

Instituto Tecnológico de Costa Rica

Escuela de Ingeniería Electrónica



**Selección de un algoritmo de localización acústica para su uso en
aplicaciones de robótica industrial**

Informe de Proyecto de Graduación para optar por el título de
Ingeniero en Electrónica con el grado académico de Licenciatura

Roberto CERDAS ROBLES

Diciembre 2012

Declaración de Autenticidad

Yo, Roberto CERDAS ROBLES, declaro que el presente documento ha sido realizado por mi persona, utilizando y aplicando literatura referente al tema, e introduciendo conocimientos propios. En los casos en que he utilizado bibliografía, he procedido a indicar las fuentes mediante las respectivas citas bibliográficas, y, por tanto, asumo la responsabilidad por el contenido de este proyecto.

Firma: Roberto Cerdas Robles

Fecha: 13/12/2012

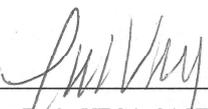
Instituto Tecnológico de Costa Rica
Escuela de Ingeniería Electrónica
Proyecto de Graduación
Tribunal Evaluador

Proyecto de Graduación defendido ante el presente Tribunal Evaluador como requisito para optar por el título de Ingeniería en Electrónica con el grado académico de Licenciatura, del Instituto Tecnológico de Costa Rica.

Estudiante: Roberto CERDAS ROBLES

Nombre del Proyecto: *Selección de un algoritmo de localización acústica para su uso en aplicaciones de robótica industrial*

Miembros del Tribunal


Ing. Paola VEGA CASTILLO
Profesora Lectora


Ing. Alfonso CHACÓN RODRÍGUEZ
Profesor Asesor



Los miembros de este Tribunal dan fe de que el presente trabajo de graduación ha sido aprobado y cumple con las normas establecidas por la Escuela de Ingeniería Electrónica.

Cartago, Diciembre 2012

Resumen

La automatización de vehículos requiere el uso de una red de localización redundante que permita determinar la posición de los objetos en todo momento. Es posible realizar esta localización mediante el uso del sonido emitido por el objeto. La estimación mediante el cálculo del *Delay Of Arrival* (DOA) es un método comúnmente utilizado para determinar el ángulo azimutal de la fuente con respecto a un arreglo de sensores. Tres algoritmos fueron examinados para un arreglo lineal de micrófonos equidistantes entre sí, y evaluados de acuerdo a su precisión, consumo de recursos y eficiencia. Se determinó que el uso del coeficiente de correlación cruzada multicanal (MCCC) es un método eficiente y preciso para la estimación del DOA.

Palabras clave: localización, acústica, TDOA, ruido, correlación, DSP

INSTITUTO TECNOLÓGICO DE COSTA RICA

Abstract

Escuela de Ingeniería en Electrónica

Licenciatura en Ingeniería Electrónica

Selección de un algoritmo de localización acústica para su uso en aplicaciones de robótica industrial

by Roberto CERDAS ROBLES

Automation of vehicles requires a redundant network of sensors that can determine the position of the objects at all times. Localization of moving objects may be performed through the use of acoustic cues. Delay of Arrival (DOA) estimation is a popular method for determining the azimuth angle of a source with respect to a sensor array. Three separate algorithms were examined for a linear array of equidistant microphones, and evaluated based on performance, accuracy and computational resources consumed. The determination of the multichannel cross correlation coefficient for the array was found to be an efficient and accurate method for DOA estimation.

Key Words: localization, acoustic, TDOA, noise, correlation, DSP

Agradecimiento

Deseo agradecer a mi familia por todo su apoyo a lo largo de los últimos años. El aliento y el cariño que me han proporcionado ha sido invaluable, y es gracias a su amor que me es posible escribir este documento. Deseo agradecer además a los profesores Alfonso Chacón, Paola Vega y Marta Vílchez, quienes han sido, y continúan siendo, una inspiración, tanto como profesionales y como personas. Finalmente, quisiera agradecer al sensei Alberto París y la sensei Patrizia Gallo, por enseñarme la importancia de la ambición para alcanzar los sueños y metas que nos proponemos en nuestras vidas.

Tabla de contenidos

Abstract	ii
Agradecimiento	iii
Índice de Figuras	vi
Índice de Tablas	vii
1 Introducción	1
1.1 Definición del problema	1
1.2 Solución propuesta	4
2 Meta y objetivos	5
2.1 Meta	5
2.2 Objetivo general	5
2.3 Objetivos específicos	5
3 Marco teórico	6
3.1 Descripción del sistema y modelo	6
3.2 Cálculo de TDOA	9
3.2.1 Correlación Cruzada Generalizada (GCC)	11
3.2.2 Forward Spatial Linear Prediction(FSLP)	12
3.2.3 Multichannel Cross Correlation Coefficient(MCCC)	14
4 Procedimiento metodológico	17
4.1 Procedimiento realizado	17
5 Metodología de implementación	18
5.1 Reconocimiento y definición del problema	18
5.2 Obtención y acondicionamiento de la señal	18
5.3 Correlación cruzada	19
5.4 Forward Spatial Linear Prediction	23
5.5 Multichannel Cross Correlation Coefficient	25
6 Análisis de resultados	27
6.1 Resultados	27

6.1.1	Correlación Cruzada	28
6.1.2	MCCC	30
6.1.3	FSLP	31
7	Conclusiones y Recomendaciones	41
7.1	Conclusiones	41
7.2	Recomendaciones	42
	 Bibliografía	 43

Índice de figuras

3.1	Triangulación	7
3.2	Reverberación	8
5.1	Convolución	20
5.2	Wrap-around	20
5.3	Zero Padding	21
5.4	Algoritmo de Correlación Cruzada	22
5.5	Algoritmo de FSLP	24
5.6	Algoritmo de MCCC	26
6.1	Señal de Control	27
6.2	Señal de Control 2	28
6.3	Señal con Ruido Agregado, $SNR = 9dB$	29
6.4	Correlación Cruzada	30
6.5	Correlación Cruzada, $SNR = 9dB$	31
6.6	Correlación Cruzada, $SNR = 6dB$	32
6.7	Correlación Cruzada, $SNR = 3dB$	33
6.8	MCCC	34
6.9	MCCC, $SNR = 6dB$	35
6.10	MCCC, $SNR = 6dB$	36
6.11	FSLP	37
6.12	FSLP, $SNR = 6dB$	38
6.13	FSLP, $SNR = 6dB$, Ampliación	39
6.14	MCCC, $SNR = 6dB$	40

Índice de tablas

6.1 Comparación de algoritmos	33
---	----

*Dedicado a mis padres, sin los cuales el presente documento no
sería posible.*

1

Introducción

1.1 Definición del problema

En la historia reciente, los procesos industriales han entrado a una era de automatización, en la cual se busca simplificar y reducir la necesidad de personal humano que realice tareas repetitivas, reemplazando a los mismos por máquinas y procesos de control automático, mejorando así el rendimiento de la producción más allá de lo alcanzable mediante capacitaciones de personal. Esta ha sido la tendencia durante los últimos 50 años [Hitomi, 1994]. Sin embargo, con el advenimiento de la era de la información, la automatización se ha vuelto cada vez más intrusiva en los campos de trabajo [Parasuraman and Riley, 1997], llevando a un nivel de interacción entre maquinaria y ser humano que no había sido visto anteriormente. Es común que los operarios interactúen con sistemas completamente automatizados en su trabajo diario, y, por tanto, los mismos deben garantizar la seguridad de los anteriores en todo momento para evitar costosos accidentes [Inagaki, 2006]. Dado que los entornos de trabajo actuales con maquinaria automatizada pesada representan una amenaza para la salud e integridad física de los trabajadores, debe contarse con monitorización electrónica precisa; a modo de ejemplo, en la industria minera, un 30% de los accidentes fatales se encuentran asociados a maquinaria pesada [Julián et al.]. Sin embargo, no es posible aún automatizar por completo todos los procesos y remover al ser humano del lugar de riesgo. Los métodos tradicionales de capacitación de empleados siempre son susceptibles a la introducción de error humano, y la creciente automatización de sistemas potencialmente peligrosos crea la necesidad de nuevos métodos de prevención de riesgos [Piggin, 2006]. Implementar protocolos de seguridad es costoso para la empresa, reduciendo el potencial beneficio del aumento en la eficiencia de producción proporcionado por la maquinaria [Boehm-Davis et al., 1983]. La localización precisa de los mismos mediante monitorización es

prerrequisito para la realización de tareas complejas automatizadas, tales como minería, mapeo, mediciones, rastreo y vigilancia [Dieudonne et al., 2010].

Una vez que se establece la necesidad de monitorización, debe elegirse el tipo de monitorización a utilizar. Es deseable utilizar más de un tipo de sensor, ya que las condiciones en el entorno de la maquinaria pueden bloquear la señal a medir, provocando mediciones erróneas o no válidas. [Julián et al.]. El rastreo de objetos en movimiento mediante GPS, en particular, es uno que ha sido utilizado a menudo con este fin en campos abiertos [Bell et al., 1996], pero que sufre de múltiples limitaciones en la práctica debido a la susceptibilidad a las condiciones de la tropósfera, las cuales podrían interrumpir la conexión con el satélite y, por tanto, ocasionar la pérdida del rastreo de la maquinaria durante un período de tiempo inaceptable [Agrawal and Konolige, 2006]. Como consecuencia, se concluyó que debe adquirirse información de múltiples sensores para ser procesada en conjunto, lo cual provee una estimación acertada de la posición de la maquinaria. Sin embargo, la cantidad de información a transmitir entre los vehículos es, al contar con redes redundantes de sensores, necesariamente grande, y, por tanto, se requieren recursos computacionales significativos en cada nodo para procesarla. Para resolver el problema, se utilizará tecnología de circuitos integrados en 3D, con múltiples obleas apiladas para cumplir distintas funciones [Julián et al.]. Esto permite un ancho de banda elevado entre las distintas obleas, permitiendo integrar circuitos lógicos y memorias en un solo chip con baja latencia [Garg and Marculescu, 2011]. Este tipo de circuitos han sido utilizados anteriormente para procesar grandes cantidades de información, por ejemplo, en la separación, purificación y evaluación de partículas biológicas a escala nanométrica [Dickerson et al., 2009].

Como parte de la red de sensores redundantes, se propuso el uso de sensores de localización acústica [Julián et al.], los cuales son capaces de determinar el ángulo o distancia con respecto a otro objeto con un bajo consumo de potencia [Chacon-Rodriguez et al., 2009]. El sistema auditivo de los mamíferos es capaz no sólo de percibir una amplia gama de sonidos, sino también de distinguir entre los mismos y responder instintivamente a aquellos de interés para su supervivencia o beneficio [Heffner et al., 2008]. El ser humano evidencia esta característica en su vida cotidiana, particularmente mediante el uso del lenguaje, respondiendo a frases u oraciones aún al estar inmerso en un entorno ruidoso como lo es el de una ciudad promedio, al tiempo que es capaz de percibir la posición desde la cual se origina la voz [Mesgarani et al., 2009, Evangelopoulos and Zlatintsi, 2009]. Esta habilidad de los mamíferos de distinción entre estímulos sensoriales separados permite diferenciar a una fuente de sonido determinada de su entorno, aún cuando el mismo pueda poseer una gran variedad de fuentes acústicas, y determinar su posición [Kayser et al., 2005]. Al emular esta habilidad, es posible crear una red de sensores que reaccione

a un sonido específico y ubique al mismo dentro de un sistema de coordenadas [Shujing, 2010, Johansson].

El proyecto propuesto consta de la selección de un algoritmo de procesamiento de datos que realice esta función, para ser posteriormente implementado como parte del sistema descrito anteriormente. Se han utilizado algoritmos anteriormente que permiten alcanzar este objetivo por distintos métodos [Benesty et al., 2008]. Es necesario evaluar los mismos y elegir aquel que se adapte al propósito del proyecto; es decir, aquel cuya implementación en hardware sea factible y cuente con un grado aceptable de especificidad y sensibilidad. El mismo debe ser, necesariamente, capaz de filtrar otras fuentes sonoras para distinguir el objeto de interés del ruido base del entorno.

El proyecto a realizar se encuentra supeditado al proyecto de investigación 3D Gigascale Integrated Circuits for Nonlinear Computation, Filter and Fusion with Applications in Industrial Field Robotics desarrollado en conjunto por la Universidad Nacional del Sur en Argentina, John Hopkins University en los Estados Unidos de América (EE. UU.), University of Maryland en ese mismo país, University of Sidney en Australia y el Instituto Tecnológico de Costa Rica. La meta establecida para el proyecto es la creación de un SoC (System on Chip) complejo utilizando tecnología tridimensional para aplicaciones en la robótica de campo industrial, específicamente en la localización de recursos mecánicos automatizados y personal. Para el desarrollo del mismo, se preparará un escenario experimental en una cantera australiana, la cual cuenta con camiones pesados y equipo auxiliar equipados con una plataforma de comunicaciones. Esta última cuenta con dos redes que operan a distinta frecuencia (2,4 GHz y 433 MHz) para reducir la posibilidad de una falla simultánea en ambas. El sistema cuenta con entradas y salidas analógicas y digitales, conexión Ethernet y puertos seriales. Este sistema almacena en un servidor la posición y orientación del vehículo, así como la de aquellos que entraron en contacto con él. Una vez se pruebe el sistema a desarrollar en el laboratorio, se procederá a realizar la interfaz entre el mismo y las redes de comunicación existentes, permitiendo así el almacenamiento en el servidor de los datos experimentales, para así comprobar la validez de la solución [Julián et al.].

El problema se define como: elegir un algoritmo de localización acústica para estimar la posición de vehículos en movimiento.

1.2 Solución propuesta

El algoritmo elegido debe ser capaz de rechazar ruido blanco gaussiano (común en micrófonos). Además, debe ser computacionalmente eficiente, de forma que su implementación final no posea una demanda alta de recursos. Se propone el uso de un arreglo lineal de micrófonos con una distancia de separación constante entre los mismos de cuatro metros. La localización de los vehículos se puede determinar a partir del tiempo de retardo entre los micrófonos, comúnmente denominado *time delay of arrival* o *time difference of arrival* (TDOA) [Jia et al., 2009]. Si se asume que el ruido en las señales capturadas por los micrófonos no se encuentra correlacionado con la señal que se desea localizar, es posible utilizar el vector de correlación cruzada entre las señales para estimar este retardo [Benesty et al., 2008]. El algoritmo propuesto permite combinar la información obtenida por el vector de correlación entre distintos pares de micrófonos para así lograr obtener una estimación del TDOA correcto, utilizando la información redundante presente en dichos vectores para mejorar la precisión del mismo y su resistencia a la introducción de ruido.

2

Meta y objetivos

2.1 Meta

La meta del proyecto es la elaboración de un SoC tridimensional para la localización de maquinaria pesada y otros recursos en entornos laborales donde interactúan sistemas automatizados y seres humanos.

2.2 Objetivo general

Seleccionar un algoritmo implementable en hardware para la localización acústica de vehículos en movimiento.

