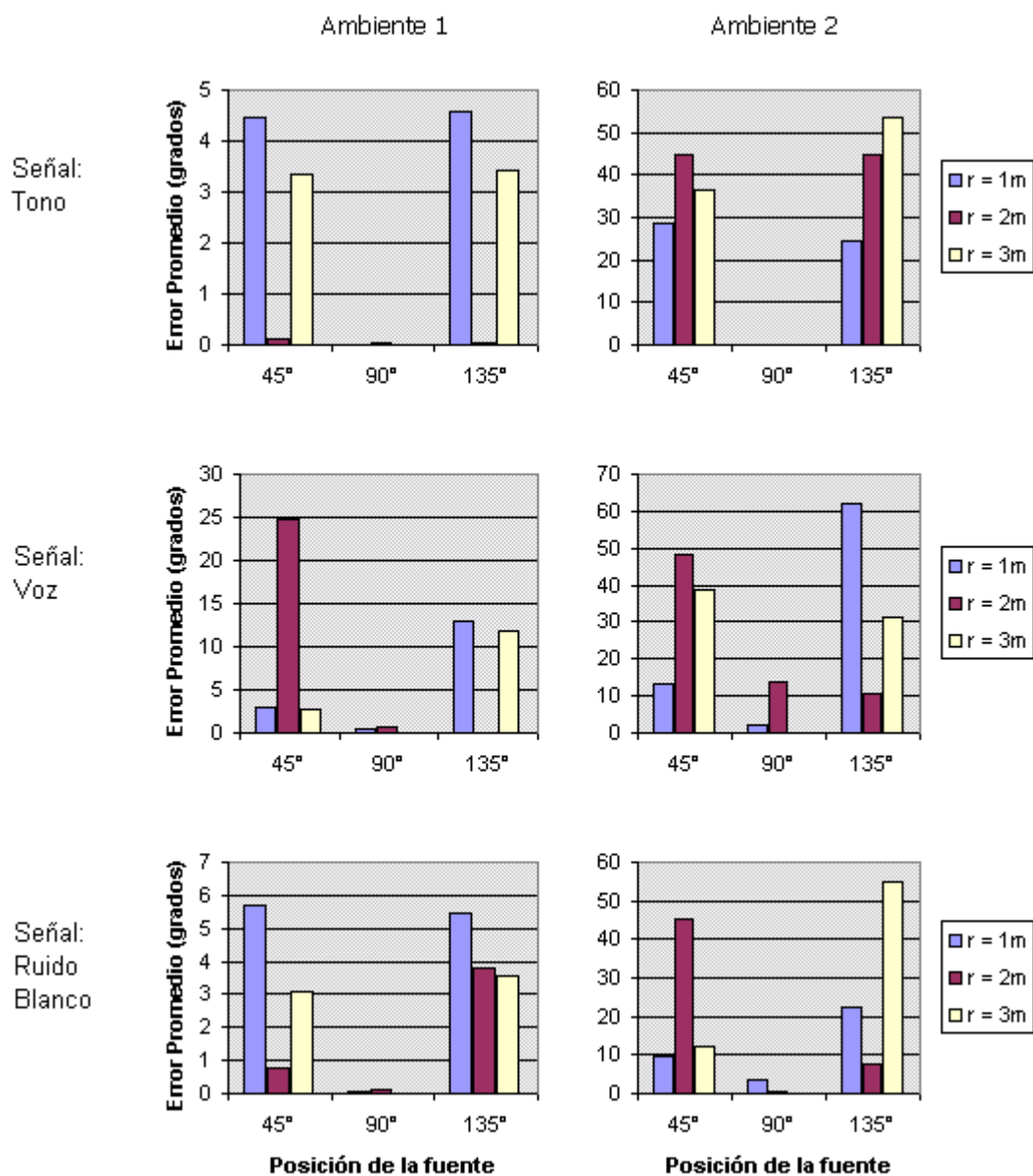


Figura 4.8 Error Promedio al usar Beamforming para la estimación del DOA.

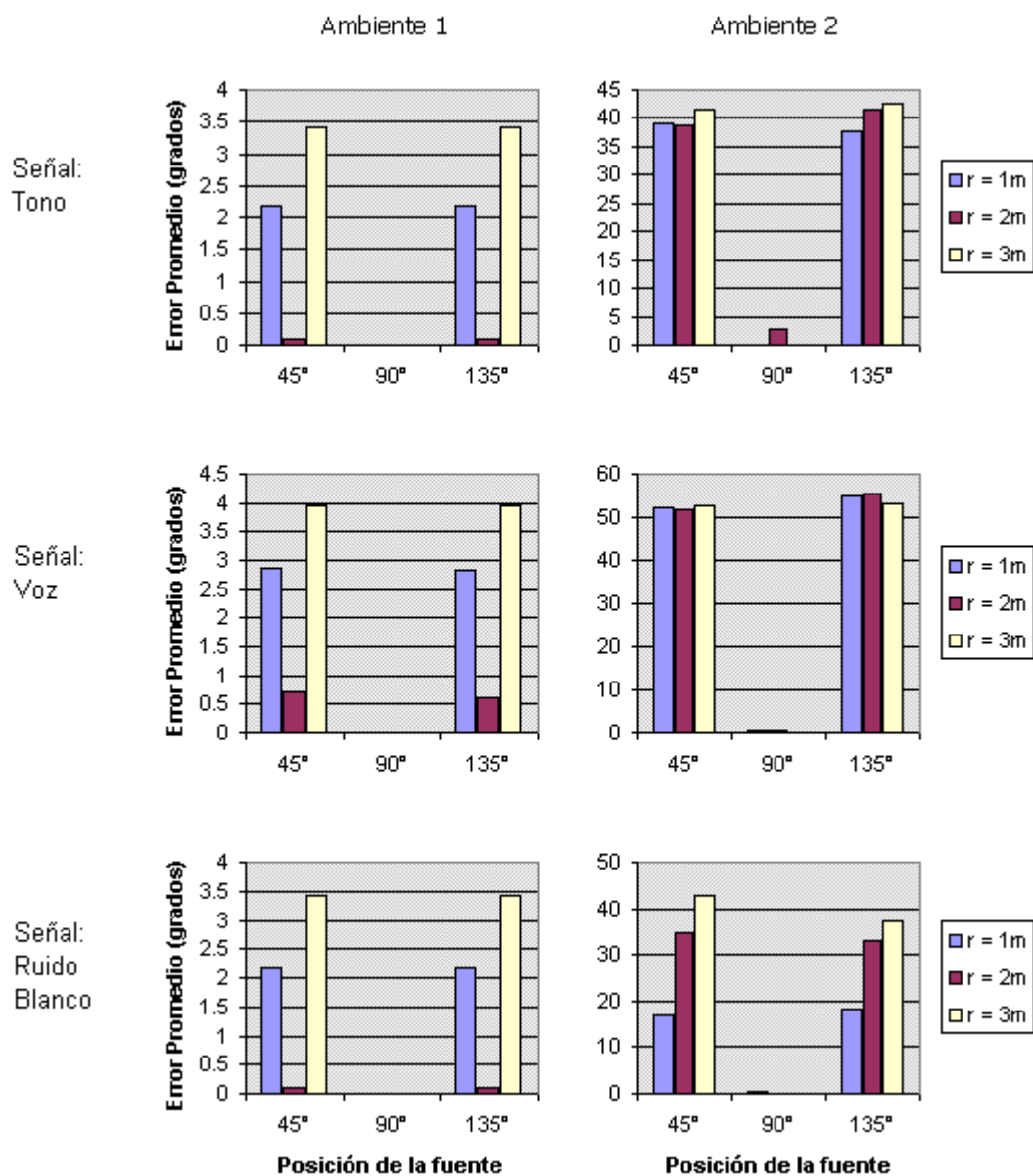


Figura 4.9 Error Promedio al usar MUSIC para la estimación del DOA.

### 4.5.2 Varianza del Error de Estimación

En las siguientes 3 figuras se muestra la varianza en cada uno de los casos, lo cual nos indica qué tanto varía el error en los diferentes experimentos realizados. Se puede ver que en la mayoría de los casos la varianza crece con el aumento del nivel de reverberación y que con un nivel bajo de reverberación la varianza es muy pequeña, lo que quiere decir, que el resultado es consistente.

