



VI Congreso Iberoamericano de Acústica - FIA 2008  
Buenos Aires, 5, 6 y 7 de noviembre de 2008

FIA2008-A146

## Localización y separación de fuentes de audio mediante el uso de un arreglo de micrófonos

Matías Chavanne <sup>(a)</sup>,  
Germán Martínez <sup>(a)</sup>,  
Álvaro Agustín Martínez Amezaga <sup>(a)</sup>,  
Mauro Suarez <sup>(a)</sup>,  
Pablo Enrique Wundes <sup>(a)</sup>.

(a) Grupo de Electronica Digital Aplicada (GEDA), Instituto Tecnologico de Buenos Aires, Av. Madero 399, Ciudad Autonoma de Buenos Aires, Argentina.

### Abstract

This document describes the design of a system that allows localizing and discerning audio sources in the same ambient, For this purpose, an evenly spaced microphone array was used. As it will be seen, the process can be divided into two separate fundamental pieces: detection and separation. For this, different techniques will be analyzed, in particular adaptive beamforming (for source separation) and MUSIC algorithm (for source localization). An important factor to take into account is that most algorithms implemented were modified and improved for this particular application, for instance the use of fractional delays. Based on the simulations run, the effectiveness of the system could be verified, meeting the expected results.

### Resumen

En este documento se describe el proceso realizado para llevar a cabo el diseño de un sistema que permita localizar y separar fuentes de sonido en un ambiente. Para dicho fin se empleó un arreglo de micrófonos equiespaciados. Como se verá a lo largo del informe, el proceso puede dividirse en dos procesos fundamentales: detección y posterior separación. Para ello se analizarán distintas técnicas como ser Beamforming adaptativo (para la separación) y la aplicación del algoritmo MUSIC para la localización de las fuentes. Un factor importante a tener en cuenta es que se han llevado a cabo modificaciones a los algoritmos estudiados, como ser la implementación de retardos fraccionarios. Al realizar las simulaciones pertinentes, se pudo verificar que el sistema funcionó, cumpliendo las expectativas planteadas.

## 1 Introducción

En este documento se abordará el tema de detección y separación de fuentes mediante un arreglo de micrófonos. Como bien se sabe, este tema es de mucho interés ya que si se lograra implementar de manera exitosa y con pocas restricciones, es decir, consiguiendo que funcione bajo cualquier tipo de entorno, se le podrían dar innumerables aplicaciones en diversos campos. Entre ellos se puede mencionar cámaras de video, robótica, sistemas de manos libres, etc.

Básicamente, la detección y separación consiste en que dado un conjunto de fuentes de sonido dentro de una sala, como ser personas hablando o altoparlantes, sea posible detectar el ángulo (medido con el arreglo de micrófonos y utilizando este último como sistema de referencia) del cual provienen cada una de dichas fuentes. Una vez logrado esto, se procede a separar o individualizar la información emitida por cada fuente.

Para obtener los algoritmos óptimos, se deberá primero fijar los objetivos a alcanzar y luego, los pasos a seguir. Se concluirá que la investigación ha sido exitosa cuando se logre detectar la máxima cantidad de fuentes admitida por el sistema (que dependerá de la cantidad de micrófonos que posea el arreglo) estando separadas entre sí no menos de  $20^\circ$  y la información recuperada sea inteligible para el oído humano. Esto último quiere decir que no se esperará obtener la señal producida por cada fuente libre de interferencias, sino que bastará con que la señal original sea considerablemente mayor al ruido (interferencia de las otras fuentes).

Para la detección de fuentes se optó por el algoritmo MUSIC (*Multiple Signal Classification*), el cual trabajando con la matriz de covarianza de las señales captadas por los micrófonos del arreglo, calcula los ángulos de donde provienen.