2.3 Objetivos específicos

- Trasladar los algoritmos a estudiar a un lenguaje de programación común.
- Diseñar un banco de pruebas que permita obtener los parámetros estadísticos necesarios para el estudio.
- Comparar los resultados del estudio.

3

Marco teórico

3.1 Descripción del sistema y modelo

Cuando un sonido es emitido en un entorno ideal, dos micrófonos distintos son capaces de capturar la misma señal, con una diferencia de tiempo entre ellas directamente relacionada a la distancia de cada micrófono al objeto [Birchfield and Gangishetty, 2005]. El tiempo que toma la señal de sonido en viajar desde la fuente al receptor se conoce en la literatura como Time-Delay (TD) [Ferguson et al., 2002]. La mayoría de los circuitos integrados (IC) desarrollados para este propósito realizan una medición de este tiempo de retardo a lo largo de dos o más sensores, utilizando la diferencia de tiempo o de fase entre las señales recibidas para determinar la posición de la fuente mediante el uso de triangulación. Esta técnica recibe el nombre de *Interaural Time Delay* (ITD) [Chacon-Rodriguez et al., 2010] o *Time Difference of Arrival* (TDOA) [Jia et al., 2009]. La figura 3.1 ilustra el concepto de triangulación bidimensional de la fuente mediante TDOA, utilizando dos micrófonos.

Es importante destacar que este modelo aplica para el caso de campos lejanos únicamente, cuando la fuente de sonido se encuentra lo suficientemente alejada para que el frente de onda sea plano; el modelo no es válido en el caso de una fuente cercana al arreglo [Li and Caldwell, 1999]. Sea τ_{ij} el TDOA entre un par de micrófonos i y j , con $i, j = 1, 2, 3, \dots, L$, siendo L el total de micrófonos en el arreglo, separados una distancia d entre sí. Sea θ_n el ángulo azimutal entre la fuente $s(k)$ y el sensor n ($n = 1, 2, \dots, L$) que describe la dirección de arribo de la señal. Nótese que para el caso en estudio, si el frente de onda es plano, el ángulo es igual para todos los sensores equidistantes entre sí. Entonces, el retardo entre dos sensores puede ser descrito como:

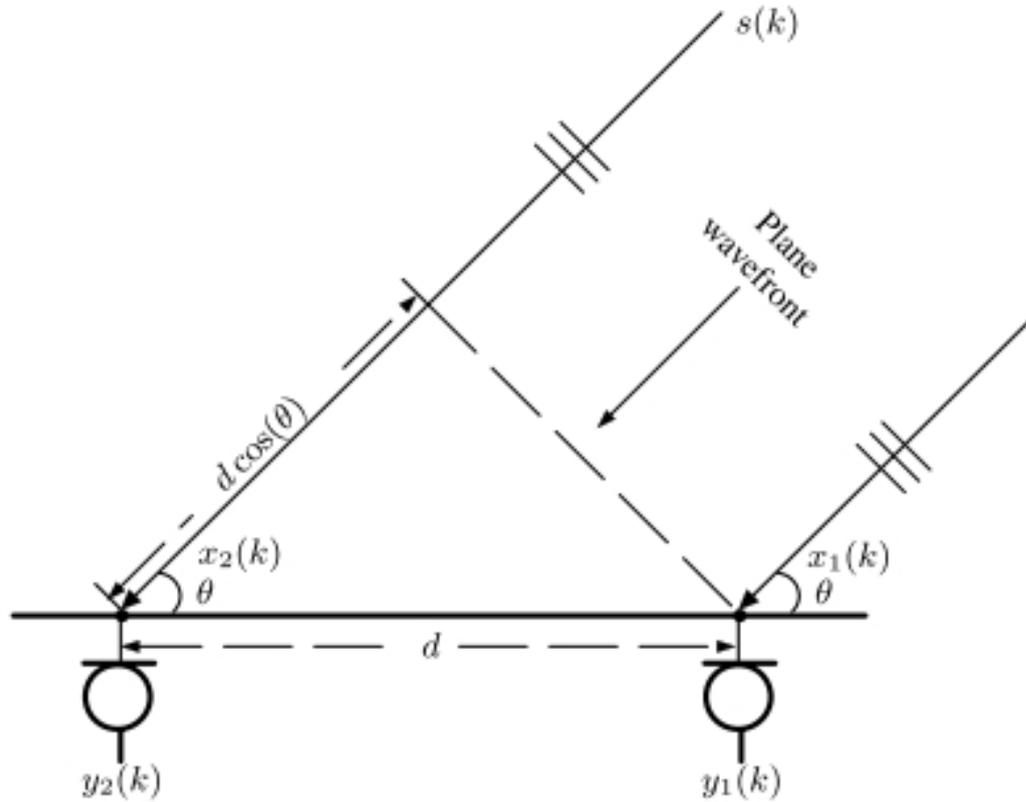


FIGURA 3.1: Triangulación de una fuente de sonido mediante TDOA. [Benesty et al., 2008].

$$\tau_{ij} = \frac{d \cos(\theta)}{c} \quad (3.1)$$

con c la velocidad del sonido en el aire, $343m/s$ a temperatura ambiente. Por tanto, si se conoce el valor de τ , es posible calcular el ángulo θ del cual proviene la señal, ya que d siempre puede ser medido una vez que el arreglo se encuentra colocado. [Shujing, 2010, Ferguson et al., 2002]. El problema se puede redefinir, entonces, como la determinación de este valor τ .

Los entornos donde los sensores son instalados usualmente no son entornos ideales, sin embargo, y por tanto presentan una dificultad adicional; la necesidad de filtrar las fuentes de sonido indeseadas y el ruido del entorno para separar los mismos de la señal de interés [DiBiase, 2000]. Otro punto a considerar es el hecho de que la ruta que sigue el sonido no es necesariamente lineal debido a la presencia de obstáculos [Tirumala, 2004]. Adicionalmente, en entornos cerrados o urbanos, los efectos de reverberación producen copias adicionales de la señal de interés, provenientes de fuentes distintas a la búsqueda originalmente [Ramamurthy, 2007]. La figura 3.2 ilustra el efecto de la reverberación.

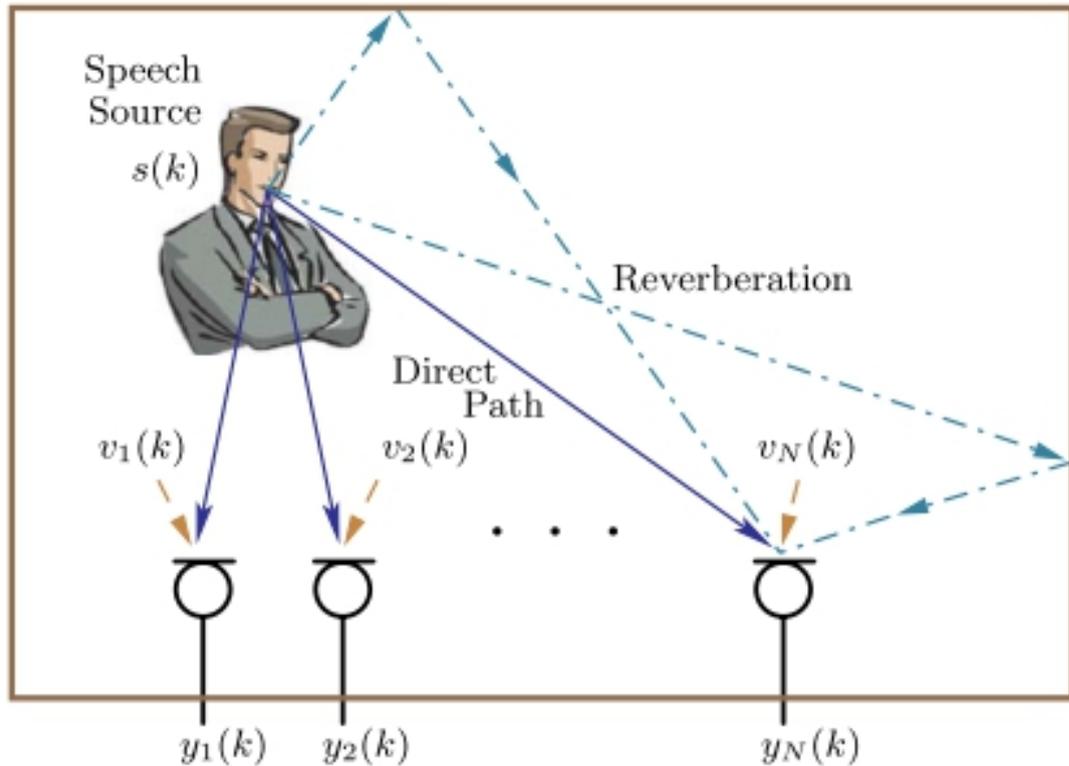


FIGURA 3.2: Reverberación de una señal acústica [Benesty et al., 2008].

Para efectos de este documento, se examinará el caso de una sola fuente sonora $s(k)$. Esta fuente cumple con las siguientes restricciones:

- La fuente emite ondas esféricas que se propagan uniformemente en todas direcciones.
- El medio es homogéneo, es decir, la velocidad del sonido c no cambia de un punto a otro. Si bien c depende de la temperatura, se asume que esta es constante a lo largo del sistema bajo estudio.

Se define ahora $y_n(k)$, con $n = 1, 2, \dots, L$ como el conjunto de las señales recibidas por los micrófonos [Benesty et al., 2008]:

$$y_n(k) = \alpha_n s(k - t - \tau_{n1}) + v_n(k) \quad (3.2)$$

Con α_n el factor de atenuación experimentado por la señal debido a su propagación y $v_n(k)$ el ruido que se suma a la señal en el micrófono n , el cual se asume que no se encuentra correlacionado con la señal de origen $s(k)$ o el ruido presente en los demás

sensores. t corresponde al tiempo que tarda el frente de onda en alcanzar al primer sensor luego de ser emitido por la fuente. τ_{n1} representa el retardo entre la señal recibida en el primer micrófono y el micrófono n , relacionado con el retardo τ por:

$$\tau_{n1} = (n - 1)\tau \quad (3.3)$$

$$= F_n(\tau) \quad (3.4)$$

para un arreglo lineal de micrófonos equidistantes entre sí. La ecuación (3.3) depende de la geometría del arreglo; sin embargo, como la geometría siempre es conocida, el único valor que se requiere encontrar es τ . Por tanto, no es de interés el retardo absoluto de una señal con respecto a la fuente, sino únicamente el retardo τ entre los sensores [Bahadirlar and Gulcur, 2001]. La ecuación (3.2) puede simplificarse como:

$$y_n(k) = \alpha_n s(k - F_n(\tau)) + v_n(k) \quad (3.5)$$

3.2 Cálculo de TDOA

Si bien existen distintos métodos para determinar el TDOA entre dos señales, el más sencillo y utilizado consiste en el cálculo de la función de correlación cruzada entre ellas [Rhudy et al., 2009]. La correlación cruzada es una medida de la similitud entre dos señales, en función de un retardo aplicado a una de ellas. Se define como:

$$(f \star g)(p) = \int_{-\infty}^{\infty} f^*(\tau)g(p + \tau)d\tau \quad (3.6)$$

$$= r_{fg}^{CC}(p) \quad (3.7)$$

Asúmase por ahora la existencia de únicamente dos micrófonos y sus señales correspondientes, $y_1(k)$ y $y_2(k)$. Al sustituir (3.5) dentro de (3.6), se obtiene [Benesty et al., 2008]:

$$r_{fg}^{CC}(p) = r_{y_1 y_2}^{CC}(p) \quad (3.8)$$

$$r_{y_1 y_2}^{CC}(p) = \alpha_1 \alpha_2 r_{ss}^{CC}(p - \tau) + \alpha_1 r_{sv_1}^{CC}(p) + \alpha_2 r_{sv_2}^{CC}(p - \tau) + r_{v_1 v_2}^{CC}(p) \quad (3.9)$$

Dado que las señales de ruido $v_1(k), v_2(k)$ no se encuentran correlacionadas con la señal de origen $s(k)$, el valor máximo de esta ecuación se alcanza cuando $p = \tau$. Entonces, el TDOA puede encontrarse como:

$$\tau = \max(\text{abs}[r_{y_1 y_2}^{\text{CC}}(p)]) \quad (3.10)$$

Este método es sencillo, pero sufre de dos defectos críticos [Valente and Wellekens, 2005]:

- No toma en consideración la autocorrelación de la señal.
- Asume la existencia de únicamente dos micrófonos. Si bien es posible combinar la información de ángulo azimutal obtenida a partir de cada par de micrófonos, realizar la estimación directamente no es eficiente, y es susceptible a variaciones en la captura de la señal en distintos micrófonos debido al ruido.

La implementación discreta de este algoritmo se puede escribir como [Do, 2009]:

$$r_{y_1 y_2}^{\text{CC}}(p) = \begin{cases} \frac{1}{K} \sum_{i=0}^{K-p-1} y_1(k+i)y_2(k+i+p) & \text{si } p \geq 0, \\ r_{y_2 y_1}^{\text{CC}}(-p) & \text{si } p < 0. \end{cases} \quad (3.11)$$

Resolver (3.11) para formar el vector de correlación en función del retardo requiere realizar p sumas, cada una de las cuales posee, en promedio, $K/2$ términos. Adicionalmente, cada uno de estos términos es un producto. Por tanto, se requiere de $O(Q^2)$ operaciones, con Q el número de muestras de la señal. Sin embargo, es posible utilizar las propiedades de la correlación y la transformada de Fourier para simplificar el cálculo [Stanacevic and Cauwenberghs, 2005]. La correlación entre dos funciones puede escribirse en función de la convolución de las mismas [Ziegel et al., 1987, Katznelson, 2004]:

$$(f \star g) = f^*(-t) * g \quad (3.12)$$

Donde $*$ denota el conjugado complejo de la función f , y $*$ es el operador de convolución. Aplicando la propiedad de convolución para la transformada de Fourier y la ecuación (3.12) se obtiene:

$$\mathcal{F}\{f \star g\} = \mathcal{F}^*\{f\}\mathcal{F}\{g\} \quad (3.13)$$

Y, por tanto, es posible simplificar (3.11). Sea $Y_n(f)$ la transformada de Fourier de la señal $y_n(k)$, entonces:

$$r_{y_1 y_2}^{\text{CC}}(p) = \mathcal{F}^{-1}\{Y_1^*(f)Y_2(f)\} \quad (3.14)$$

La transformada de Fourier y su inversa pueden ser calculadas eficientemente mediante el uso de la transformada rápida de Fourier. Esta permite reducir la cantidad de operaciones necesaria para la ejecución del algoritmo de $O(Q^2)$ a $O(Q \log Q)$ [Welch, 1967].