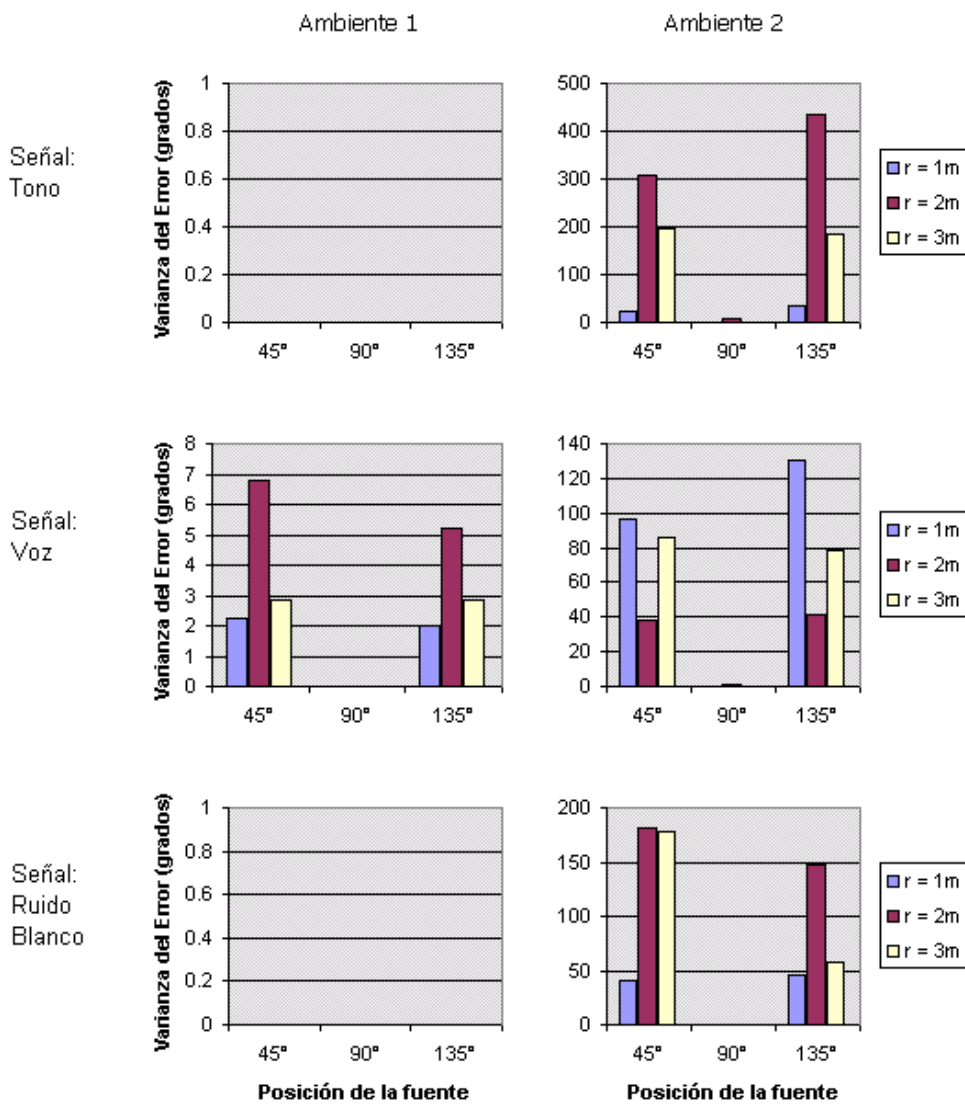


Figura 4.10 Varianza del Error de estimación del DOA al usar Correlación.

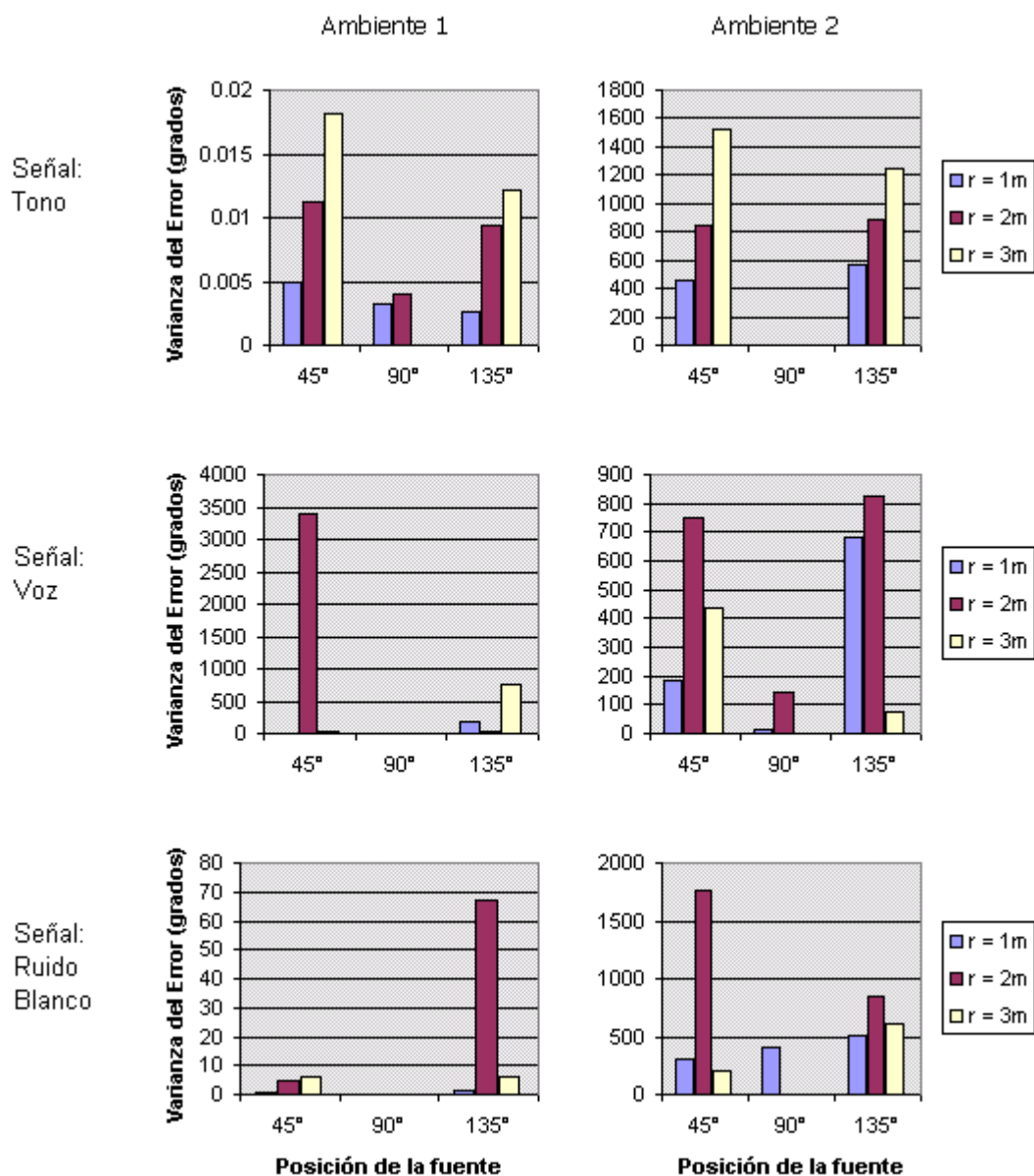


Figura 4.11 Varianza del Error de estimación del DOA al usar Beamforming.

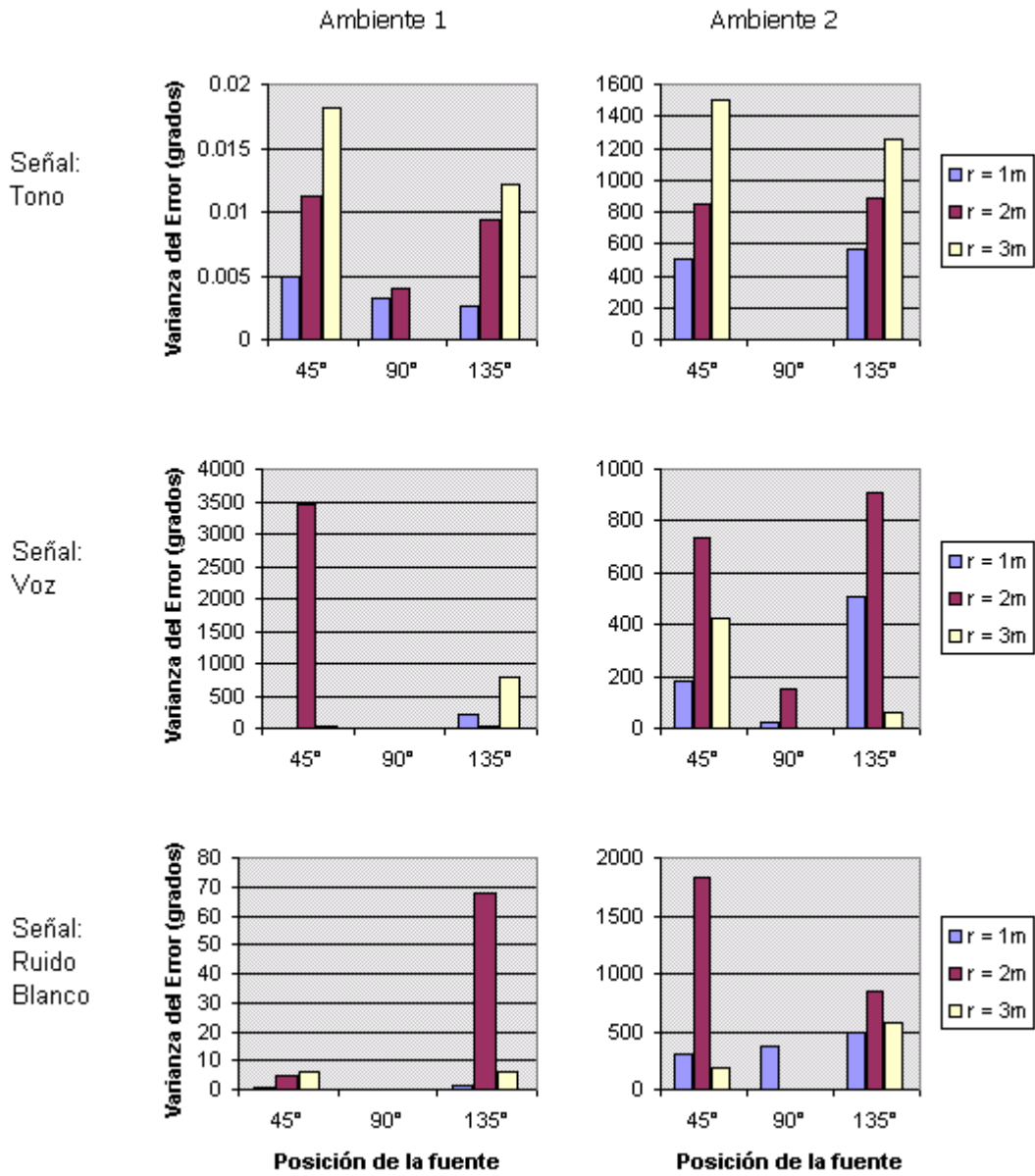


Figura 4.12 Varianza del Error de estimación del DOA al usar MUSIC.