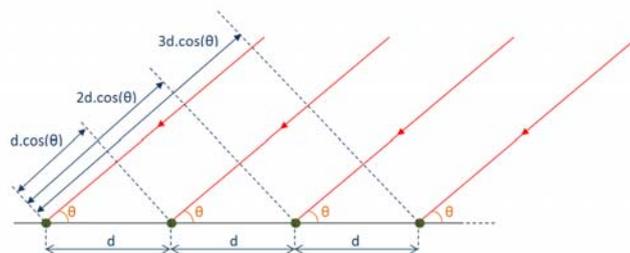
La separación de fuentes se llevará a cabo mediante la técnica de *beamforming*. Como se analizará a lo largo del informe, las técnicas convencionales de *beamforming* no son eficientes; por lo cual es necesaria la implementación de un algoritmo adaptativo.

## 2 Algoritmos de DOA

Se han desarrollado múltiples enfoques para estimar de la DOA de múltiples señales, los cuales se pueden separar en dos grandes grupos: Los llamados *métodos del subespacio* y los *frequency wavenumber spectrum*. Dentro del primero, los algoritmos que sobresalen son el *Multiple Signal Classification* o MUSIC [1] y el *Estimation of signal parameters via rotational invariance techniques* o ESPRIT [2]. Dentro del segundo se puede destacar el *Maximum Likelihood* o ML [3]. Mientras que en el segundo grupo, los parámetros de localización son estimados buscando picos, el primer grupo utiliza las propiedades de los subespacios para extraer toda la información necesaria de la matriz de covarianza de la señal recibida por los micrófonos.

### 2.1 Arreglo de micrófonos

Suponga que a un arreglo de micrófonos equiespaciados arriba una señal de frecuencia  $f$  y que debido a la distancia a la cual se encuentra la fuente puede considerarse un frente de onda plano. Cada micrófono captará distintas versiones de cada una de las señales incidentes  $s_K$ , es decir con un cierto atraso o adelanto. Es por esto que se requiere poner un punto de referencia dentro de nuestro sistema con el fin de poder referenciar estos desfasajes.



**Figura 1.** Arreglo de micrófonos.

Si se trata a la señal captada por cada uno de los micrófonos como compleja y de banda angosta con frecuencia central  $f$ , entonces, esta puede ser escrita como un cambio de fase de la señal incidente, resultando ser

$$x_i(t) = s(t)e^{-j2\pi ft_i} = s(t)e^{-j2\pi fd_i \cos\theta/v} \quad (1)$$

Para poder utilizar esta última ecuación se debe aplicar la Transformada de Gilbert a cada una de las salidas de los micrófonos de manera de hacerlas analíticas. Por otra parte si se tienen  $K$  fuentes emisoras de sonido, la señal recibida por cada micrófono  $m_i$  será una combinación lineal de las señales emitidas por las fuentes y del ruido que se encuentre en el sistema en ese instante.

$$x_i(t) = \sum_{k=0}^K s_k(t) e^{-j2\pi fd_i \cos\theta_k/v} \quad (2)$$

Es de vital importancia comentar que cuando se muestrea espacialmente una señal con un arreglo de micrófonos equiespaciados una distancia  $d$ , la frecuencia angular espacial de Nyquist corresponde a  $K_N = \pi/d$  con longitud de periodo igual a  $2d$  [4]. De esta forma, las ondas planas con longitud de onda menores a  $\lambda_{min} = 2d$  no podrán ser inequívocamente reconstruidas a partir de las muestras espaciales. Implicando que la mayor frecuencia  $f_{max}$  con suficiente densidad de muestreo sea  $f_{max} = v/(2d)$ . En general, las ondas planas con frecuencias angulares  $K \pm pK_s$ ,  $p=1,2,\dots$ , no podrán ser distinguidas cuando sean muestreadas con frecuencia angular espacial  $K_s = 2K_N$ , y consecuentemente contribuirán cuando se estime de alguna otra componente en el campo de sonido espacialmente muestreado.