3.2.1 Correlación Cruzada Generalizada (GCC)

El método de correlación cruzada generalizada es una extensión de la correlación cruzada sencilla [Carter and Knapp, 1976], y es un método eficiente y popular para el cálculo de TDOA en un par de micrófonos [Shujing, 2010, Johansson, Jia et al., 2009, Do, 2009, Tang and Chang, 2007, Dmochowski et al., 2007a]. Este algoritmo aplica una función de peso $\vartheta(f)$ a la señal recibida por los sensores para reducir la influencia del ruido externo y el valor de la autocorrelación de la señal en el resultado.

$$r_{y_1 y_2}^{\text{GCC}}(p) = \mathcal{F}^{-1}\{\psi_{y_1 y_2}(f)\} \quad (3.15)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \vartheta(f) Y_1^*(f) Y_2(f) e^{j2\pi fp} df \quad (3.16)$$

Cuando la función de peso $\vartheta(f) = 1$, el resultado es el método de correlación cruzada sencilla. Típicamente, se opta por utilizar una función de peso igual al inverso de la magnitud del espectro en frecuencia de las funciones, o:

$$\vartheta(f) = \frac{1}{|Y_1^*(f)Y_2(f)|} \quad (3.17)$$

A esta variante del algoritmo de correlación cruzada generalizada se le denomina *GCC-PHAT* (*Generalized Cross Correlation-Phase Transform*) [DiBiase, 2000, Do, 2009]. La elección de $\vartheta(f)$ se basa en el hecho de que la información del TDOA buscado se encuentra en la fase del espectro de frecuencia, y no en su magnitud. Por tanto, se descarta la amplitud y se retiene únicamente la fase. Al sustituir la función de peso $\vartheta(f)$ presentada en (3.17) dentro de (3.15), se obtiene:

$$r_{y_1 y_2}^{\text{GCC-PHAT}}(p) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{\{Y_1^*(f)Y_2(f)\}} Y_1^*(f)Y_2(f)e^{j2\pi fp} df \quad (3.18)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-j2\pi f\tau} e^{j2\pi fp} df \quad (3.19)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j2\pi f(p-\tau)} df \quad (3.20)$$

$$r_{y_1 y_2}^{\text{GCC-PHAT}}(p) = \begin{cases} \infty & \text{si } p = \tau, \\ 0 & \text{si } p \neq \tau. \end{cases} \quad (3.21)$$

Esta variante es particularmente robusta a la introducción de ruido [Silverman et al., 2005], aunque aún no resuelve el problema presentado por entornos con reverberación alta, y considera únicamente un par de micrófonos a la vez. [Ramamurthy, 2007, Zarifeh et al., 2007]

3.2.2 Forward Spatial Linear Prediction(FSLP)

El efecto de la reverberación en el cálculo de TDOA se puede reducir utilizando un arreglo con múltiples micrófonos. Los algoritmos presentados en la sección anterior permiten encontrar la correlación entre la señal capturada en un par de micrófonos cualquiera, pero no contribuyen a combinar esta información para extraer la posición de la fuente. Por ejemplo, sea $L = 3$. Considérese ahora los pares de micrófonos y sus respectivos retardos, los cuales pueden ser calculados mediante los algoritmos de la sección anterior: τ_{12} , τ_{23} , τ_{13} . Para una geometría de arreglo conocida, es posible obtener un ángulo azimutal estimado θ a partir de cada uno de estos pares. Sin embargo, este método no toma en cuenta la siguiente relación:

$$\tau_{13} = \tau_{12} + \tau_{23} \quad (3.22)$$

La ecuación (3.22) muestra que existe información redundante en los vectores de correlación que se obtienen a partir de un arreglo de micrófonos con $L \geq 3$. Esta información puede utilizarse para reducir el efecto de la reverberación en la predicción del sistema [Dmochowski et al., 2007b,c]. Se busca un algoritmo de fusión que permita combinar los datos obtenidos para generar una predicción más útil.

Se introduce a continuación el algoritmo *Forward Spatial Linear Prediction*(FSLP), el cual se basa en minimizar el error entre la señal del primer micrófono y la señal capturada por los micrófonos restantes. Para ello, considérese primero la siguiente relación:

$$y_n[k + F_n(\tau)] = \alpha_n s(k - t) = y_1(k) \quad (3.23)$$

por tanto, es posible alinear la señal del micrófono n con la señal del micrófono 1 realizando un adelanto de la señal; el valor de este adelanto se encuentra determinado por la ecuación (3.3) [Tang and Chang, 2007, Benesty, 2003]. Los valores adelantados de la señal n pueden ser predichos a partir de la señal del micrófono 1, razón por la cual el algoritmo recibe su nombre. Se define primero la matriz de correlación espacial:

$$\mathbf{R}_a(p) = \begin{bmatrix} \sigma_{y_1}^2 & r_{a, y_1 y_2}(p) & \cdots & r_{a, y_1 y_L}(p) \\ r_{a, y_2 y_1}(p) & \sigma_{y_2}^2 & \cdots & r_{a, y_2 y_L}(p) \\ \vdots & \vdots & \ddots & \vdots \\ r_{a, y_L y_1} & r_{a, y_L y_2}(p) & \cdots & \sigma_{y_L}^2(p) \end{bmatrix} \quad (3.24)$$

Cuyas posiciones i, j corresponden a la correlación cruzada entre las señales $y_i(k), y_j(k)$. Ahora, se define el error como:

$$e_1(k, p) = y_1(k) - \mathbf{y}_{a,2:L}^T(k, p) \mathbf{a}_{2:L}(p) \quad (3.25)$$

Con:

$$\mathbf{y}_{a,2:L}(k, p) = \begin{bmatrix} y_2[k + F_2(\tau)] \\ y_3[k + F_3(\tau)] \\ \vdots \\ y_L[k + F_L(\tau)] \end{bmatrix} \quad (3.26)$$

Y $\mathbf{a}_{2:L}(p)$ los coeficientes de predicción lineal espacial, definidos como:

$$\mathbf{a}_{a,2:L}(k, p) = \begin{bmatrix} a_2(p) \\ a_3(p) \\ \vdots \\ a_L(p) \end{bmatrix} \quad (3.27)$$

Si el error es mínimo, estos coeficientes cumplen con [Benesty, 2003]:

$$\mathbf{R}_{a,2:L}(p)\mathbf{a}_{2:L}(p) = \mathbf{r}_{a,2:L}(p) \quad (3.28)$$

Donde $\mathbf{r}_{a,2:L}(p)$ se define como:

$$\mathbf{r}_{a,2:L}(p) = \begin{bmatrix} r_{a, y_1 y_2}(p) \\ r_{a, y_1 y_3}(p) \\ \vdots \\ r_{a, y_1 y_L}(p) \end{bmatrix} \quad (3.29)$$

La ecuación (3.29) puede resolverse como:

$$\mathbf{a}_{2:L}(p) = \mathbf{R}_{a,2:L}^{-1}(p)\mathbf{r}_{a,2:L}(p) \quad (3.30)$$

Sustituyendo (3.30) en (3.25) se obtiene el error mínimo:

$$e_{1, \min}(k, p) = y_1(k) - \mathbf{y}_{a,2:L}^T(k, p)\mathbf{R}_{a,2:L}^{-1}(p)\mathbf{r}_{a,2:L}(p) \quad (3.31)$$

A continuación, se busca el error medio cuadrático [Ziegel et al., 1987]:

$$J_{1, \min}(p) = E[e_{1, \min}^2(k, p)] = \sigma_{y_1}^2 - \mathbf{r}_{a,2:L}^T(p)\mathbf{R}_{a,2:L}^{-1}(p)\mathbf{r}_{a,2:L}(p) \quad (3.32)$$

el cual depende únicamente de p . Cuando el error es mínimo, la correlación entre los dos primeros micrófonos es máxima, es decir:

$$\tau = \min(\text{abs}[J_{1, \min}(p)]) \quad (3.33)$$

3.2.3 Multichannel Cross Correlation Coefficient(MCCC)

El coeficiente de correlación cruzada multicanal (MCCC) es otra técnica de fusión de los vectores de correlación obtenidos entre pares de micrófonos. En la práctica, produce mejores resultados que el algoritmo FSLP presentado en la sección anterior [Shujing, 2010, Benesty et al., 2008, Benesty, 2003]. Se define el vector completo de señales:

$$\mathbf{y}_a(k, p) = \begin{bmatrix} y_1(k) \\ y_2[k + F_2(\tau)] \\ y_3[k + F_3(\tau)] \\ \vdots \\ y_L[k + F_L(\tau)] \end{bmatrix} \quad (3.34)$$

a partir del cual se calcula la matriz de correlación espacial definida en (3.24):

$$\mathbf{R}_a(p) = \begin{bmatrix} \sigma_{y_1}^2 & r_{a, y_1 y_2}(p) & \cdots & r_{a, y_1 y_L}(p) \\ r_{a, y_2 y_1}(p) & \sigma_{y_2}^2 & \cdots & r_{a, y_2 y_L}(p) \\ \vdots & \vdots & \ddots & \vdots \\ r_{a, y_L y_1} & r_{a, y_L y_2}(p) & \cdots & \sigma_{y_L}^2(p) \end{bmatrix}$$

Se define además la matriz Σ como:

$$\Sigma = \begin{bmatrix} \sigma_{y_1} & 0 & \cdots & 0 \\ 0 & \sigma_{y_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{y_L}(p) \end{bmatrix} \quad (3.35)$$

La matriz $\mathbf{R}_a(p)$ se factoriza entonces como [Benesty et al., 2008]:

$$\mathbf{R}_a(p) = \Sigma \hat{\mathbf{R}}_a(p) \Sigma \quad (3.36)$$

donde $\hat{\mathbf{R}}_a(p)$ es la matriz de coeficientes de correlación normalizados, definida como:

$$\hat{\mathbf{R}}_a(p) = \begin{bmatrix} 1 & \rho_{a, y_1 y_2}(p) & \cdots & \rho_{a, y_1 y_L}(p) \\ \rho_{a, y_2 y_1}(p) & 1 & \cdots & \rho_{a, y_2 y_L}(p) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{a, y_L y_1}(p) & \rho_{a, y_L y_2}(p) & \cdots & 1 \end{bmatrix} \quad (3.37)$$

Con $\rho_{a, y_i y_j}(p)$ denotando el vector de correlación $r_{a, y_i y_j}(p)$ normalizado.

A continuación, se define el coeficiente de correlación cruzada multicanal(MCCC):

$$\rho_{a, y_1:y_L}(p) = 1 - \det[\hat{\mathbf{R}}_a(p)] \quad (3.38)$$

El MCCC posee un valor de 0 si y sólo si todas las señales no se encuentran correlacionadas entre sí, y 1 si dos o más señales están perfectamente correlacionadas. Cabe mencionar que si una de las señales no se encuentra correlacionada con el resto (debido a la corrupción de la señal debido a ruido, por ejemplo), el coeficiente retorna la correlación de las $L - 1$ señales restantes [Benesty et al., 2004]. Por tanto, se puede encontrar el TDOA como:

$$\tau = \max(\text{abs}[\rho_{a, y_{1:y_L}}(p)]) \quad (3.39)$$

O, alternativamente:

$$\tau = \min(\text{abs}[\det[\hat{\mathbf{R}}_a(p)]]]) \quad (3.40)$$

4

Procedimiento metodológico

4.1 Procedimiento realizado

Se debe elegir un algoritmo de procesamiento de datos que realice la determinación del ángulo azimutal de arribo de la señal de origen al arreglo de sensores. El mismo debe ser capaz de rechazar ruido blanco gaussiano. Para el análisis y comparación de los algoritmos expuestos anteriormente, se asume un arreglo hipotético lineal de micrófonos equidistantes entre sí. La señal de sonido fue extraída a partir del video provisto, e importada a *MATLAB* para su análisis. Se extrajo un segmento continuo de continuo de la misma para ser utilizado como señal de referencia. Se implementó un filtro de retardo discreto para simular la captura de la señal retardada por parte de los demás micrófonos. Finalmente, se introdujo una señal aditiva de ruido blanco a cada una de las señales de los demás micrófonos.

Los algoritmos fueron implementados en lenguaje *MATLAB*; se graficó su respuesta y comparó el resultado con el valor esperado de τ para determinar la precisión de los mismos, de acuerdo con las ecuaciones (3.10), (3.33), (3.40). Finalmente, se evaluó la cantidad de operaciones requerida para la ejecución de cada algoritmo.

5

Metodología de implementación

5.1 Reconocimiento y definición del problema

Se debe elegir un algoritmo de procesamiento de datos que realice la determinación del ángulo azimutal de arriba de la señal de origen al arreglo de sensores. El mismo debe ser capaz de rechazar ruido blanco gaussiano. Para el análisis y comparación de los algoritmos expuestos anteriormente, se asume un arreglo hipotético de dos, cuatro, y seis micrófonos, con una distancia d de cuatro metros entre sí. La señal utilizada para realizar las pruebas es una muestra de video y audio tomada de un camión en movimiento, con una frecuencia de muestreo de $45,7kHz$. Se asume que la señal se aproxima al arreglo con un ángulo θ de 67° , lo cual corresponde a un retardo teórico τ de aproximadamente $4,578ms$, o 200 muestras. Este es el τ hipotético que se introduce como retardo a las señales de los micrófonos $y_2(k)$, $y_3(k)$, $y_4(k)$, $y_5(k)$ y $y_6(k)$, escalado apropiadamente por la ecuación (3.3), presentada en el capítulo anterior.

5.2 Obtención y acondicionamiento de la señal

La señal de sonido fue extraída a partir del video provisto, e importada a *MATLAB* para su análisis. Se extrajo un segmento continuo de 10000 muestras de la señal; este representa la señal de referencia. Se implementó un filtro de retardo discreto para simular la captura de la señal retardada por parte de los demás micrófonos. Finalmente, se introdujo una señal aditiva de ruido blanco a cada una de las señales $y_2(k)$, $y_3(k)$, $y_4(k)$, $y_5(k)$ y $y_6(k)$, con el fin de simular el ruido adicional introducido por cada sensor.

5.3 Correlación cruzada

Se optó por implementar los algoritmos en *MATLAB*, debido a la flexibilidad del lenguaje en el manejo de vectores y matrices. Inicialmente, se implementó la correlación cruzada en forma discreta, utilizando el algoritmo descrito por la ecuación (3.11):

$$r_{y_1 y_2}^{\text{CC}}(p) = \begin{cases} \frac{1}{K} \sum_{i=0}^{K-p-1} y_1(k+i)y_2(k+i+p) & \text{si } p \geq 0, \\ r_{y_2 y_1}^{\text{CC}}(-p) & \text{si } p < 0. \end{cases}$$

Sin embargo, al probar la ejecución del mismo, la cantidad de operaciones necesaria resultó excesiva. Se optó por utilizar la transformada rápida de Fourier, descrita en el capítulo 3.

Se implementó el cálculo de correlación cruzada mediante el uso de la FFT, utilizando la ecuación (3.14). Sin embargo, la validez de la ecuación (3.14) se basa en el teorema de convolución para transformadas de Fourier, el cual asume dos características con respecto a las señales de entrada:

- La longitud de ambos vectores es equivalente.
- El espectro de las señales es periódico.