### 4.5.3 Tiempos de Procesamiento

Se debe hacer notar que los tiempos de procesamiento están en el orden de los milisegundos para la estimación en ambos dominios, lo cual sugiere que una implementación de este sistema en tiempo real es factible, además de la posibilidad de utilizar algoritmos más robustos ante la reverberación.

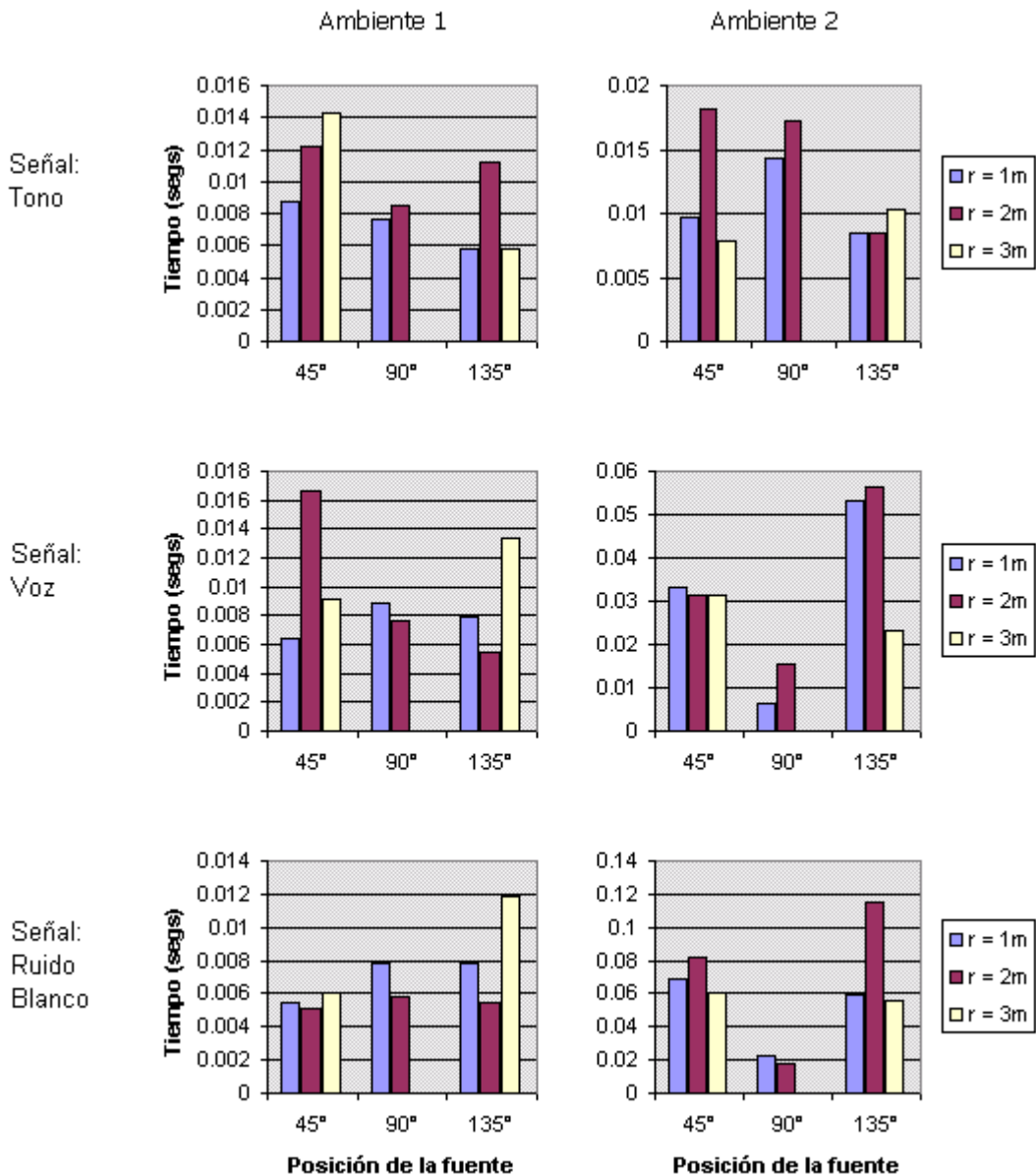


Figura 4.13 Tiempo de procesamiento de un experimento en el dominio del tiempo.

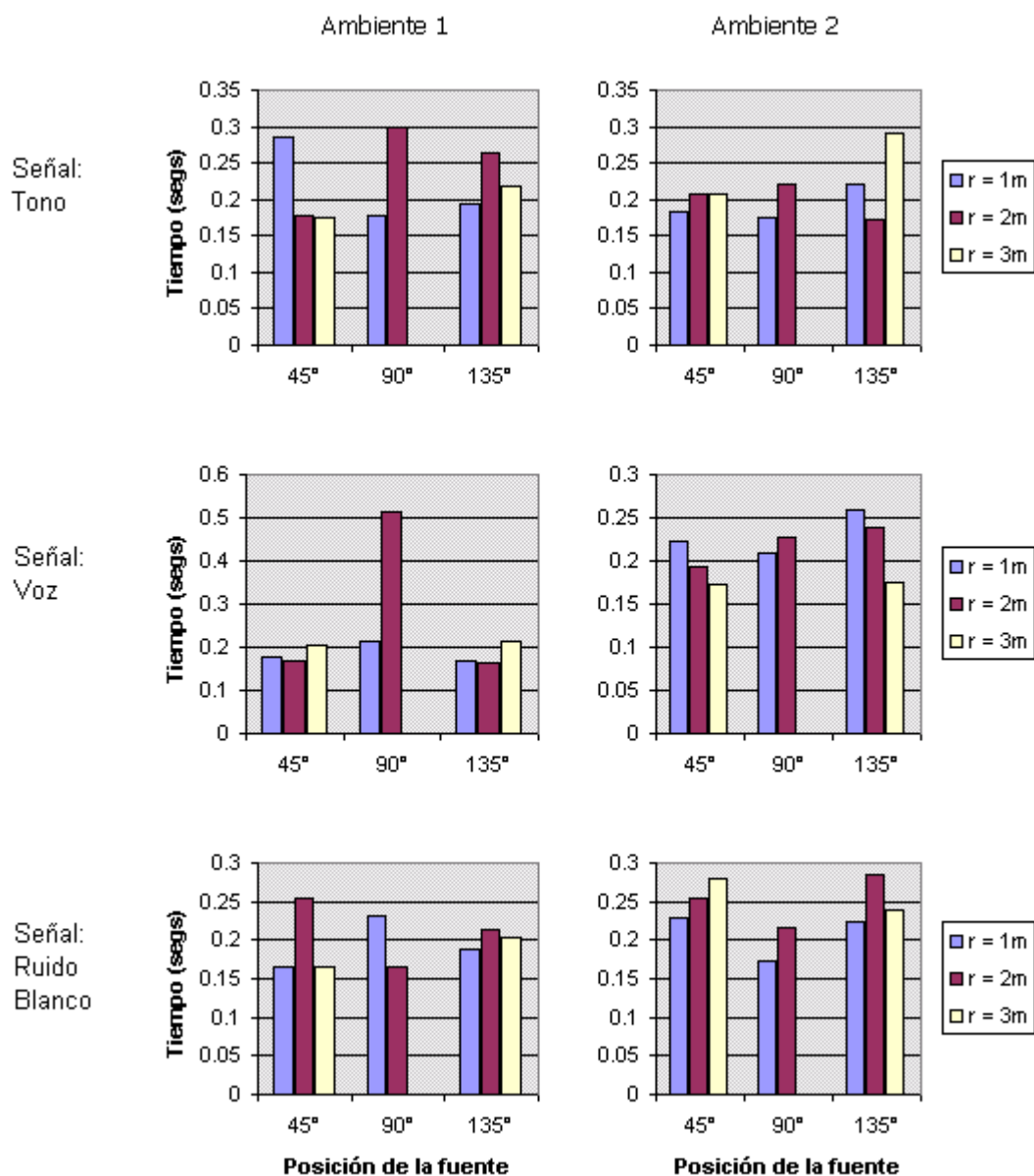


Figura 4.14 Tiempo de procesamiento de un experimento en el dominio de la frecuencia.