Por lo tanto para no tener alias temporal, el cual no puede ser quitado por un filtro, se deben colocar los micrófonos a una distancia tal que se cumpla que

$$d \leq \frac{v}{2f_{max}} \quad (3)$$

## 2.2 Estimación de la cantidad de fuentes

Para poder estimar la cantidad de fuentes existen varios métodos, los más utilizados son el *Akaike's Information Criterion* (AIC) y el *Minimum Description Length Criterion* (MDL) [5]. Estos métodos se basan en encontrar el  $k$  que minimice la ecuación que los describe.

En la práctica el estimador  $\hat{K}_{AIC}$  no presentó muy buenos resultados mientras que el estimador  $\hat{K}_{MDL}$  arrojaba el resultado correcto en la gran mayoría de los experimentos

realizados. Cabe destacar que estos estimadores fueron desarrollados para otras aplicaciones y no específicamente para el algoritmo MUSIC. Por esto se decidió experimentar con estimadores alternativos hasta hallar uno que funcionase correctamente en nuestro caso. Este estimador resultó funcionar bien en todos los casos analizados en banda angosta.

### 2.3 Utilización del algoritmo en señales de banda ancha

Debido a que el algoritmo MUSIC solo es aplicable para una señal de banda angosta ya que la longitud de onda de la señal debe ser fija para poder realizar la estimación de la cantidad de fuentes. Si se contempla que la longitud de onda puede variar por un 4 % respecto a un valor central, se observó experimentalmente que el error producido no supera los 2°. Por lo tanto se puede extender la utilización del algoritmo a señales de banda ancha mediante el uso de filtros pasa-banda.

Bajo la suposición de una variación del 4% en la longitud de onda se generaron filtros FIR pasa-banda comprendidos dentro de la banda 400 Hz – 3 kHz. Estos filtros van aumentando su BW a medida que la frecuencia central de los mismos aumenta. La Figura 1.4 muestra la respuesta en frecuencia de los filtros antes mencionados. Las frecuencias centrales de los filtros fueron fijadas en 416 Hz, 515 Hz, 640 Hz, 793 Hz, 983 Hz, 1220 Hz, 1512 Hz, 1875 Hz, 2325 Hz y 2883 Hz.

### 2.4 Simulaciones

En primer lugar se procedió a testear el funcionamiento del algoritmo implementado para señales senoidales puras para diferentes relaciones señal-ruido. En principio se evaluó el error cometido por el algoritmo para detectar una única fuente cuya frecuencia coincidiese con la longitud de onda de referencia. En la tabla 1.1 se muestran los resultados obtenidos de la simulación realizada con una fuente ubicada a una distancia de 10m emitiendo una señal senoidal de 500Hz y 4 micrófonos separados 5cm entre sí. Como puede observarse en la Tabla 1, el algoritmo funciona aun con condiciones extremas como es el caso de -5dB de SNR. El máximo error detectado para una relación señal ruido superior a 10dB ronda los 0,45° lo que resulta despreciable para la aplicación en cuestión.

**Tabla 1.** Eficiencia del algoritmo de separación con distintos SNR

Angulo de incidencia [°]	SNR=-5dB	SNR=0dB	SNR=10dB	SNR=20dB
10	9,35	10,35	9,95	10,05
30	30,45	30,2	30,2	30,2
40	40,2	40,35	40,3	40,25
50	50,3	50,2	50,3	50,35
70	70,2	70,5	70,35	70,4
90	90,15	90,5	90,45	90,4
100	100,45	100,5	100,4	100,4
140	140,95	140,25	140,25	140,25
160	159,6	160,05	160,15	160,15

En la Tabla 2 se muestra la performance del algoritmo para detectar el ángulo de arribo de dos fuentes para diferentes ángulos de las mismas. Estas mediciones fueron realizadas con un arreglo de 4 micrófonos separados 5cm y dos fuentes ubicadas a 10m emitiendo con igual intensidad y una SNR de 40dB. Las frecuencias de las mismas fueron 495 Hz y 505 Hz respectivamente mientras que la frecuencia central utilizada por el algoritmo fue de 500 Hz.