Si bien las señales a utilizar son vectores de igual longitud, ya que se toma un número equivalente de muestras a partir de cada micrófono, la señal capturada no es periódica. Sea $v(k)$ el vector de largo N estacionario original, y $r(k)$ el vector de largo M a desplazar. El teorema asume que la señal a desplazar posee un espectro en frecuencia desde $-M/2$ a $M/2$. Por convención, se asume que los primeros $M/2$ datos de un vector son los valores positivos de su espectro en frecuencia, y los demás corresponden a los valores negativos, con la posición $M-1$ correspondiente a $r(-1)$, y así sucesivamente. La figura 5.1 ilustra gráficamente el almacenamiento de los índices restantes de la señal a desplazar.

Esto implica que la convolución de los últimos $M/2$ términos del vector *estacionario* $v(k)$ y la señal desplazada $r(k)$ será sumada a los primeros índices de la señal resultante, y la convolución de los primeros $M/2$ términos se sumará a los últimos términos del resultado. La figura 5.2 ilustra este efecto, conocido como *wrap-around*:

Para corregir este problema, es necesario extender el largo de la señal original, agregando ceros al vector estacionario de forma que los mismos eliminen el efecto.

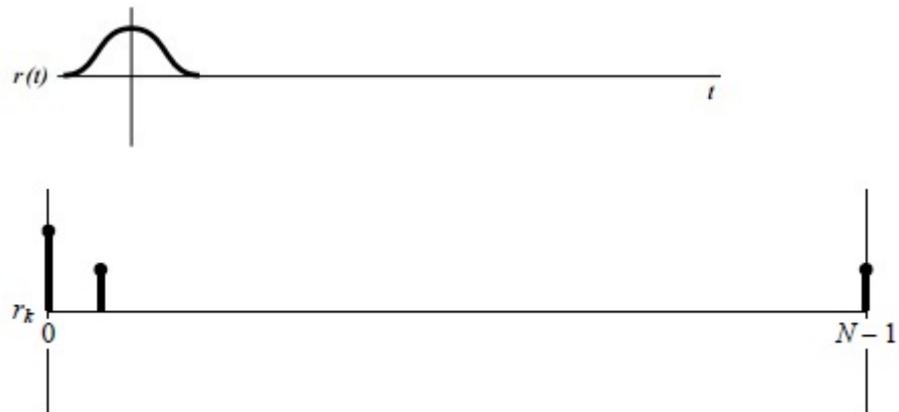


FIGURA 5.1: Señal de respuesta a desplazar. Nótese que los valores negativos se almacenan en la segunda mitad del vector [Ziegel et al., 1987].

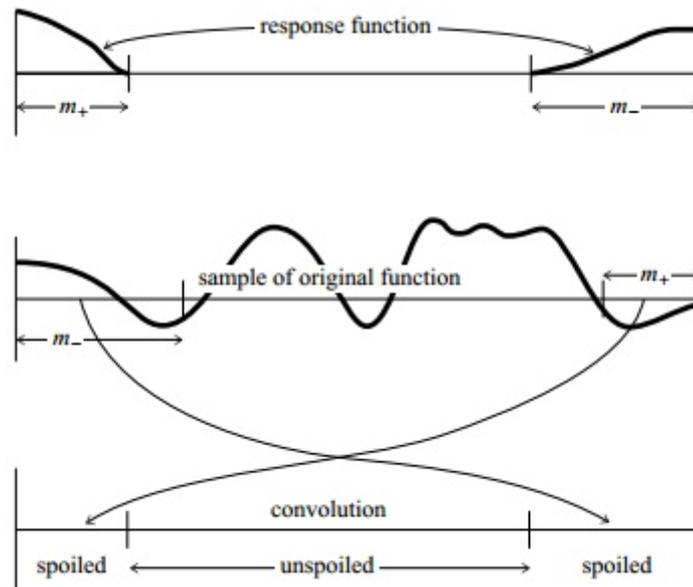


FIGURA 5.2: Respuesta de la convolución de una señal no periódica [Ziegel et al., 1987].

Dado que el vector a desplazar en el caso bajo estudio es de largo N , se requiere un relleno mínimo de $N - 1$ ceros agregados al vector original. Sin embargo, debe tomarse en consideración además la primera restricción del teorema, que exige que las señales sean de largo equivalente; por tanto, el relleno debe aplicarse también al segundo vector. Si se aplica el relleno del primer vector al inicio del mismo y el relleno del segundo vector al final de este, puede eliminarse por completo el efecto del *wrap-around*, obteniendo dos vectores de longitud $(2N - 1)$. Los vectores pueden escribirse entonces como:

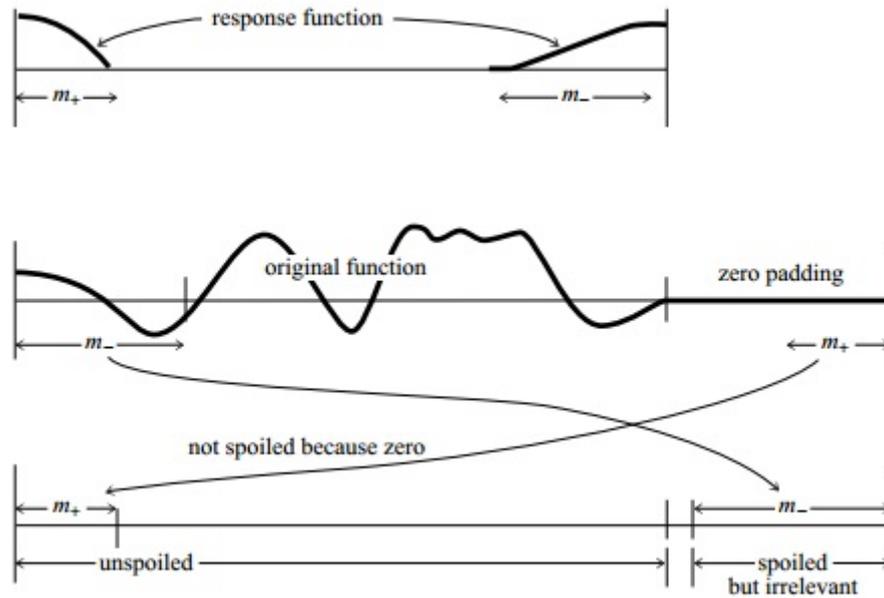


FIGURA 5.3: Ilustración del uso de ceros para eliminar el efecto del *wrap-around* [Ziegel et al., 1987].

$$v_1(k) = [\underbrace{0, \dots, 0}_{N-1}, v_1(1), v_1(2), \dots, v_1(N)] \quad (5.1)$$

$$v_2(k) = [v_2(1), v_2(2), \dots, v_2(N), \underbrace{0, \dots, 0}_{N-1}] \quad (5.2)$$

El cálculo de la FFT se realiza utilizando el algoritmo de *Cooley-Tukey* en base 2, que divide una transformada de Fourier discreta de K términos en dos transformadas de $K/2$ términos, separando los términos pares de los impares recursivamente. Para optimizar el tiempo de ejecución del algoritmo, se fijó la longitud de los vectores a la potencia de dos superior más cercana a $(2N - 1)$, definiendo una nueva longitud $Q = 2^{\text{nextpow}2(2N-1)}$. Los vectores resultantes utilizados son:

$$v_1(k) = [\underbrace{0, \dots, 0}_{N-1}, v_1(1), v_1(2), \dots, v_1(N), \underbrace{0, \dots, 0}_{P-2N+1}] \quad (5.3)$$

$$v_2(k) = [v_2(1), v_2(2), \dots, v_2(N), \underbrace{0, \dots, 0}_{P-N}] \quad (5.4)$$

Finalmente, se procedió a calcular la correlación entre las señales utilizando la ecuación (3.14). La figura 5.4 describe el flujo del algoritmo.

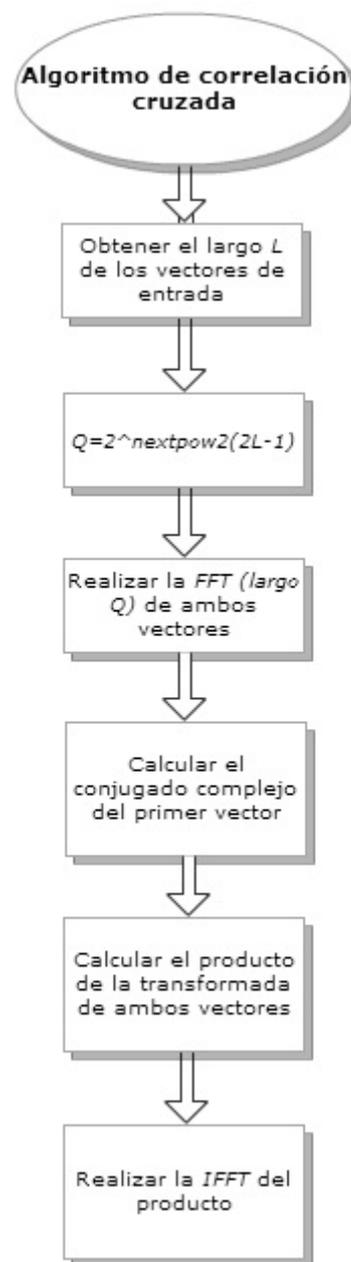


FIGURA 5.4: Diagrama de flujo del algoritmo de correlación cruzada utilizado.

5.4 Forward Spatial Linear Prediction

Dado que se fija la distancia entre los micrófonos del arreglo inicialmente, es posible definir un retardo máximo p_{max} esperado, el cual se obtiene examinando el caso límite, en el cual el frente de onda se aproxima al arreglo con un ángulo de 0° . Este viene a ser definido por la ecuación:

$$\tau_{max} = \frac{d \cos(0^\circ)}{c} \quad (5.5)$$

Con $d = 4m$, $c = 343m/s$, se obtiene $\tau_{max} = 11,66 ms$, lo cual equivale a un retardo de aproximadamente 498 muestras. Se fija entonces:

$$[p_{min}, p_{max}] = [-500, 500] \quad (5.6)$$

Esto permite reducir el número de cálculos al limitar la búsqueda de p únicamente dentro de los valores posibles para el arreglo de estudio. Se procedió a adelantar las señales de los micrófonos, de la forma:

$$y_n[k + F_n(p)] = y_n[k + (n - 1)p] \quad (5.7)$$

A continuación, se calculó para cada iteración de p hipotético la matriz de correlación espacial definida en las ecuaciones (3.24) y (3.29):

$$\mathbf{R}_{a,2:L}(p) = \begin{bmatrix} \sigma_{y2}^2 & r_{a, y_2 y_3}(p) & \cdots & r_{a, y_2 y_L}(p) \\ r_{a, y_3 y_2}(p) & \sigma_{y2}^2 & \cdots & r_{a, y_2 y_L}(p) \\ \vdots & \vdots & \ddots & \vdots \\ r_{a, y_L y_1} & r_{a, y_L y_2}(p) & \cdots & \sigma_{yL}^2(p) \end{bmatrix}$$

Dado que la solución del algoritmo es matemáticamente compleja, y requiere el cálculo de la inversa de la matriz de correlación espacial, se utilizó un método alternativo para encontrar el error al método mostrado anteriormente. Puede demostrarse que [Benesty et al., 2008]:

$$J_{1, \min}(p) = \sigma_{y1}^2 \frac{\det[\hat{\mathbf{R}}_a(p)]}{\det[\hat{\mathbf{R}}_{a,2:L}(p)]} \quad (5.8)$$

La figura 5.5 describe el flujo del algoritmo.

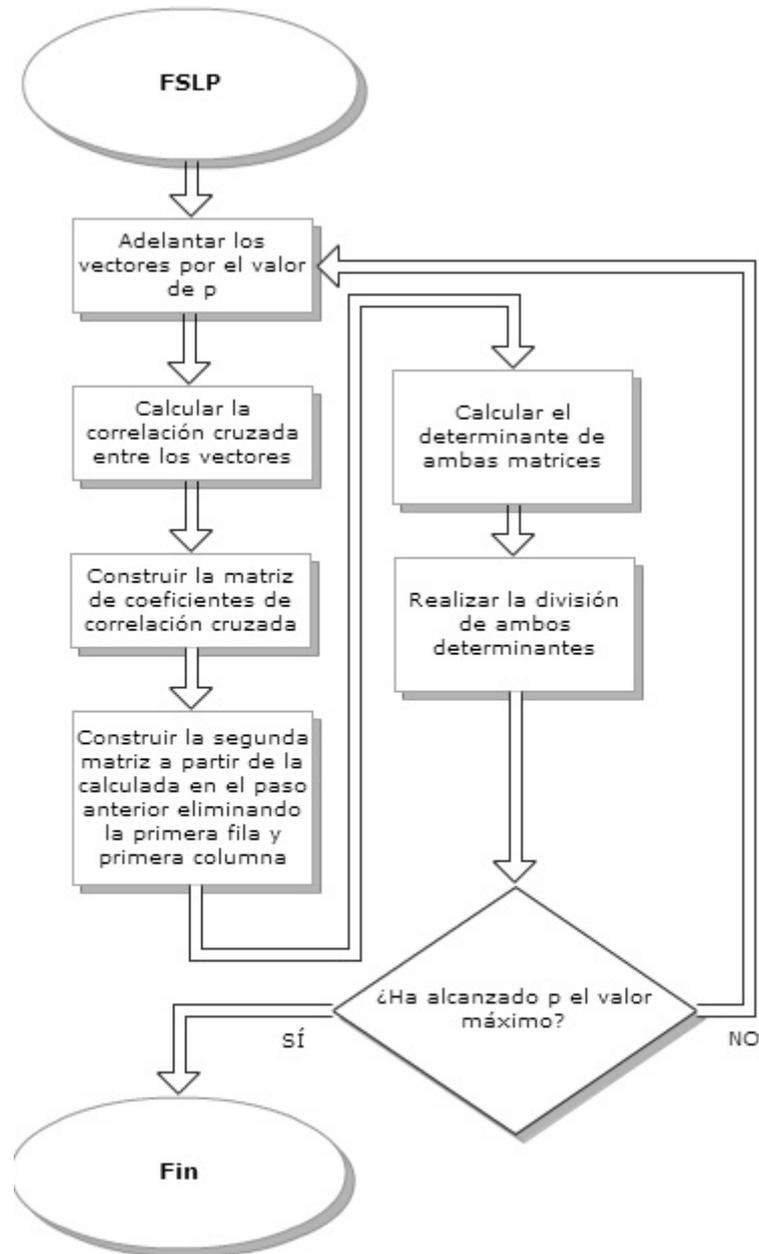


FIGURA 5.5: Diagrama de flujo del algoritmo FSLP utilizado.