**Capítulo 5**  
**Pruebas y Resultados**

En este capítulo se mostrarán los experimentos realizados con fuentes reales en diferentes condiciones. Los resultados de estos experimentos serán la media y la varianza del error entre los valores reales y la estimación de los algoritmos de la ubicación de la fuente.

De acuerdo a lo comentado en el capítulo anterior, se mostrarán los resultados de las estimaciones en 2 ambientes: los exteriores de la sección de Ing. Electrónica y el estudio de grabación de radio de la facultad de Ciencias y Artes de la Comunicación. Lugares que se diferencian principalmente en el nivel de reverberación al cual está sometida la fuente de sonido.

En la figura 5.1 se puede apreciar una vista del primer ambiente con la disposición del arreglo lineal de micrófonos y la fuente de sonido lista para generar un frente de onda plano a  $45^\circ$ . Nótese que el efecto de reverberación es casi nulo ya que no hay muchas superficies que reflejen el sonido hacia el arreglo.



**Figura 5.1** Exteriores de la sección de Ing. Electrónica

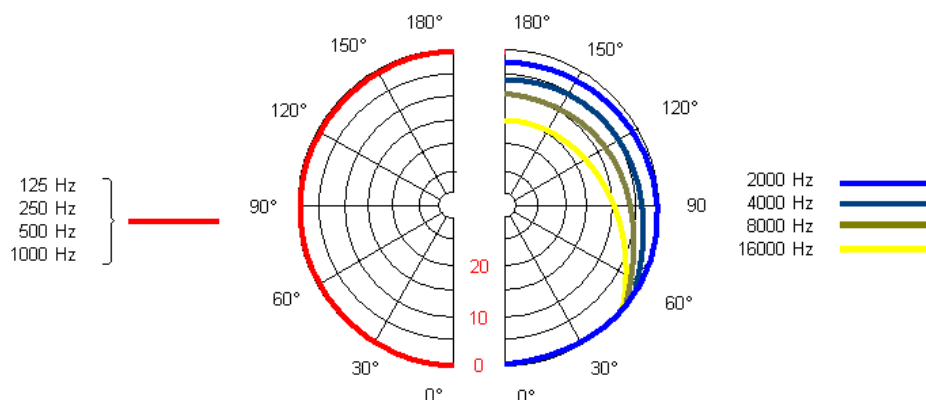
Del mismo modo que en la figura 5.1, en la figura 5.2 se puede apreciar una vista del segundo ambiente con la fuente de sonido y el arreglo lineal de micrófonos. En este caso, sí existe el problema de reverberación y tal como se mostró en la tabla 4.5, los coeficientes de absorción son considerables.



**Figura 5.2** Estudio de Grabación de la facultad de Ciencias y Artes de la Comunicación.

Las posiciones donde se ubicó la fuente de sonido se pueden apreciar en la figura 4.6. Las distancias de la fuente al centro del arreglo fueron 1, 2 y 3 metros para las direcciones de arribo de 45° y 135° y de 1 y 2 metros para la dirección de arribo de 90°. Se seleccionaron estas posiciones pues se requiere trabajar con ondas planas para obtener una buena estimación.

Como se indicó anteriormente los micrófonos tienen un patrón omnidireccional, (ver figura 53) siendo esta característica la que permite un arreglo lineal y no todo un sistema electromecánico para analizar las posibles direcciones de donde proviene la señal.



**Figura 5.3** Patrón omnidireccional de los micrófonos.

Las señales que se analizaron fueron 3 como se mostró en la figura 4.1, siendo éstas un tono (a 500 Hz), voz (vocal 'a') y ruido blanco. La diferencia entre estas señales radica en el contenido frecuencial. El tono consiste en una señal que concentra su energía en una sola frecuencia, la voz no tiene un patrón específico de componentes de frecuencia y el ruido blanco contiene todas las frecuencias con un mismo nivel. En cuanto a la generación de estas señales, se utilizó un parlante para el tono y el ruido blanco; mas no para la voz, la cual es producida naturalmente.

Los resultados que se mostrarán en este capítulo provienen de analizar señales muestreadas a 8khz y como se mencionó antes, la frecuencia de muestreo puede influir en una mejora de la precisión de los resultados. Cabe mencionar que se realizaron grabaciones con una frecuencia de muestreo de 44khz, con las cuales se obtuvo resultados muy similares a los que mostrarán aquí. Si se desea comparar estos resultados se podrá utilizar los archivos incluidos en el CD adjunto a este trabajo de tesis.

Como se mencionó al inicio el tiempo de reverberación nos puede dar una idea del ambiente y según las mediciones realizadas en el estudio de grabación se obtuvo un tiempo de reverberación T60 de aproximadamente 1 segundo, si se usa como base el T30 @500hz, cercano al valor usado para crear las señales sintéticas en el caso de la simulación en MATLAB de la estimación de DOA's.

	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz
T30 (MLS)	0.857	0.721	0.503	0.301	0.297	0.273
T30 (Sweep)	0.828	0.677	0.518	0.29	0.291	0.273

**Tabla 5.1** Tiempo de reverberación T30 (en segundos).

### 5.1 Relación señal a ruido de los datos experimentales

A continuación se mostrará la relación señal a ruido para cada una de las posiciones y ambientes de prueba de las señales escogidas. Como el propósito de esta investigación se centra en hallar la dirección de arribo, la potencia de la señal emitida no fue exactamente igual en todos los casos; lo que se buscó fue un buen SNR para hacer la estimación correcta.

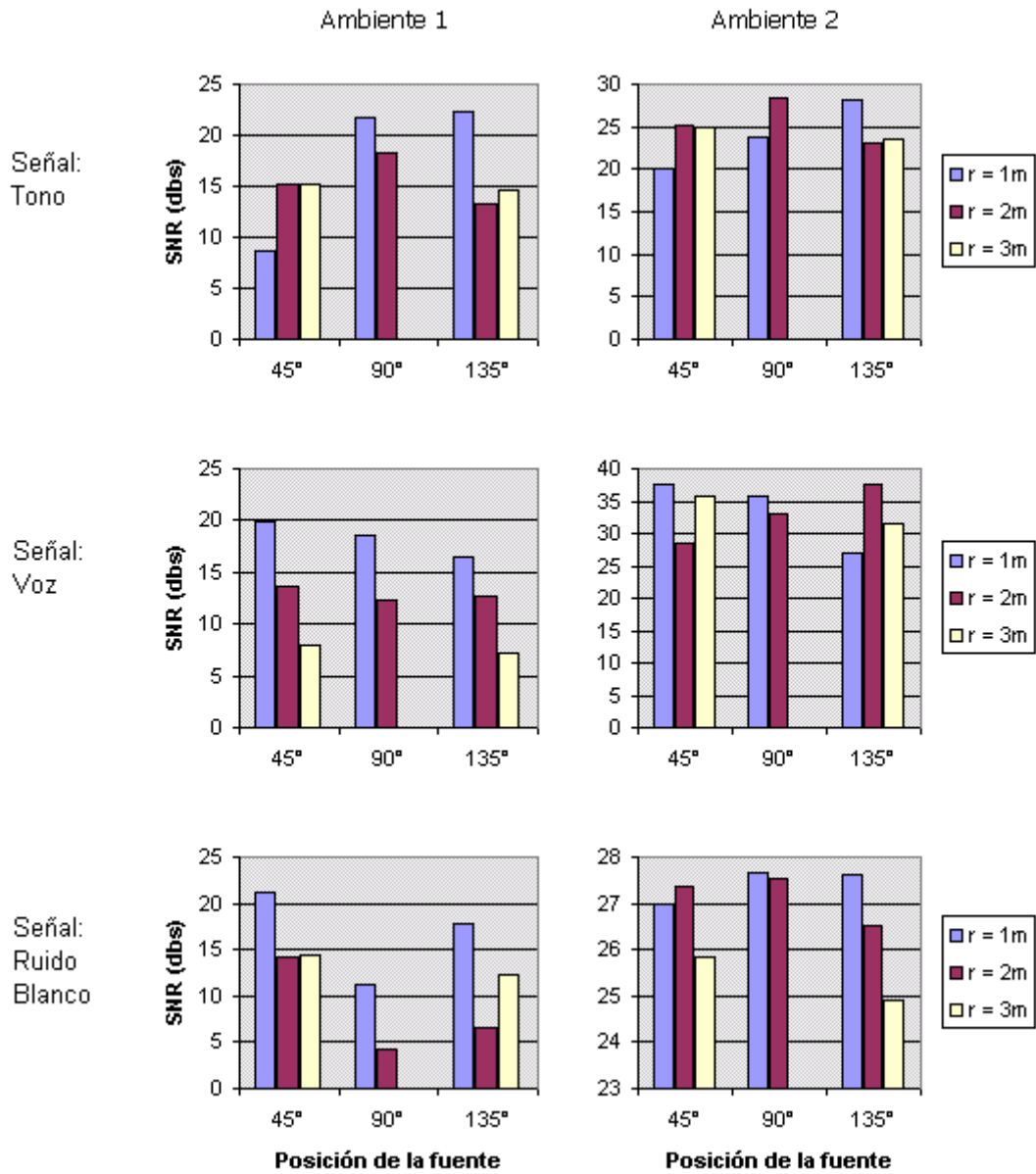


Figura 5.4 Relación señal a ruido (SNR).