**Tabla 2.** Eficiencia del algoritmo de separación con dos fuentes

Ángulo de incidencia [°]		Ángulo detectado [°]	
fFuente 1	fFuente 2	fFuente 1	fFuente 2
10	50	12,25	49,65
20	40	21,55	38,8
45	135	45,85	135,9
70	100	70,6	100,55
90	110	90,45	110,65
20	170	21,65	173,15
50	65	50,6	64,8
150	170	150	173,25
0	180	8,15	180

Como se puede ver, el error cometido es aceptable exceptuando en casos críticos como el último caso en el que las fuentes se ubican en la zona donde el arreglo de micrófonos posee menor sensibilidad. También se puede observar que para ángulos cercanos a los extremos el algoritmo comete un error mayor que para ángulos cercanos a noventa grados así como también que la SNR juega un papel fundamental cuando se trata de distinguir dos o más fuentes. Cabe la pena aclarar que con una relación señal ruido menor a 40dB y fuentes cercanas (20° o menos), el estimador desarrollado por nosotros puede fallar, pero, en ese caso también lo harán el AIC y MDL por un margen aun mayor. Se puede concluir que la mínima separación angular promedio distinguible para un arreglo de cuatro micrófonos ronda los 20°. Por otra parte si se utiliza una mayor cantidad de micrófonos se podrán distinguir fuentes más cercanas con la desventaja de un incremento en los tiempos de procesamiento.

### 3 Algoritmo de separación

#### 3.1 Introducción

La separación de fuentes, como su nombre lo indica, busca distinguir de manera precisa y eficiente, la información proveniente de distintas fuentes que llegan en forma simultánea a un arreglo de micrófonos. Por preciso y eficiente, se entiende que no se anulará completamente la interferencia causada por las fuentes restantes, sino que se buscará obtener una señal inteligible con un ruido aceptable. En otras palabras, se buscará maximizar la señal proveniente de la fuente de sonido requerida, minimizando los efectos de las señales no deseadas (interferencia).

#### 3.2 Beamforming convencional

La técnica de *beamforming* convencional consiste en retrasar y sumar las señales. Debido a que existe cierta separación física entre los micrófonos del arreglo, la señal proveniente de cada fuente de sonido, llega en diferentes instantes de tiempo a cada uno de ellos. Por lo tanto, se deberá realizar una compensación que permita contemplar dichos retardos.

De las simulaciones realizadas utilizando este método se pudo concluir que el simple hecho de retrasar y promediar las señales que arriban a cada micrófono no es suficiente como para poder discernir entre dos señales provenientes de dos fuentes de sonido. Por tal motivo,

resulta evidente la necesidad de implementar un algoritmo que permita minimizar los efectos de las fuentes de sonido que no sean las deseadas.

### 3.3 Beamforming adaptativo

La técnica de *beamforming* adaptativo consiste básicamente en ponderar las señales que inciden en cada uno de los micrófonos del arreglo, con el objeto de que el mismo sólo reciba señales provenientes de la dirección deseada, minimizando los efectos causados por la interferencia creada por el resto de fuentes de sonido existentes [6] [7].

La salida del *beamformer* se puede expresar como

$$y_k(\theta) = \bar{w}^H \bar{x}(k), \quad (4)$$

dónde  $x$  es la señal que arriba a cada micrófono y  $w$  es el vector de pesos, que representa la ponderación de cada una de dichas señales.

#### 3.3.1 Algoritmo LMS (Least Square Mean)

El algoritmo de mínimo error cuadrático (LMS) [8] es un algoritmo basado en un filtro adaptativo lineal, el cual consta de dos procesos básicos: Primero un proceso de filtrado, el cual involucra el cálculo de la salida de un filtro transversal producido por un conjunto de entradas retrasadas (*taps*), y la generación de una señal error a partir de la comparación de dicha salida con la señal deseada. Luego, un proceso adaptativo, el cual involucra un el ajuste automático de los pesos del filtro, de acuerdo a la estimación del error.