5.5 Multichannel Cross Correlation Coefficient

Se inició calculando para cada iteración de p hipotético la matriz de correlación espacial definida en las ecuaciones (3.24) y (3.29), de igual manera que en el algoritmo expuesto anteriormente, y la matriz de coeficientes de correlación espacial. A diferencia del algoritmo anterior, como se observa en la ecuación (3.38), únicamente es necesario el cálculo de la matriz $\hat{\mathbf{R}}_a(p)$. La figura 5.6 describe el flujo del algoritmo.

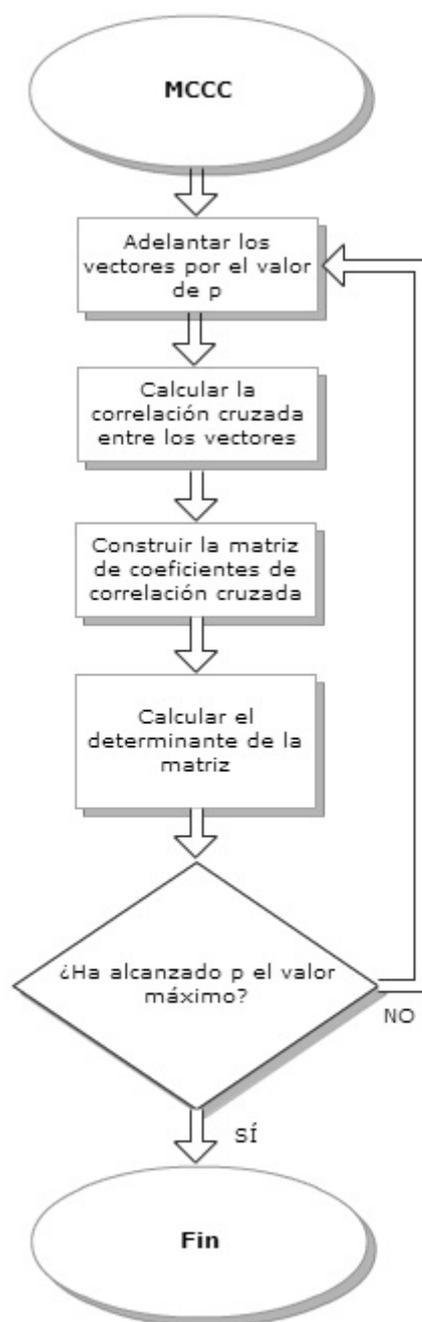


FIGURA 5.6: Diagrama de flujo del algoritmo MCCC utilizado.

6

Análisis de resultados

6.1 Resultados

Se presentan en las figuras 6.1, 6.2, y 6.3 muestras de las señales utilizadas para realizar las pruebas de los algoritmos.

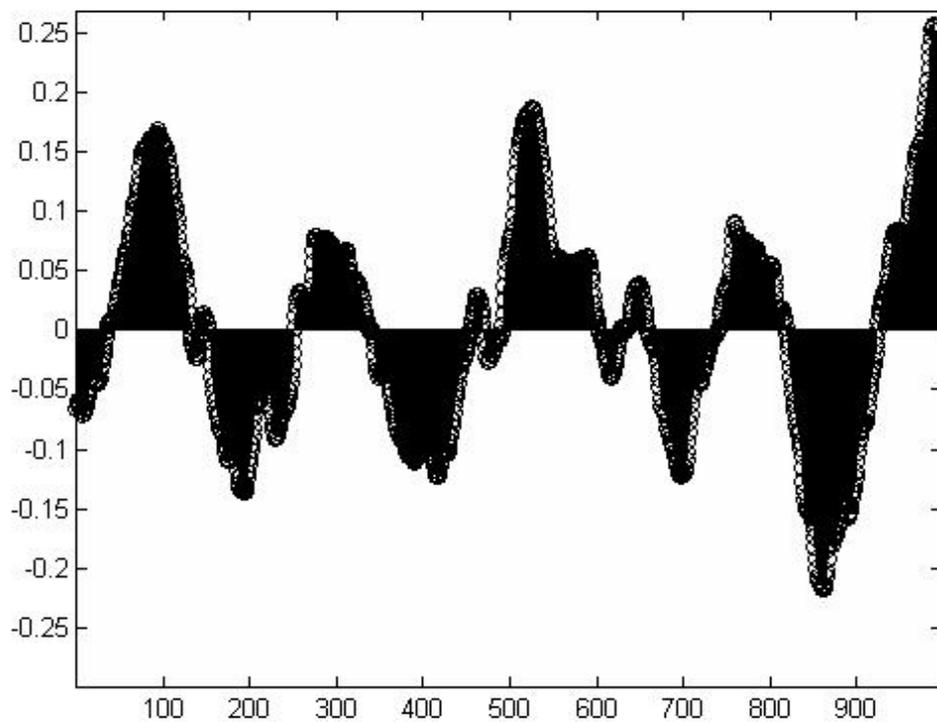


FIGURA 6.1: Muestra de control de la señal $y_1(k)$.

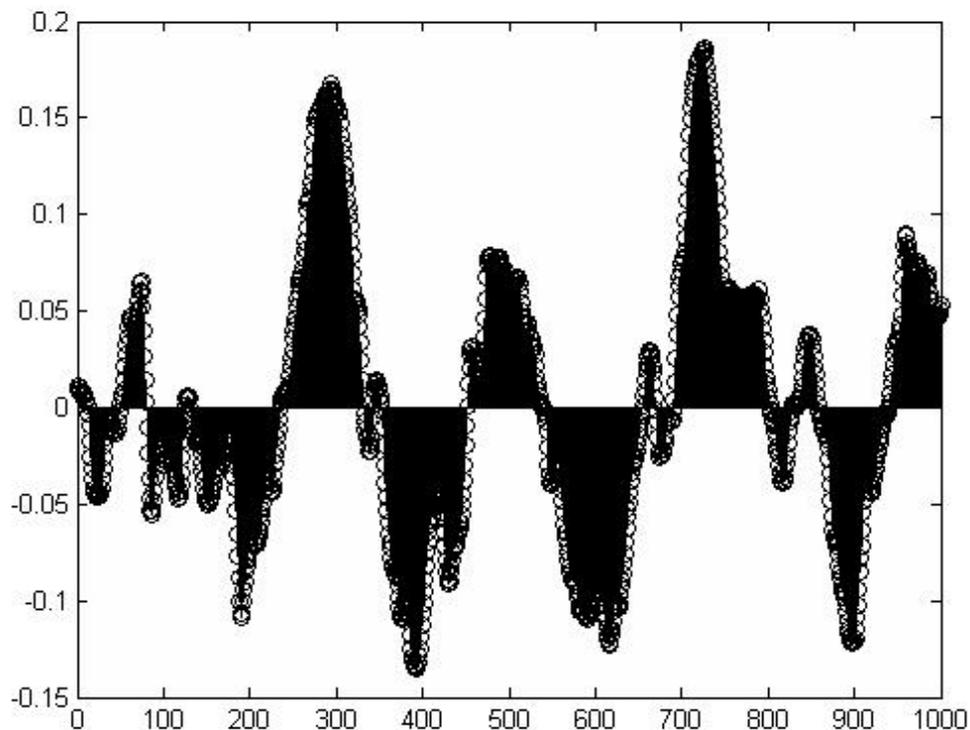


 FIGURA 6.2: Muestra de control de la señal $y_2(k)$.

6.1.1 Correlación Cruzada

Se presentan inicialmente, a modo de control, los resultados obtenidos al calcular la correlación entre las señales $y_1(k)$ y $y_2(k)$ en la figura 6.4.

Las figuras 6.5 6.6 6.7 corresponden a las pruebas realizadas utilizando ruido aleatorio de media cero. La razón señal a ruido indica el nivel de ruido introducido a la señal $y_2(k)$.

Cabe destacar que si bien el algoritmo logra encontrar el TDOA para valores de SNR mayores o iguales a $6dB$, se observa una descomposición rápida del vector p al incrementar la amplitud del valor de ruido introducido. El impulso claramente observable en la figura 6.4 se vuelve menos definido conforme se incrementa la amplitud del ruido, y el algoritmo falla en encontrar el TDOA apropiado cuando $SNR < 6dB$.

Se analizó a continuación el costo computacional de implementar el algoritmo, utilizando para ello el número de operaciones requeridas para ejecutarlo. Sea L el número de micrófonos utilizados en el arreglo. La operación más utilizada es la DFT, la cual puede implementarse utilizando el algoritmo de FFT de *Cooley-Tukey*. Se fijó el largo

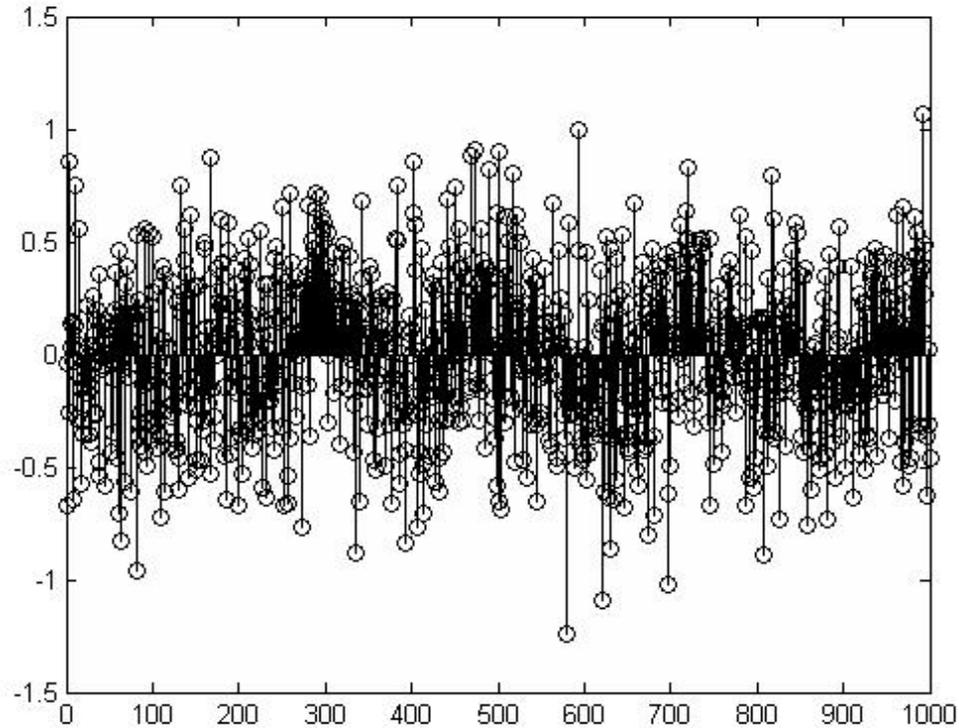


FIGURA 6.3: Muestra de la señal $y_2(k)$, luego de agregar ruido. $SNR = 9dB$.

del vector anteriormente a una potencia de 2, por tanto, para vectores de tamaño Q potencia de 2 como los definidos en la ecuación (5.3), el costo de realizar la transformada es de $\frac{Q}{2}\log_2(Q)$ multiplicaciones complejas y $Q\log_2(Q)$ sumas complejas, donde cada multiplicación compleja equivale a cuatro multiplicaciones sencillas y dos sumas simples, y cada suma compleja equivale a dos sumas simples. Por tanto, el costo de implementar la FFT para un vector de largo Q es de $5Q\log_2(Q)$ operaciones sencillas. La transformada debe aplicarse a todas las señales recibidas de cada sensor, por tanto, el costo total es de $5LQ\log_2(Q)$ [Do, 2009].

Para obtener el vector de correlación cruzada, es necesario realizar una multiplicación de las transformadas de los vectores, lo cual equivale a una multiplicación compleja, o seis operaciones sencillas. Finalmente, debe obtenerse la transformada inversa del par de micrófonos bajo estudio; esto equivale a $5\frac{L}{2}Q\log_2(Q)$ operaciones. Sumando estas al total obtenido para la FFT, se obtiene $15\frac{L}{2}Q\log_2(Q)$ operaciones. Si se asume que no se implementa una función de peso ϑ , el costo computacional del método de correlación cruzada es:

$$Costo_{CC} = 15\frac{L}{2}Q\log_2(Q) \quad (6.1)$$

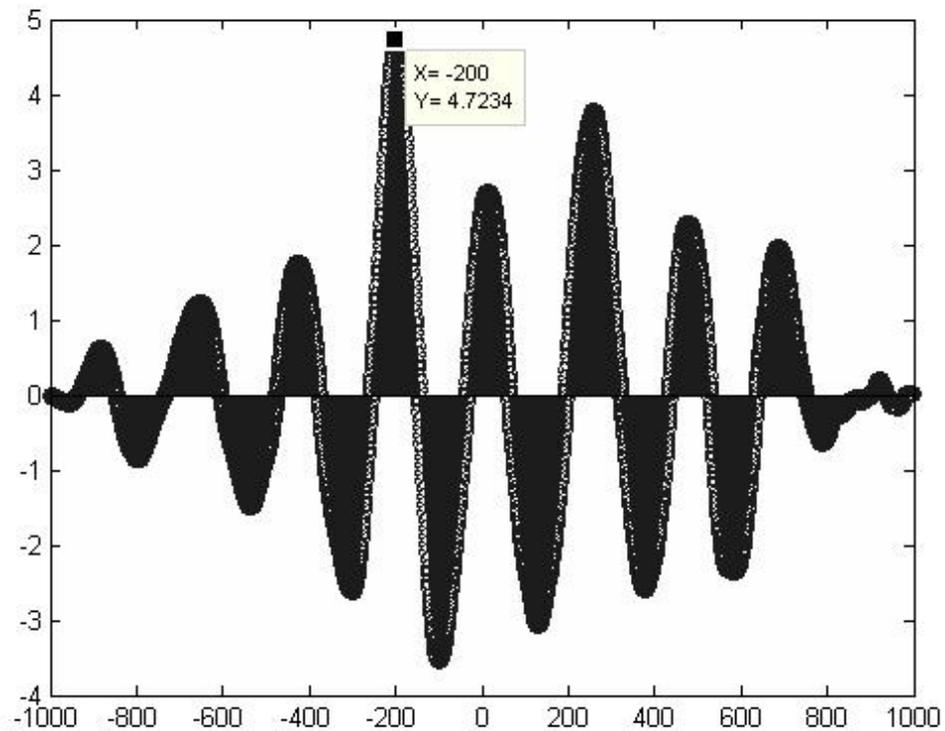


FIGURA 6.4: Vector p de correlación obtenido mediante el uso de CC entre las señales $y_1(k)$ y $y_2(k)$. No se introdujo ruido a la señal $y_2(k)$.

6.1.2 MCCC

Se presenta inicialmente en la figura 6.8 el resultado de implementar el algoritmo MCCC con 2, 4 y 6 micrófonos en el arreglo, sin introducción de ruido a las señales.

A continuación, se presenta en la figura 6.9 el resultado de realizar la misma prueba, introduciendo una señal de ruido, con $SNR = 6dB$.