### 5.2 Error de Estimación Promedio

En las siguientes figuras se muestra la respuesta de los algoritmos de estimación. Como es de esperar, se puede apreciar en los tres métodos de estimación, que el error tiende a incrementarse debido a un aumento del nivel de reverberación y de la distancia entre la fuente de sonido y el centro del arreglo de micrófonos.

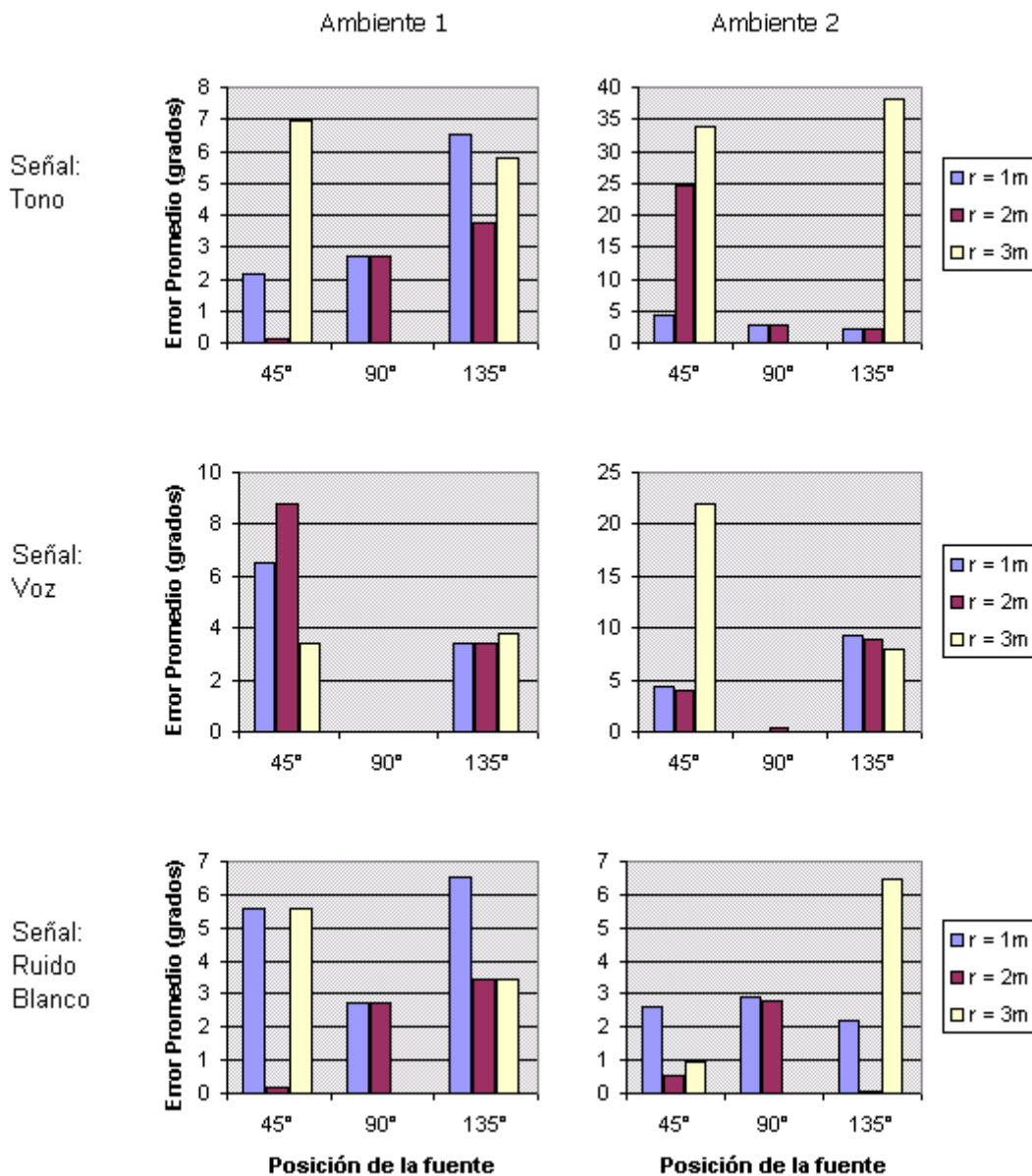


Figura 5.5 Error Promedio al usar Correlación para la estimación del DOA.

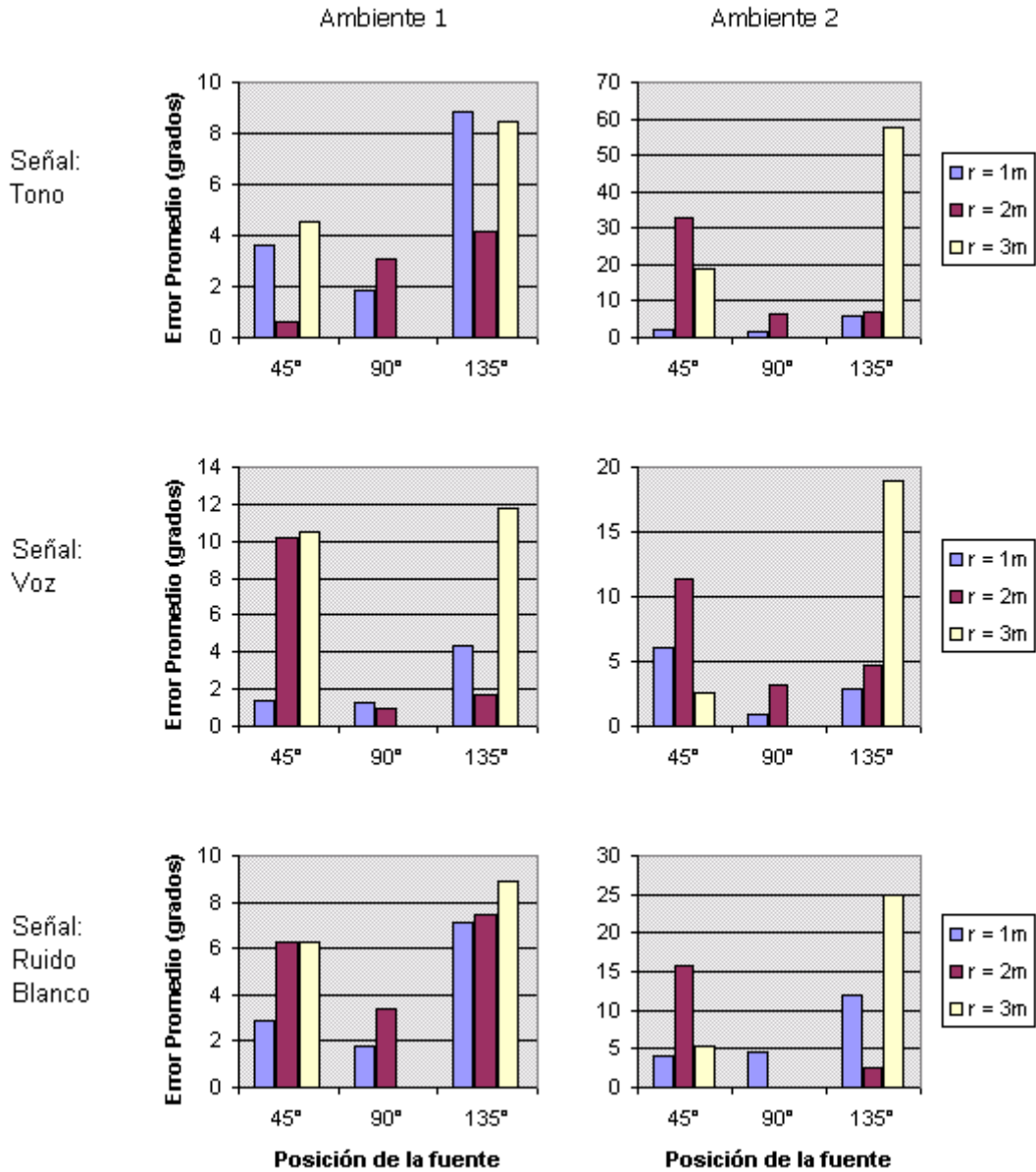
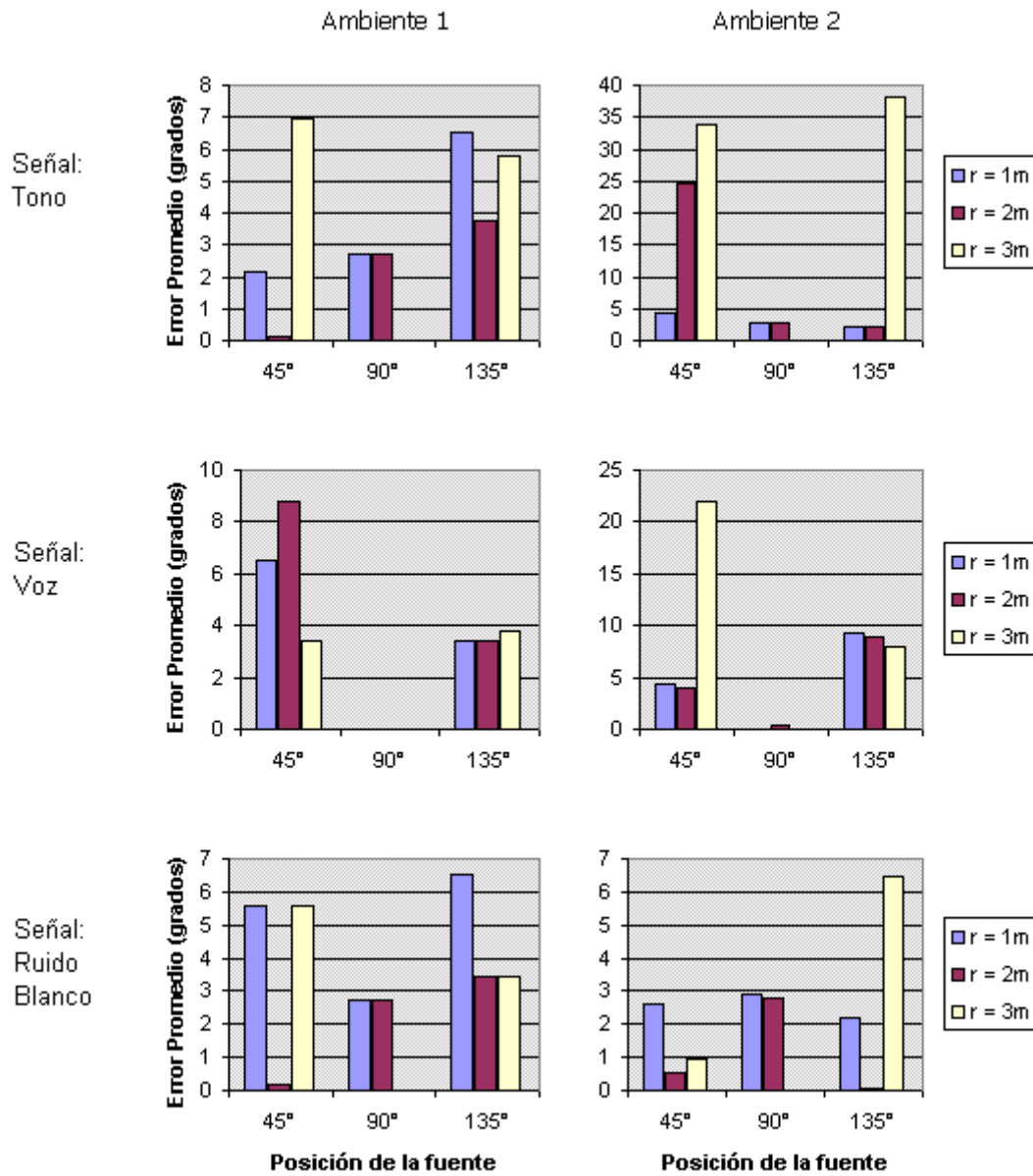