Finalmente, se comprobó la propiedad de rechazo de canales no correlacionados del algoritmo MCCC. Para ello, se introdujeron dos señales completamente compuestas de ruido no correlacionado a la señal de origen $s(k)$, así como cuatro canales de información válida con $SNR = 6dB$. El resultado de la prueba se muestra en la figura 6.10.

El algoritmo alcanza el valor mínimo cuando $p = \tau$, aún ante la introducción de ruido y señales no correlacionadas con la señal de control, siendo la respuesta directamente proporcional a la cantidad de micrófonos utilizados en el arreglo.

Para el análisis de costo del algoritmo MCCC, es importante notar que la matriz $\hat{\mathbf{R}}_a(p)$ es una matriz Toeplitz, con todos los términos de su diagonal equivalentes entre sí.

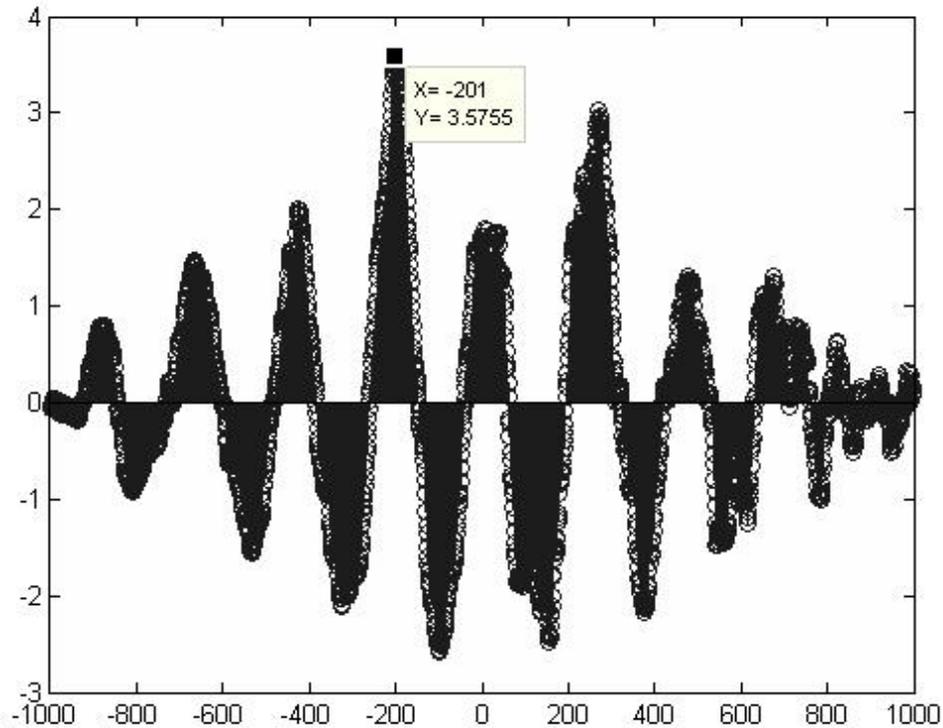


FIGURA 6.5: Vector p de correlación obtenido mediante el uso de CC entre las señales $y_1(k)$ y $y_2(k)$. $SNR = 9dB$.

Por tanto, una descomposición LU para encontrar su determinante requiere únicamente de L^2 operaciones. En el caso de MCCC, se requiere calcular τ_{max} determinantes. Típicamente, el valor de L (cantidad de micrófonos) es pequeño comparado con el valor de Q (cantidad de muestras, extendidas para satisfacer las condiciones del teorema de convolución). Por tanto, el incremento en el número de operaciones comparado al método de correlación cruzada no es significativo.

$$Costo_{MCCC} = PL^2 + 15\frac{L}{2}Q\log_2(Q) \quad (6.2)$$

6.1.3 FSLP

Se presenta inicialmente el resultado de implementar el algoritmo FSLP con 2, 4 y 6 micrófonos en el arreglo, sin introducción de ruido a las señales, en la figura 6.11.

A continuación, se presenta en las figuras 6.12 6.13 el resultado de realizar la misma prueba, introduciendo una señal de ruido, con $SNR = 6dB$.

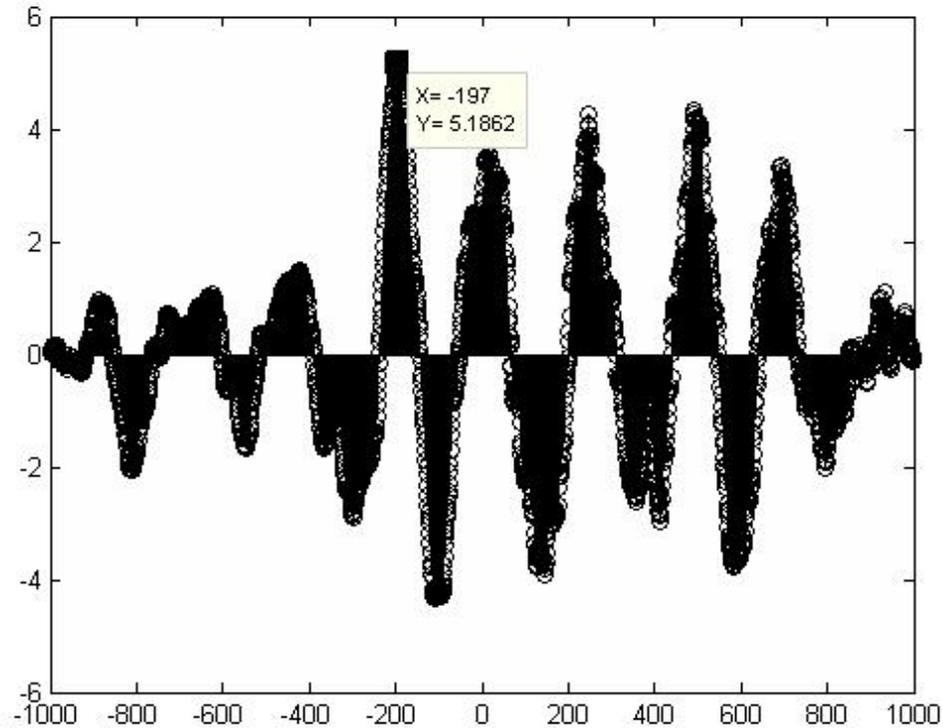


FIGURA 6.6: Vector p de correlación obtenido mediante el uso de CC entre las señales $y_1(k)$ y $y_2(k)$. $SNR = 6dB$.

Se probó a continuación la propiedad de rechazo de señales no correlacionadas del algoritmo. Para ello, se introdujeron dos señales completamente compuestas de ruido no correlacionado a la señal de origen $s(k)$, así como cuatro canales de información válida con $SNR = 6dB$. El resultado de la prueba se muestra en la figura 6.10.

Se observa una respuesta clara en $p = \tau$ aún ante la presencia de ruido, con la precisión de la respuesta siendo directamente proporcional a la cantidad de micrófonos utilizados en el arreglo, aunque la respuesta al ruido presenta mayor cantidad de respuestas de mayor amplitud al ser comparada con el algoritmo anterior. El algoritmo logra rechazar señales no correlacionadas, pero sufre inestabilidad cuando la matriz $\hat{\mathbf{R}}_{a,2:L}(p)$ tiende a valores cercanos a cero.

El algoritmo FSLP requiere el cálculo de P determinantes adicionales y P divisiones con respecto al cálculo del MCCC:

$$Costo_{FSLP} = 2PL^2 + P + 15\frac{L}{2}Q\log_2(Q) \quad (6.3)$$

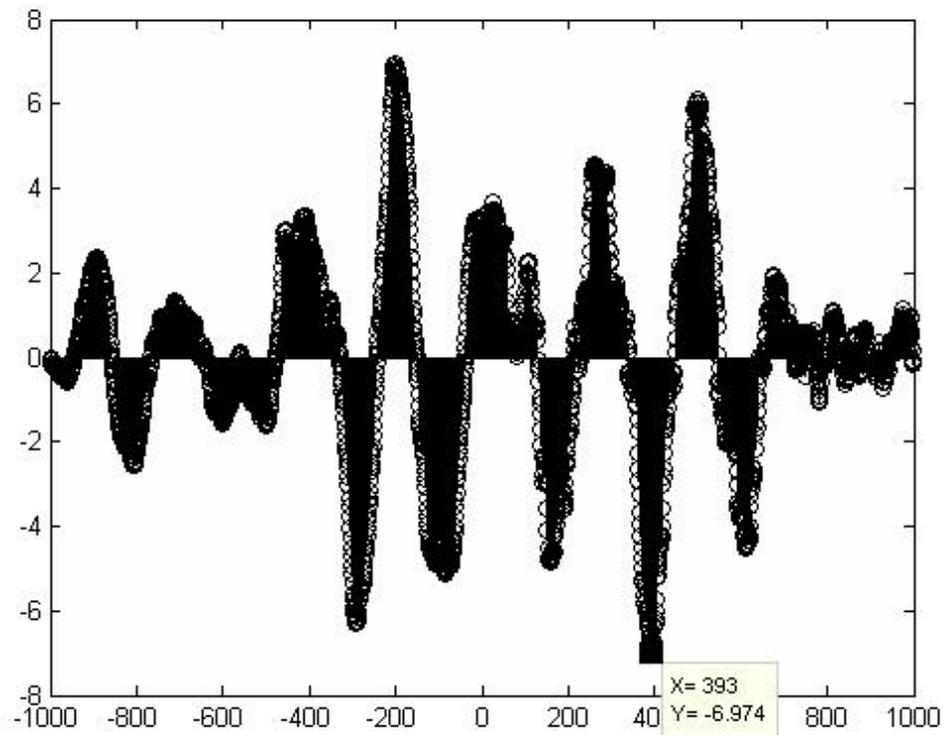


FIGURA 6.7: Vector p de correlación obtenido mediante el uso de CC entre las señales $y_1(k)$ y $y_2(k)$. $SNR = 3dB$.

Se presenta a modo de resumen la siguiente tabla con las propiedades de cada algoritmo, comparadas utilizando una razón señal-ruido $SNR = 6dB$.

Porcentaje de error observado	Cantidad de operaciones	Rechazo de canal
1,5%	$15\frac{L}{2}Q\log_2(Q)$	No
-2%	$2PL^2 + P + 15\frac{L}{2}Q\log_2(Q)$	Limitado (Inestable)
0,5%	$PL^2 + 15\frac{L}{2}Q\log_2(Q)$	Sí

TABLA 6.1: Tabla de comparación de los algoritmos estudiados.

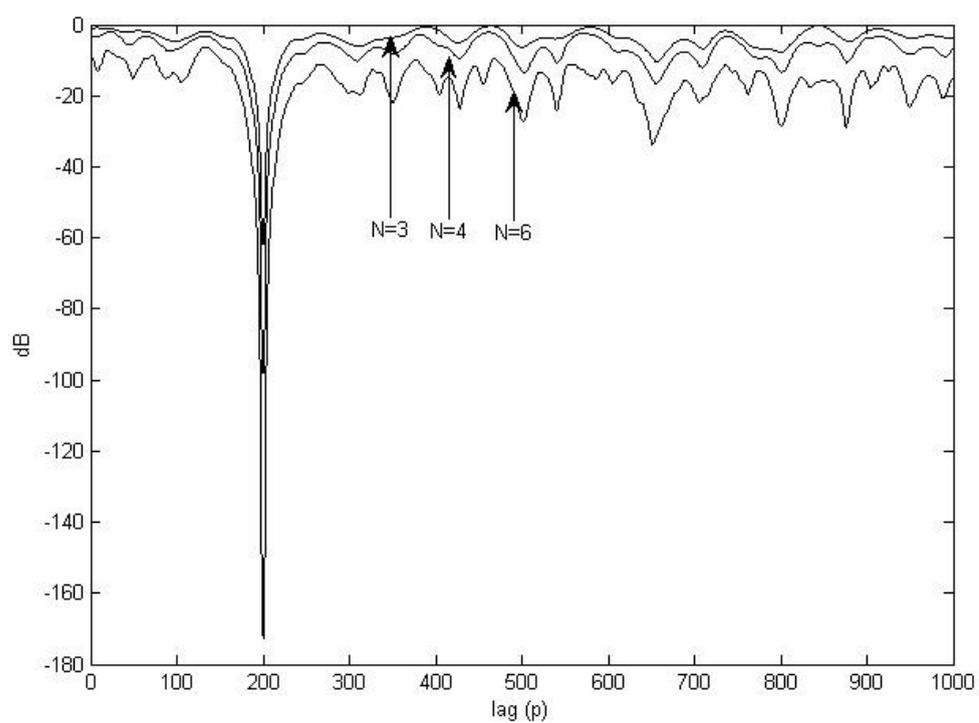


FIGURA 6.8: Vector p de correlación obtenido mediante el uso de MCCC para $N = 2$, $N = 4$, $N = 6$. No se introdujo ruido a las señales.

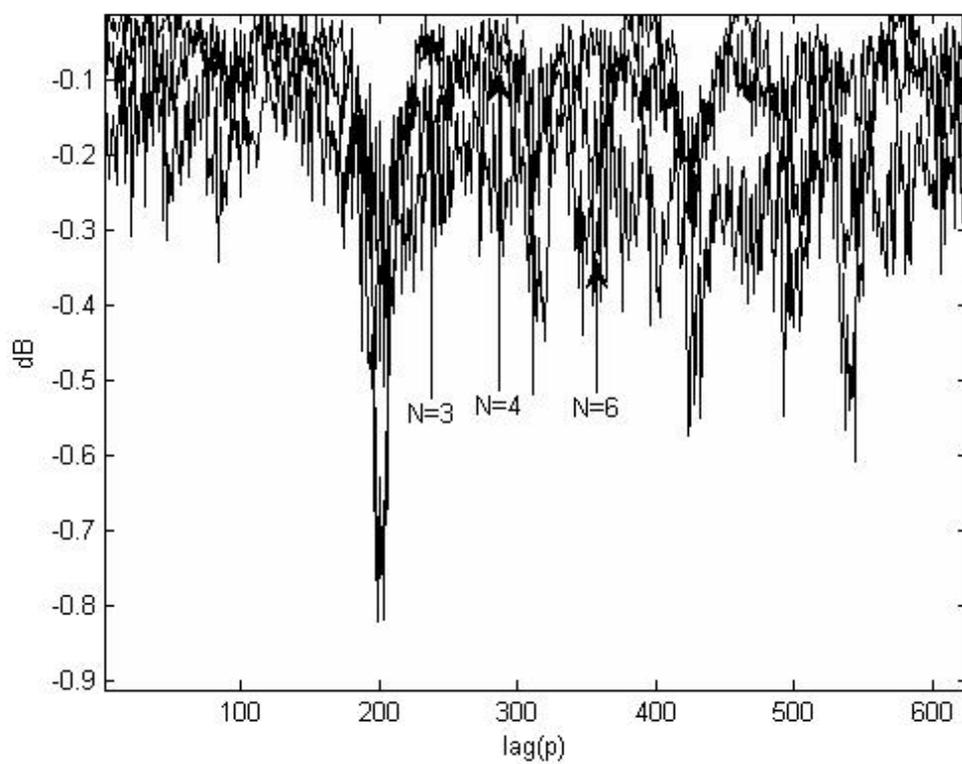


FIGURA 6.9: Vector p de correlación obtenido mediante el uso de MCCC para $N = 2$, $N = 4$, $N = 6$. $SNR = 6dB$.

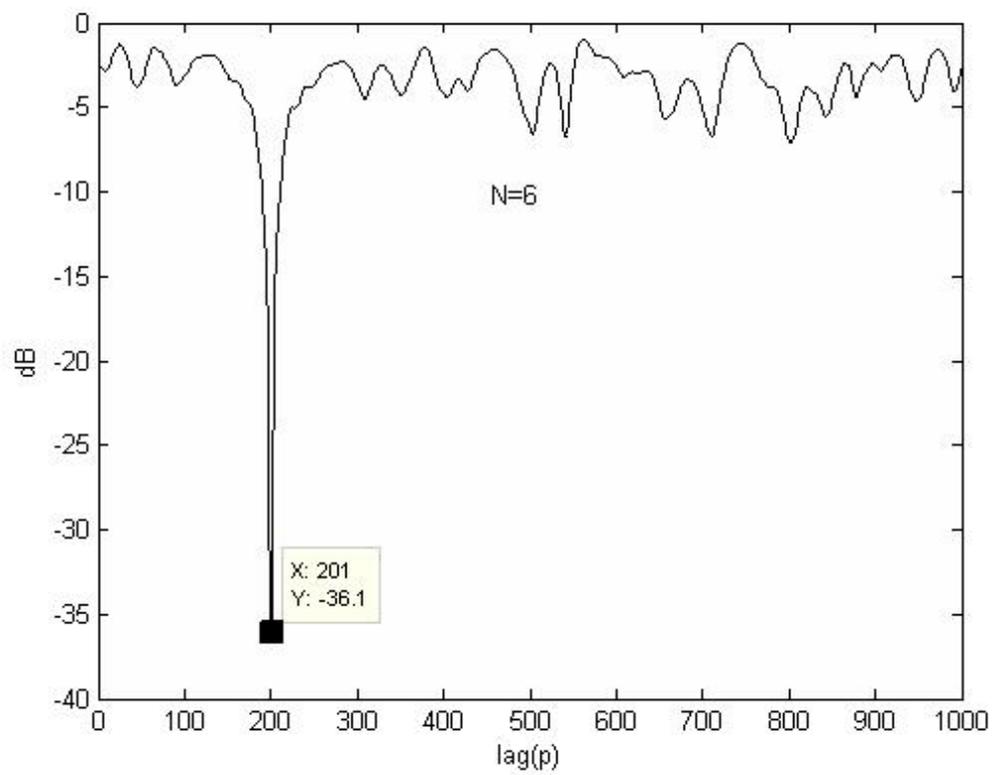


FIGURA 6.10: Vector p de correlación obtenido mediante el uso de MCCC para $N = 6$, con $y_5(k)$, $y_6(k)$ señales de ruido no correlacionadas con la señal $s(k)$. $SNR = 6dB$.

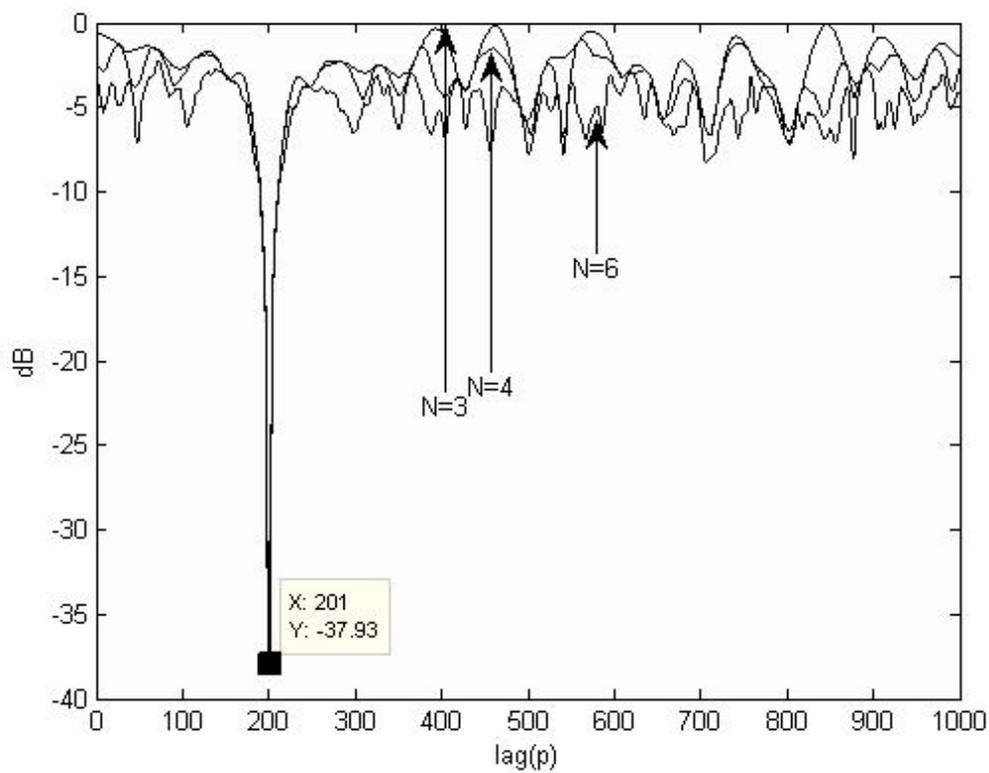


FIGURA 6.11: Vector p de correlación obtenido mediante el uso de FSLP para $N = 2$, $N = 4$, $N = 6$. No se introdujo ruido a las señales.

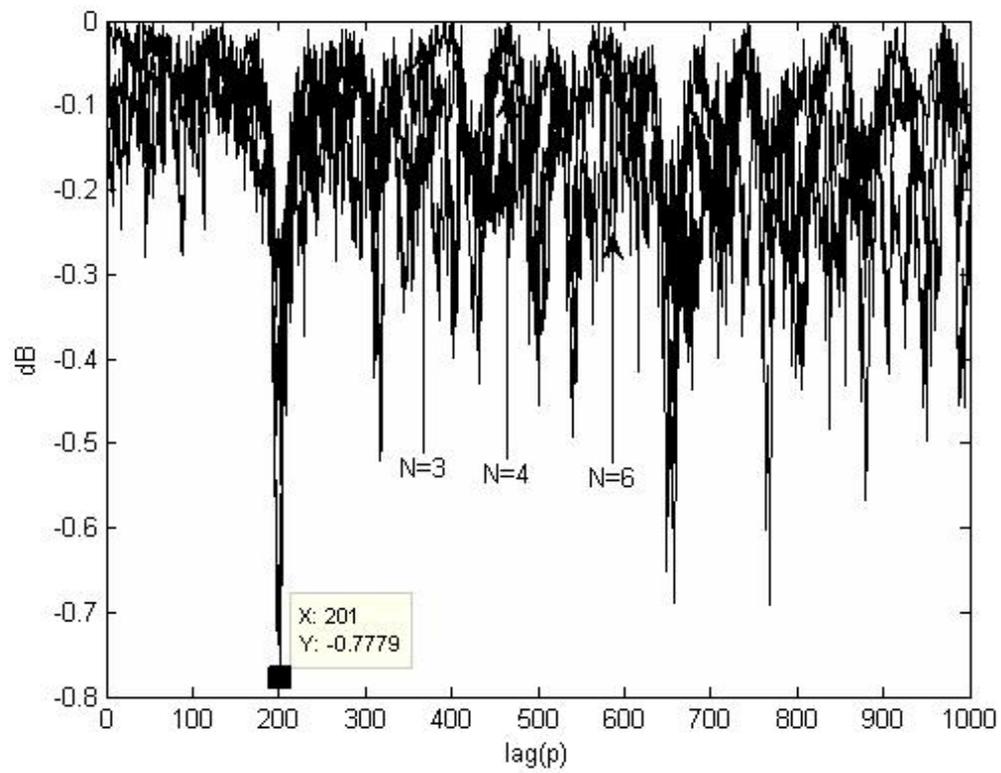


FIGURA 6.12: Vector p de correlación obtenido mediante el uso de FSLP para $N = 2$, $N = 4$, $N = 6$. $SNR = 6dB$.

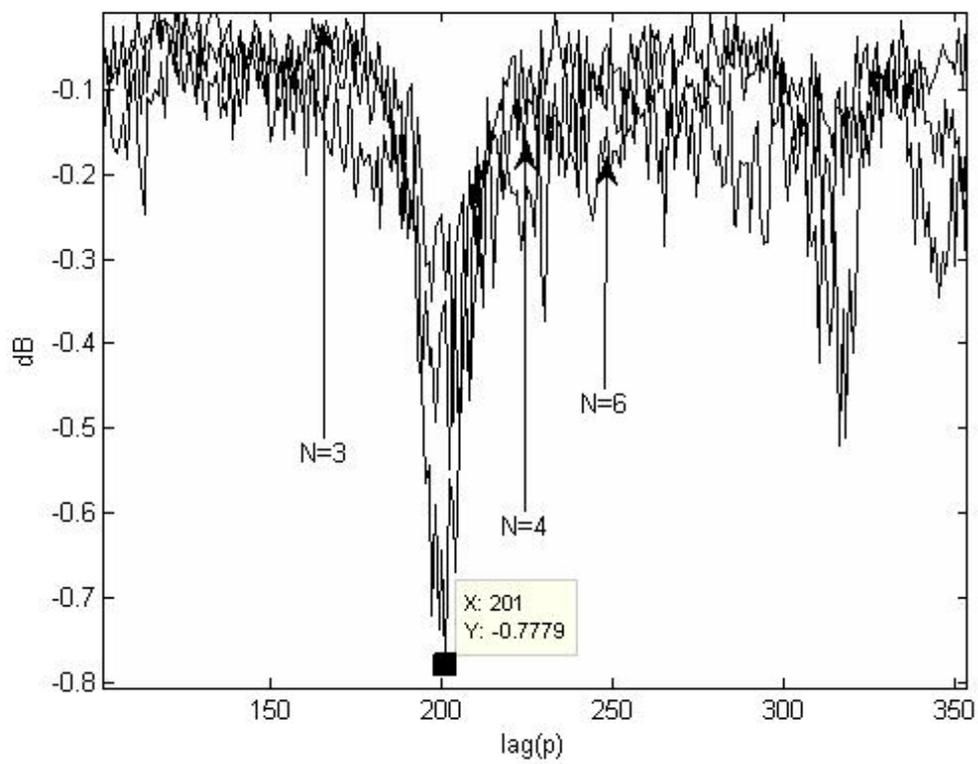


FIGURA 6.13: Vector p de correlación obtenido mediante el uso de FSLP para $N = 2$, $N = 4$, $N = 6$. $SNR = 6dB$.

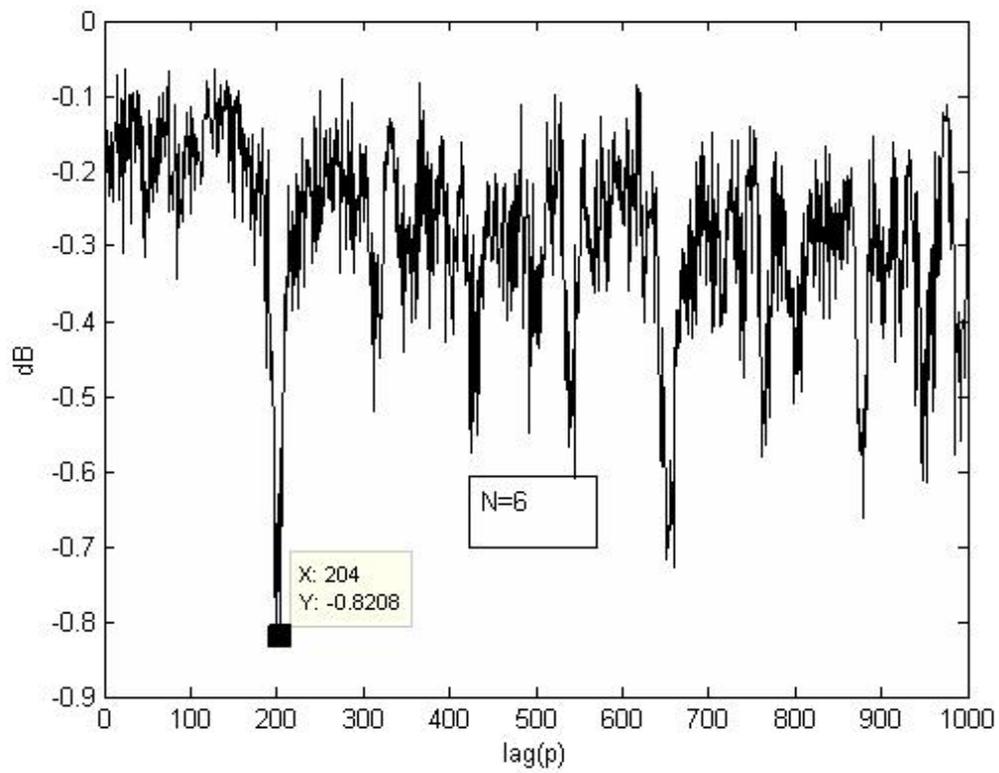


FIGURA 6.14: Vector p de correlación obtenido mediante el uso de FSLP para $N = 6$, con $y_5(k)$, $y_6(k)$ señales de ruido no correlacionadas con la señal $s(k)$. $SNR = 6dB$.

Conclusiones y Recomendaciones

7.1 Conclusiones

Los algoritmos de fusión de vectores de correlación cruzada son capaces de encontrar el TDOA planteado aún bajo condiciones adversas de ruido externo introducido a los micrófonos, obteniéndose una desviación máxima de cuatro muestras, lo cual equivale a un error de 2%. Se observa además que la respuesta de los algoritmos se torna más definida al incrementar el número de sensores, pero a su vez, aumenta el número de operaciones necesarias para la ejecución del mismo.

Se comprobó además la propiedad de rechazo de canales no correlacionados del algoritmo MCCC, lo cual lo vuelve robusto ante la influencia de ruido [Benesty et al., 2008]. El algoritmo FSLP no sólo requiere de una mayor cantidad de operaciones al necesitar el cálculo de un segundo determinante y una división por cada valor de p a estimar, sino también posee el riesgo de experimentar inestabilidades cuando el valor de la matriz $\hat{\mathbf{R}}_{a,2:L}(p)$ tiende a cero. Esto ocurre cuando una o más de las señales de entrada no se encuentra correlacionada con las demás, razón por la cual este algoritmo no puede realizar el rechazo de canales no correlacionados en forma confiable.