Figura 5.6 Error Promedio al usar Beamforming para la estimación del DOA.



**Figura 5.7** Error Promedio al usar MUSIC para la estimación del DOA.

Dependiendo de la aplicación deseada, ahora se puede tener una idea del error de estimación, notando un error no muy grande en el caso de voz y ruido blanco en ambos ambientes; y un error pequeño en un ambiente de bajo nivel de reverberación para cualquier tipo de señal.

### 5.3 Varianza del Error de Estimación

Como complemento del error promedio de la estimación, se presenta también la varianza de este error, la cual muestra un aumento en el caso de mayor reverberación.

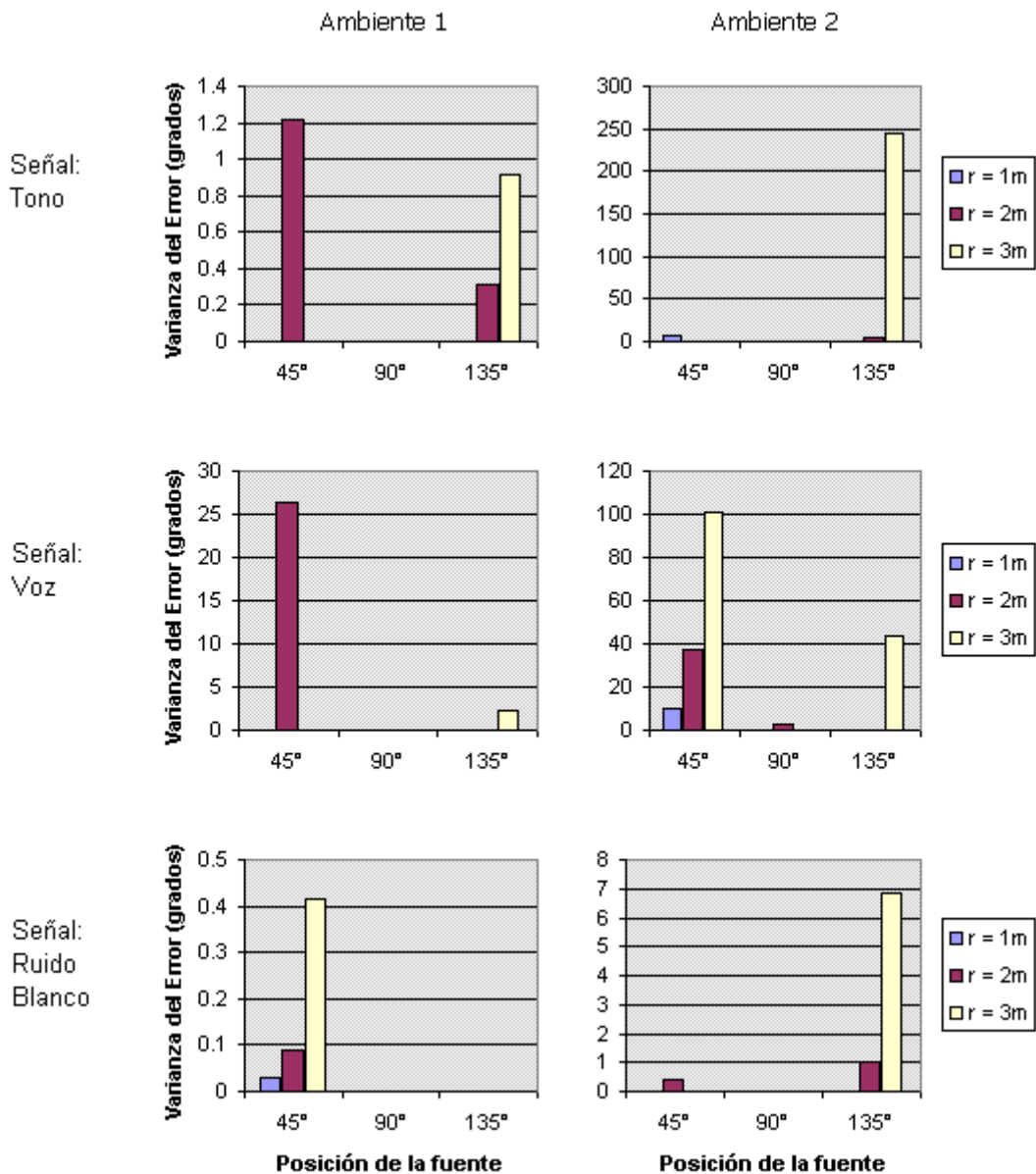


Figura 5.8 Varianza del Error de estimación del DOA al usar Correlación.

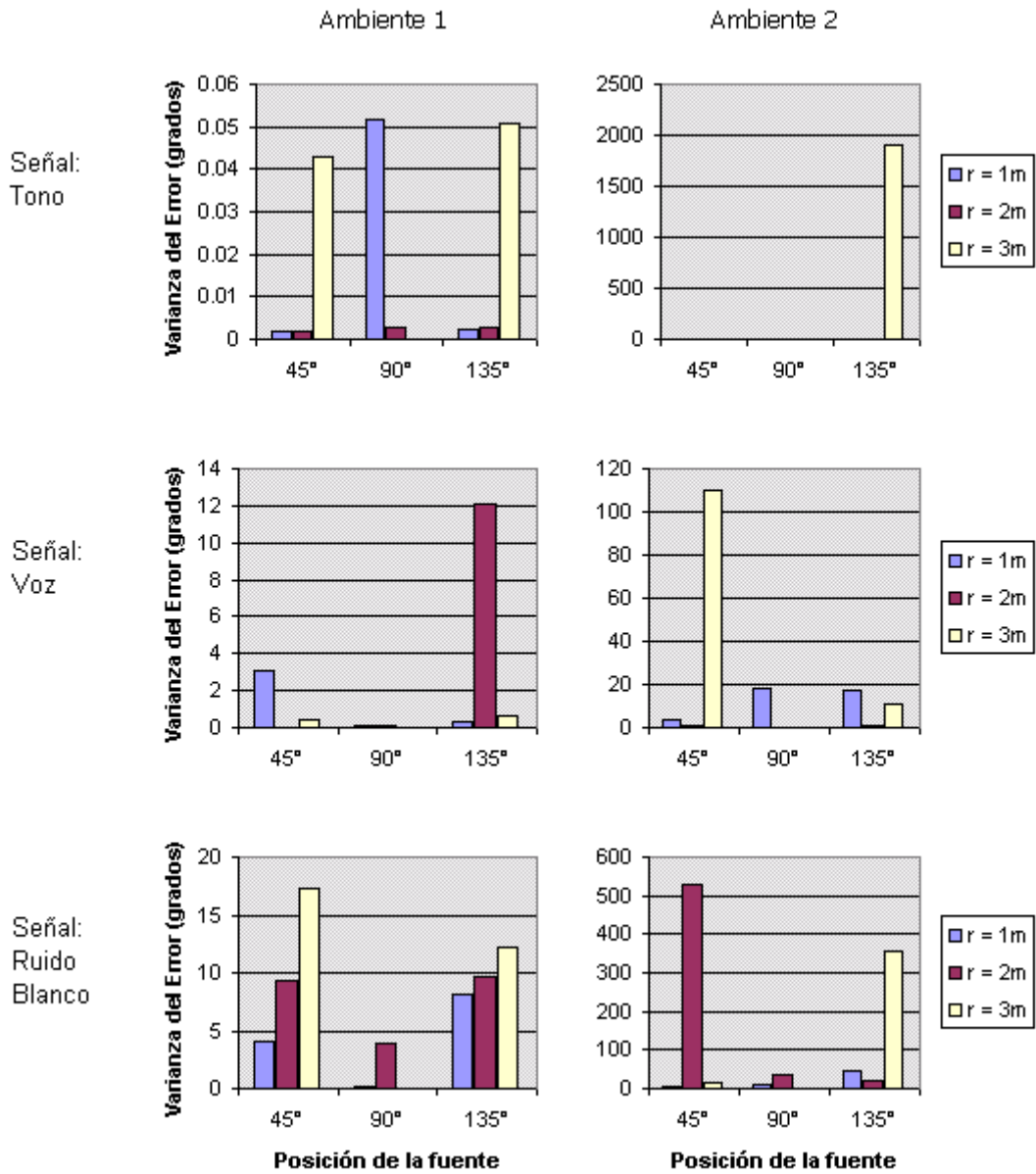


Figura 5.9 Varianza del Error de estimación del DOA al usar Beamforming.

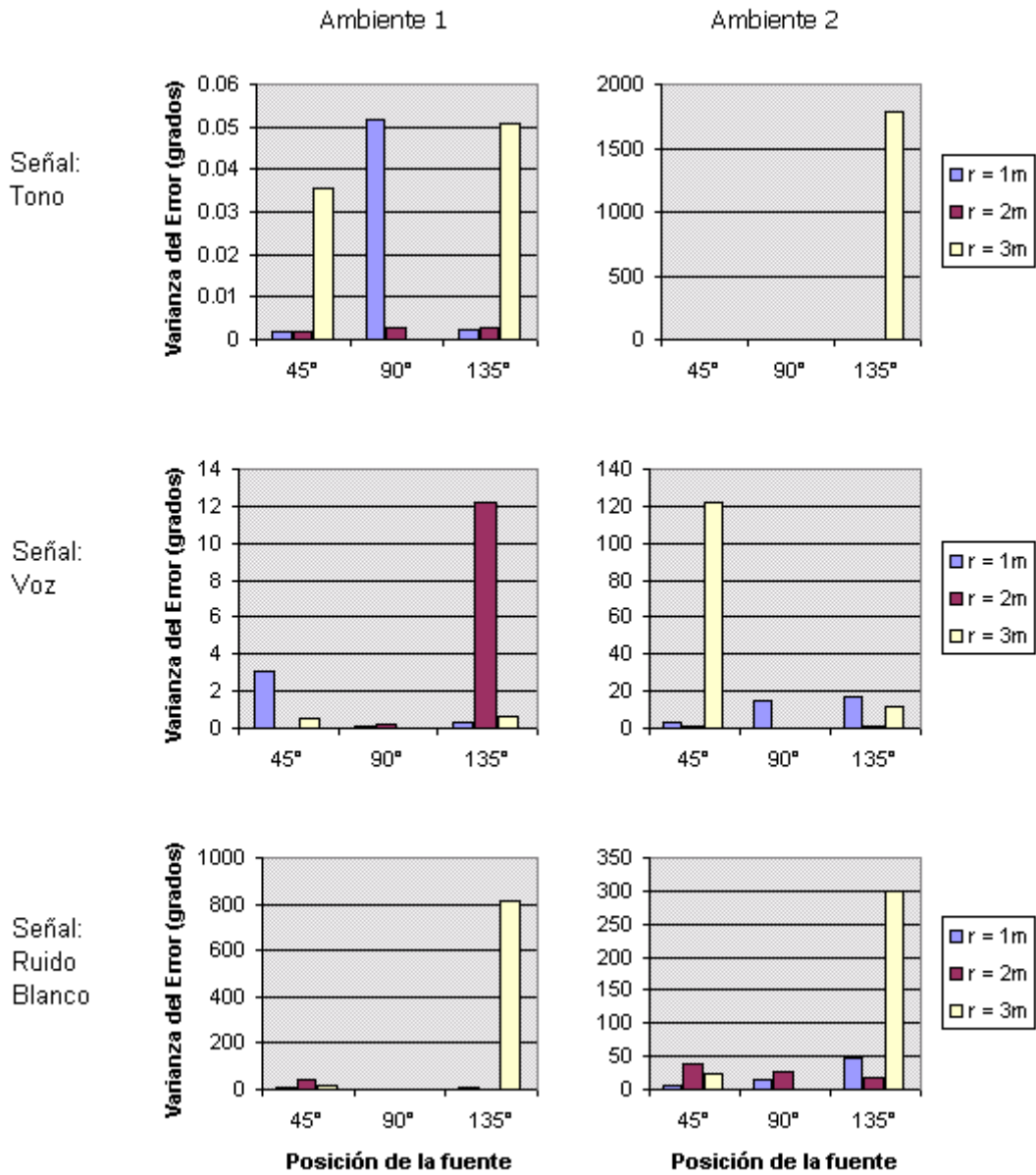


Figura 5.10 Varianza del Error de estimación del DOA al usar MUSIC.



**Conclusiones**

- La implementación del sistema de localización muestra que el modelo de datos y las técnicas de localización son válidas, aunque con ciertas limitaciones ante interferencias intrínsecas de la localización de fuentes de sonido. El rendimiento obtenido que ofrece este sistema es comparable con investigaciones previas en el área de procesamiento de señales con arreglos, siendo este trabajo el primer paso dado en implementar sistemas más complejos y robustos en la Pontificia Universidad Católica del Perú. Dentro de nuestro contexto tecnológico, los resultados obtenidos son adecuados.
- Se ha podido comprobar que el modelo de datos es funcional en cuanto a la descripción de la forma de propagación del sonido con un patrón esférico. Como se mostró en las simulaciones, es posible definir el campo cercano y el campo lejano haciendo referencia a la distancia entre la fuente y el arreglo de sensores. De esta manera, si la fuente de sonido se encuentra en el campo lejano, las ondas recibidas por el arreglo se podrán caracterizar como frentes de onda planos y se podrá determinar la dirección de arribo.
- El efecto de las perturbaciones que sufre el sonido en un ambiente con cierto nivel de reverberación es notable con ciertas señales. Al obtener los resultados de las simulaciones y de los experimentos, se puede afirmar la necesidad de algoritmos más complejos para conseguir una estimación con menos cantidad de error. Estos sistemas más robustos requerirían de conceptos como filtraje adaptivo u otros que permitan analizar la influencia del ambiente en el cual se trabaja de una manera más específica que la tratada aquí.
- Se comprobó la necesidad del uso de restricciones de la posición de la fuente para los experimentos con el fin de evitar ambigüedades. Estas restricciones pueden variar dependiendo de la aplicación deseada. Si bien es cierto que se utilizó un arreglo lineal de 4 micrófonos para localizar la fuente de sonido en un semiplano, existe la posibilidad de la ubicación tridimensional al usar un arreglo no lineal.

- El teorema de Nyquist del muestreo del tiempo tiene su contraparte en su caso espacial, lo cual resalta la importancia de realizar un muestreo adecuado para que los datos mantengan un significado físico.
- Los resultados obtenidos en las simulaciones y en los experimentos con una fuente de sonido real en MATLAB, indican una alta confiabilidad porque no se encontraron grandes diferencias entre ellas en todas las pruebas realizadas.
- El uso del promediado y del enventanamiento son conceptos muy usados en el procesamiento digital de señales y su aplicación aquí demuestra su utilidad para obtener mejores resultados en cualquiera de los métodos de estimación escogidos. Particularmente, aquí se usaron para una mejor estimación de la matriz de covarianza, de los coeficientes de Fourier; y también para hallar el error promedio y la varianza del error de estimación de una serie de 10 experimentos.
- La correlación cruzada que se utilizó para la estimación en el dominio del tiempo del retardo entre dos secuencias discretas es muy útil y práctica como se aprecia a través de este documento. Pero hay que tener cuidado al escoger punto de referencia para poder determinar la dirección de arribo correcta.
- El trabajo con modelos de datos, llevan a aproximaciones que consideran la situación real del problema y su debido tratamiento para una buena estimación. Estas aproximaciones incluyen la discriminación comentada anteriormente para la correlación y la segmentación de la señal en cuadros para una estimación de la matriz de covarianza estimada.
- Los resultados favorecen la viabilidad para desarrollar trabajos más complejos en el futuro, que incluyan mejora de recepción de señales, sistemas en tiempo real, localización de 2 ó más fuentes de sonido en el espacio ó generación de campos sonoros por nombrar algunos.