A partir de los datos experimentales obtenidos y el análisis de costo de los algoritmos estudiados, se recomienda el uso del algoritmo MCCC como punto de partida para resolver la meta propuesta, debido a las siguientes características:

- Es robusto frente a la introducción de ruido.
- Rechaza señales no correlacionadas con la señal buscada.
- Posee una precisión aceptable que puede escalarse de acuerdo a la cantidad y posicionamiento de los micrófonos pertenecientes al arreglo.

- Su costo computacional es similar al costo de implementar la correlación cruzada sencilla.

El algoritmo presenta una respuesta visiblemente pronunciada en el retardo esperado, aún ante la presencia de ruido aleatorio.

7.2 Recomendaciones

Es posible implementar un filtro PHAT como el descrito en las ecuaciones (3.18) y (3.21), a un costo de Q divisiones adicionales para cada par de sensores, lo cual permite desensibilizar al algoritmo ante las variaciones de volumen en la fuente. [Zarifeh et al., 2007]. Existen además otros algoritmos de fusión basados en el uso del cálculo del vector propio correspondiente a la matriz de correlación espacial [Benesty et al., 2008], así como métodos de mínima entropía. No se estudió en el documento presente la opción de realizar la localización mediante la información de potencia del espectro en frecuencia de la señal de origen [Birchfield and Gangishetty, 2005].

Bibliografía

- Katsundo Hitomi. Automation its concept and a short history. *Technovation*, 14(2): 121–128, March 1994. ISSN 01664972. doi: 10.1016/0166-4972(94)90101-5. URL [http://dx.doi.org/10.1016/0166-4972\(94\)90101-5](http://dx.doi.org/10.1016/0166-4972(94)90101-5).
- Raja Parasuraman and Victor Riley. Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 39(2):230–253, June 1997. ISSN 00187208. doi: 10.1518/001872097778543886. URL <http://hfs.sagepub.com/cgi/content/abstract/39/2/230>.
- T. Inagaki. Design of humanmachine interactions in light of domain-dependence of human-centered automation. *Cognition, Technology & Work*, 8(3):161–167, April 2006. ISSN 1435-5558. doi: 10.1007/s10111-006-0034-z. URL <http://www.springerlink.com/content/m6158053666r6u41/>.
- Pedro Julián, Masson Favio, Pablo Mandolesi, Andreas Andreou, Shihab Shamma, Eduardo Nebot, and Alfonso Chacón. 3D Gigascale Integrated Circuits for Nonlinear Computation, Filter and Fusion with Applications in Industrial Field Robotics. Technical report.
- Richard Piggin. What’s happening with machine safety standards and networks? *Assembly Automation*, 26(2):104–110, 2006. ISSN 0144-5154. doi: 10.1108/01445150610658077. URL <http://www.emeraldinsight.com/10.1108/01445150610658077>.
- Deborah A. Boehm-Davis, Renwick E. Curry, Earl L. Wiener, and R. Leon Harrison. Human factors of flight-deck automation: Report on a NASA-industry workshop. *Ergonomics*, 26(10):953–961, October 1983. ISSN 0014-0139. doi: 10.1080/00140138308963424. URL <http://www.tandfonline.com/doi/abs/10.1080/00140138308963424>.
- Y. Dieudonne, O. Labbani-Igbida, and F. Petit. Deterministic Robot-Network Localization is Hard. *IEEE Transactions on Robotics*, 26(2):331–339, April 2010. ISSN 1552-3098. doi: 10.1109/TRO.2010.2042753. URL <https://doi.org/10.1109/TRO.2010.2042753>.

//sibdi.ucr.ac.cr/http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=5427034&contentType=Journals+%26+Magazines&searchField%3DSearch_All%26queryText%3Ddeterministic+robot-network+localization+is+hard+workhttp://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5427034.

Thomas Bell, Gabriel Elkaim, and Dr. Bradford Parkinson. Automatic steering of farm vehicles using GPS. In *IN 3RD INTL. CONF. PRECISION AGRICULTURE*, 1996. URL <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.115.3909>.

M Agrawal and K Konolige. Real-time Localization in Outdoor Environments using Stereo Vision and Inexpensive GPS. In *18th International Conference on Pattern Recognition (ICPR'06)*, volume 3, pages 1063–1068. IEEE, 2006. ISBN 0-7695-2521-0. doi: 10.1109/ICPR.2006.962. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1699709http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1699709.

Siddharth Garg and Diana Marculescu. Special session 4A: New topics parametric yield and reliability of 3D integrated circuits: New challenges and solutions. In *29th VLSI Test Symposium*, pages 99–99. IEEE, May 2011. ISBN 978-1-61284-657-6. doi: 10.1109/VTS.2011.5783764. URL https://sibdi.ucr.ac.cr/http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=5783764&contentType=Conference+Publications&searchField%3DSearch_All%26queryText%3D3D+integrated+circuitshhttp://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5783764.

Samuel J. Dickerson, Steven P. Levitan, and Donald M. Chiarulli. 3D integrated circuits for lab-on-chip applications. In *2009 IEEE International Conference on 3D System Integration*, pages 1–8. IEEE, September 2009. ISBN 978-1-4244-4511-0. doi: 10.1109/3DIC.2009.5306597. URL https://sibdi.ucr.ac.cr/http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=5306597&contentType=Conference+Publications&searchField%3DSearch_All%26queryText%3D3D+integrated+circuitshhttp://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5306597.

A Chacon-Rodriguez, F Martin-Pirchio, S. Sanudo, and P. Julian. A Low-Power Integrated Circuit for Interaural Time Delay Estimation Without Delay Lines. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 56(7):575–579, July 2009. ISSN 1549-7747. doi: 10.1109/TCSII.2009.2023281. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5075638http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5075638.

- R S Heffner, G Koay, and H E Heffner. Sound localization acuity and its relation to vision in large and small fruit-eating bats: II. Non-echolocating species, *Eidolon helvum* and *Cynopterus brachyotis*. *Hearing research*, 241(1-2):80–6, July 2008. ISSN 0378-5955. doi: 10.1016/j.heares.2008.05.001. URL <http://dx.doi.org/10.1016/j.heares.2008.05.001>.
- Nima Mesgarani, Stephen V David, Jonathan B Fritz, and Shihab A Shamma. Influence of context and behavior on stimulus reconstruction from neural activity in primary auditory cortex. *Journal of neurophysiology*, 102(6):3329–39, December 2009. ISSN 1522-1598. doi: 10.1152/jn.91128.2008. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2804432&tool=pmcentrez&rendertype=abstract>.
- G Evangelopoulos and A Zlatintsi. Video event detection and summarization using audio, visual and text saliency. *Speech and Signal*, (2):3553–3556, 2009. doi: 10.1109/ICASSP.2009.4960393. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4960393.
- Christoph Kayser, Christopher I Petkov, Michael Lippert, and Nikos K Logothetis. Mechanisms for allocating auditory attention: an auditory saliency map. *Current biology : CB*, 15(21):1943–7, November 2005. ISSN 0960-9822. doi: 10.1016/j.cub.2005.09.040. URL <http://dx.doi.org/10.1016/j.cub.2005.09.040>.
- Su Shujing. Research on passive acoustic localization based on array field. *2010 International Conference on Computer Application and System Modeling (IC-CASM 2010)*, (Iccasm):V13–481–V13–485, October 2010. doi: 10.1109/ICCASM.2010.5622863. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5622863>.
- AM Johansson. Implementation Strategies for Array Localisation Algorithms Based on the SRP-PHAT. URL <http://www.fishdsp.com/research/rvk2008.pdf>.
- Jacob Benesty, Chen Jingdong, and Huang Yiteng. *Microphone Array Signal Processing*, volume 1 of *Springer Topics in Signal Processing*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN 978-3-540-78611-5. doi: 10.1007/978-3-540-78612-2. URL <http://www.springerlink.com/index/10.1007/978-3-540-78612-2>.
- Jingjing Jia, Mingjie Liu, and Xiaofeng Li. Acoustic Localization Algorithm Using Wireless Sensor Networks. *2009 Second International Conference on Intelligent Computation Technology and Automation*, pages 434–437, 2009. doi: 10.1109/ICICTA.2009.340. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5287822>.

- ST Birchfield and Rajitha Gangishetty. Acoustic Localization By Interaural Level Difference. In *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, volume 4, pages 1109–1112. IEEE, 2005. ISBN 0-7803-8874-7. doi: 10.1109/ICASSP.2005.1416207. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1416207http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1416207.
- Brian G. Ferguson, Lionel G. Criswick, and Kam W. Lo. Locating far-field impulsive sound sources in air by triangulation. *The Journal of the Acoustical Society of America*, 111(1):104, 2002. ISSN 00014966. doi: 10.1121/1.1402618. URL <http://link.aip.org/link/JASMAN/v111/i1/p104/s1&Agg=doi>.
- A. Chacon-Rodriguez, P. Julian, and F. Masson. Fast and low power integrated circuit for impulsive sound localisation using Kalman filter approach. *Electronics Letters*, 46(7):533, 2010. ISSN 00135194. doi: 10.1049/el.2010.2669. URL <http://link.aip.org/link/ELLEAK/v46/i7/p533/s1&Agg=doi>.
- L Li and G E Caldwell. Coefficient of cross correlation and the time domain correspondence. *Journal of electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology*, 9(6):385–9, December 1999. ISSN 1050-6411. URL <http://www.ncbi.nlm.nih.gov/pubmed/10597051>.
- JH DiBiase. *A high-accuracy, low-latency technique for talker localization in reverberant environments using microphone arrays*. Doctoral thesis, Brown University, 2000. URL <http://www.glat.info/ma/av16.3/2000-DiBiaseThesis.pdf>.
- Ajay Tirumala. Lightning: A Fast and Lightweight Acoustic Localization Protocol Using Low-End Wireless Micro-Sensors. In *25th IEEE International Real-Time Systems Symposium*, pages 371–381. IEEE, 2004. ISBN 0-7695-2247-5. doi: 10.1109/REAL.2004.29. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1381323http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1381323.
- Anand Ramamurthy. *Experimental Evaluation of Modified Phase Transform for sound source detection*. Masters, University of Kentucky, 2007. URL http://uknowledge.uky.edu/gradschool_theses/478/.
- Y Bahadirlar and H.O. Gulcur. Time-frequency cardiac passive acoustic localization. In *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1850–1853. IEEE, 2001. ISBN 0-7803-7211-5. doi: 10.1109/IEMBS.2001.1020583. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1020583http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1020583.

- Matthew Rhudy, Brian Bucci, Jeffrey Vipperman, Jeffrey Allanach, and Bruce Abraham. Microphone Array Analysis Methods Using Cross-Correlations. In *Volume 15: Sound, Vibration and Design*, pages 281–288. ASME, 2009. ISBN 978-0-7918-4388-8. doi: 10.1115/IMECE2009-10798. URL <http://link.aip.org/link/ASMECP/v2009/i43888/p281/s1&Agg=doi>.
- F. Valente and C. Wellekens. Variational Bayesian Feature Saliency for Audio Type Classification. *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, 1:513–516, 2005. doi: 10.1109/ICASSP.2005.1415163. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1415163>.
- HTH Do. *Real-time SRP-PHAT source location implementations on a large-aperture microphone array*. Masters, Brown University, 2009. URL http://www.lems.brown.edu/~hdo/mythesis_online.pdf.
- M Stanacevic and G. Cauwenberghs. Micropower gradient flow acoustic localizer. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 52(10): 2148–2157, October 2005. ISSN 1057-7122. doi: 10.1109/TCSI.2005.853356. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1519628http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1519628.
- Eric Ziegel, William Press, Brian Flannery, Saul Teukolsky, and William Vetterling. *Numerical Recipes: The Art of Scientific Computing*, volume 29. November 1987. ISBN 0521431085. doi: 10.2307/1269484. URL <http://www.jstor.org/stable/1269484?origin=crossref>.
- Yitzhak Katznelson. *An introduction to harmonic analysis*. Stanford University, California, third edition, 2004. URL http://books.google.com/books?hl=en&lr=&id=gkpUE_m5vvsC&oi=fnd&pg=PA1&dq=An+Introduction+to+Harmonic+Analysis&ots=A4iBK3RFvz&sig=bDyCmdqmgk2F1B8YHxic1hRwx6I.
- P. Welch. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15(2):70–73, June 1967. ISSN 0018-9278. doi: 10.1109/TAU.1967.1161901. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1161901>.
- G. Carter and C. Knapp. The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24(4):320–327, August 1976. ISSN 0096-3518. doi: 10.1109/TASSP.1976.1162830. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1162830>.

- Z. S. Tang and S. H. Chang. A novel DOA estimation based on signal subspace structure of uniform linear array. In *2007 IEEE Antennas and Propagation International Symposium*, pages 5279–5282. IEEE, June 2007. ISBN 978-1-4244-0877-1. doi: 10.1109/APS.2007.4396738. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4396738>.
- Jacek Dmochowski, Jacob Benesty, and Sofine Affes. Direction of Arrival Estimation Using the Parameterized Spatial Correlation Matrix. *IEEE Transactions on Audio, Speech and Language Processing*, 15(4):1327–1339, May 2007a. ISSN 1558-7916. doi: 10.1109/TASL.2006.889795. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4156192>.
- H.F. Silverman, J.M. Sachar, and W.R. Patterson. Performance of real-time source-location estimators for a large-aperture microphone array. *IEEE Transactions on Speech and Audio Processing*, 13(4):593–606, July 2005. ISSN 1063-6676. doi: 10.1109/TSA.2005.848875. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1453602>.
- R. Zarifeh, N. Alinier, S. Krishnan, and A. Anpalagan. Interference Detection in Spread Spectrum Communication Using Polynomial Phase Transform. In *2007 IEEE International Conference on Communications*, pages 2979–2984. Ieee, June 2007. ISBN 1-4244-0353-7. doi: 10.1109/ICC.2007.495. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4289166>.
- Jacek Dmochowski, Jacob Benesty, and Sofiene Affes. Direction of Arrival Estimation using Eigenanalysis of the Parameterized Spatial Correlation Matrix. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, number X, pages I-1–I-4. IEEE, 2007b. ISBN 1-4244-0727-3. doi: 10.1109/ICASSP.2007.366601. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4217001>.
- Jacek Dmochowski, Jacob Benesty, and Sofine Affes. Direction of Arrival Estimation Using the Parameterized Spatial Correlation Matrix. *IEEE Transactions on Audio, Speech and Language Processing*, 15(4):1327–1339, May 2007c. ISSN 1558-7916. doi: 10.1109/TASL.2006.889795. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4156192>.
- Jacob Benesty. Robust time delay estimation exploiting redundancy among multiple microphones. *IEEE Transactions on Speech and Audio Processing*, 11(6):549–557, November 2003. ISSN 1063-6676. doi: 10.1109/TSA.2003.818025. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1255443>.

-
- J. Benesty, J. Chen, and Y. Huang. Time-Delay Estimation via Linear Interpolation and Cross Correlation. *IEEE Transactions on Speech and Audio Processing*, 12(5): 509–519, September 2004. ISSN 1063-6676. doi: 10.1109/TSA.2004.833008. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1323087>.