Recomendaciones

- Un arreglo de sensores ha permitido una menor complejidad del equipo necesario para resolver el problema propuesto, en comparación con un sistema en base a un solo micrófono. El equipo usado fue muy básico. Se contó con 4 micrófonos omni-direccionales, 4 cables balanceados, una consola de audio y un computador personal para almacenar las señales digitalizadas. Además, permitió una alta flexibilidad para el trabajo de estimación mediante diversos algoritmos, dejando abierta la posibilidad de próximas investigaciones sobre el tema. Se recomienda, trabajar con diferentes geometrías de arreglos; de esta manera se podrían hacer incluso estimaciones de fuentes de sonido en el espacio.
- Los resultados muestran la posibilidad de nuevas investigaciones como localización de fuentes en el espacio, Para hacer esto es recomendable utilizar un arreglo de sensores más grande (en términos de cantidad de sensores) y además una geometría distinta. Utilizando arreglos de sensores más grandes es posible realizar aplicaciones más complejas, tales como estudios de campos sonoros como los que se realizan en la industria automotriz y aeronáutica.
- Agrandando el número de sensores en un arreglo, es posible mejorar la directividad del mismo y por tanto se podría hacer estimaciones para más de una fuente. En este caso al contar con cuatro micrófonos en un arreglo lineal, lo adecuado fue estimar la dirección de arribo de una sola fuente.
- Es necesario un estudio de la reverberación, tema que se viene estudiando en la actualidad. De esta manera, se tendrá que hacer una adaptación adecuada para una mejor estimación de la dirección de arribo de una señal. Actualmente las técnicas que permiten ubicar fuentes de sonido u otro tipo de señal utilizan filtros adaptivos permitiendo así, la eliminación de interferencias.
- Probar algoritmos con arreglos no lineales, para no restringir la fuente a ciertas posiciones. De esta manera se podría hacer estimaciones en tres dimensiones evitando ambigüedades. Esto resalta la posibilidad del estudio de las ambigüedades dependiendo del arreglo de sensores que se use.

- Investigaciones en tiempo real para contemplar las consideraciones necesarias para en una aplicación real, tales como el orden adecuado de las rutinas a ejecutarse y su uso efectivo en el tiempo. Además, esto permitirá experimentar luego con algoritmos como filtraje adaptivo, algoritmo muy usado en estos casos, pues permite como se dijo anteriormente la eliminación de interferencias.
- Analizar algoritmos en otros dominios, como el dominio de los wavelets, con el fin de comparar resultados y así escoger la técnica más eficiente. Cabe mencionar el trabajo de tesis de Paúl Rodríguez [23], en el cual se utiliza un filtrado en este dominio con el fin de eliminar ruido. Pues como se pudo observar en este trabajo de tesis, el ruido afecta la estimación del ángulo de arribo.
- Una de las aplicaciones más evidente es el seguimiento de un locutor en una videoconferencia para optimizar el enfoque de la toma de su voz o su imagen. Obviamente, las recomendaciones dadas anteriormente podrán mejorar los resultados de esta aplicación.
- Considerando la mejora de SNR de señales, es evidente el potencial de la teoría de arreglos de sensores para diversas aplicaciones. Por ejemplo, en el caso de las comunicaciones satelitales, es posible el uso de arreglos de antenas para reemplazar una antena más grande. De esta manera, se puede lograr reducir los costos de diversos sistemas.

## Bibliografía

1. El sonido: Bases físicas para su aplicación en ecografía  
<http://medsalud.com.ar/principal.htm>
2. D. Davis y C. Davis, "Sound System Engineering", Howard W. Sams & Co., Inc. Segunda Edición. May 1997.
3. H. Kuttruff, "Room acoustics", E & FN Spon, Cuarta Edición, New York 2000.
4. Diccionario bilingüe de Audio Profesional / Translation Dictionary  
URL: [www.doctorproaudio.com](http://www.doctorproaudio.com)
5. H. Krim y M. Viberg, "Two Decades of Array Signal Processing Research", IEEE Signal Processing Magazine, pp 67-94. Julio 1996.
6. A. Hero, H. Messer, J. Goldberg, D. J. Thomson, M.G. Amin, G. Giannakis, A. Swami, J. K. Tugnait, A. Nehorai, A. L. Swindlehurst, J. F. Cardoso, Lang Tong, J. Krolik, "Highlights of Statistical Signal and Array Processing", IEEE Signal Processing Magazine, vol. 15, p. 21-64, Setiembre 1998.
7. D.H. Johnson y D.E. Dudgeon, "Array Signal Processing: Concepts and Techniques.", Prentice Hall, New Jersey 1993.
8. J.L. Spiesberger, "Hyperbolic location errors due to insufficient numbers of receivers", J. Acoust. Soc. Am., Vol-109:3076-3079, Junio 2001.
9. B.H. McGuyer, S.T. Lionetti, y J.D. Thompson, "Small microphone array sound localization techniques and Applications". Octubre 2001
10. H. Wang, y P. Chu "Voice Source Localization Automatic Camera Pointing System in Videoconferencing" In Proceedings of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing: Munich, 1997.
11. J. González Rodríguez, J.L. Sánchez Bote y J. Ortega, "Mejora de voz con arrays de micrófonos mediante descomposición en componentes de fase mínima y paso todo", Actas de TecniAcústica 99, Ávila – España 1999.

12. M. Omologo, M. Matassoni, P. Svaizer y D. Giuliani, "Microphone array based speech recognition with different talker-array positions.", Proceedings of ICASSP-97, pp. 227–230, 1997.
13. Q. Lin, E. Jan y J. Flanagan. "Microphone arrays and speaker identification.", IEEE Transactions on Speech and Audio Processing, vol 2(4):622-629, Octubre 1994.
14. J.E. Greenberg and P.M. Zurek. "Evaluation of an adaptive beamforming method for hearing aids.", J. Acous. Soc. Am. 91:1662-1676, Marzo 1992.
15. E. Smith, "Parameter Estimation in Array Signal Processing Via Nonlinear Semidefinite Programming With Application to Wireless Communications", The Telecommunications Review pp. 93-99, 2001.
16. J.S. Mosher, R.M. Leahy y P.S. Kewis, "Biomagnetic Localization from transient Quasi-Static Events." In Proc. ICASSP93, Vol 1, pp 91-94, Minneapolis, MN, 1993.
17. G.M. Barbieri, E.D. Di. Claudio, R. Parisi, "Sensor Arrays for improved instrumental landing systems.", Italia 2000.
18. J.G. Proakis y D.G. Manolakis, "Tratamiento Digital de Señales", tercera edición, Prentice Hall, España, 2000.
19. C.S. McInnes, "Procesamiento Digital de Señales", Apuntes de clases de postgrado del Ing. Andrés Flores, Pontificia Universidad Católica del Perú, Lima-Perú, 2002
20. C. S. Burrus, J. H. McClellan, A. V. Oppenheim, T. W. Parks, R.W. Schafer y H.W. Schuessler, "Tratamiento de la Señal Utilizando MATLAB v.4". Pearson Educación, España, 1998.
21. D.C. Lay, "Álgebra lineal y sus aplicaciones.", Pearson Educación, segunda edición, México 2001.
22. P. Athanasios, "Probability, Random Variables and Stochastic Processes". McGraw-Hill, Tercera Edición, NY 1991.

23. P. Rodríguez, "Estimación del tiempo de retardo de arribo y de la frecuencia instantánea de señales provenientes de un arreglo de fase (sonar) utilizando técnicas de prefiltrado de ruido en el dominio de los wavelets.", Tesis de licenciatura, Pontificia Universidad Católica Del Perú, Facultad de Ciencias e Ingeniería, Lima-Perú, 1997
24. A.G. Piersol, "Time Delay Estimation Using Phase Data.", IEEE Trans. ASSP, Vol ASSP-29:471-477, Junio 1981.
25. B.D. Van Veern y K.M. Buckley. "Beamforming: A Versatile Approach to Spatial Filtering". IEEE ASSP Magazine. pp 4-24. Abril 1988.
26. P. Stoica y A. Nehorai. "MUSIC, Maximum Likelihood, And Cramer-Rao Bound." IEEE Trans. ASSP, ASSP-37:720-741, Mayo 1989.
27. M. Torlak, Notes from Signal and Image Processing Seminar. EE381K Multidimensional Digital Signal Processing - Spatial Array Processing. Universidad de Texas en Austin. EEUU. Nov 1998.
28. B.L. Evans, "EE381K-9 Advanced Signal Processing.", Lectures and notes, Universidad de Texas en Austin. EEUU. 1997
29. J.C. Chen, "Wideband Source Localization For Sensor Networks: Theory And Practice" University Of California, Los Angeles, Febrero 2002.
30. F. Asano, H. Asoh y T. Matsui, "Sound Source Localization And Separation In Near Field.", IEEE Trans. Fundamentals, Vol-E83-A:2286-2294, Noviembre 2000.
31. S. Lehman, "Reverberation". 1996  
URL: <http://www.harmony-central.com/Effects/Articles/Reverb/#ref>
32. J. B. Allen y D. A. Berkley, "Image method for efficiently simulating small-room acoustics," J. Acoust. Soc. Am., vol. 65, pp. 943–950, Apr. 1979.