### 3.4 Implementación

El algoritmo de separación se puede dividir en dos bloques fundamentales: una etapa de *beamforming* convencional, y otra etapa de procesamiento adaptativo de la señal deseada. A continuación se detallará el funcionamiento de cada uno de dichos bloques.

Al analizar cada uno de los algoritmos adaptativos, se pudo comprobar que el RLS es alrededor de diez veces más rápido que el LMS, aunque posee una complejidad computacional mucho mayor. Al realizar las simulaciones pertinentes, se pudo comprobar que la complejidad del RLS hace que el algoritmo no llegue a converger en la mayoría de las pruebas, ya que los cálculos del mismo y la cantidad de iteraciones realizadas provocaron el desborde de la memoria disponible en el sistema operativo. Comparando el CMA con el LMS, se concluye que debido a que la función de costo involucra momentos de cuarto orden, el gradiente es muy pronunciado. Esto hace que diste mucho del óptimo en comparación con el LMS, así como también que el  $\mu$  máximo que garantice estabilidad sea más pequeño. Por lo tanto, el algoritmo LMS resulta ser más preciso que el de módulo constante (CM).

De esta manera, el algoritmo escogido para el bloque adaptativo resultó ser el LMS. El diagrama en bloques de la implementación realizada se puede apreciar en la Figura 2.

La rama superior del circuito anterior, es la encargada de generar la señal de *beamforming* convencional. Por otro lado, la rama inferior consiste básicamente en un “cancelador de lóbulo lateral” (GSC) adaptativo. El objetivo de la rama inferior es generar la señal que permitirá la cancelación del lóbulo lateral. La idea es que dicha señal sea lo más parecido a la señal interferencia o ruido existente en la señal a la salida del *beamformer* convencional. De esta manera, al realizar la resta entre esta última señal y la señal generada por el bloque adaptativo, se obtendrá una buena estimación de la señal deseada. Obviamente dicha estimación será más cercana a la deseada cuando el error existente entre dicha señal y la

salida del bloque adaptativo, sea mínimo. La idea de la matriz B es bloquear la señal deseada de la rama inferior, es decir, tratar de que en dicha rama exista la menor cantidad de componentes de la señal deseada.

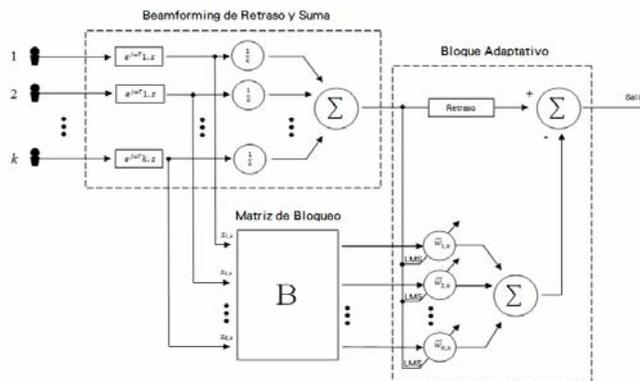


Figura 2. Esquema del beamformer adaptativo.

Una vez que la señal atraviesa la matriz “bloqueadora”, se ingresa al bloque adaptativo. Dicho bloque implementa un algoritmo LMS, utilizando filtros FIR, cuyos coeficientes (pesos) son variables. Dicho bloque toma la referencia de la salida del beamformer convencional, y su entrada es la salida de la matriz, la cual da una idea de la señal interferencia. El filtro LMS adaptativo minimizará el error entre su salida y la señal de referencia, a partir de su señal de entrada. De esta manera, cuánto más parecidas sean dichas señales, significará que la señal de salida del bloque adaptativo posee una señal de interferencia muy similar a la interferencia que existe a la salida del beamformer convencional. Por lo tanto, al restar dicha señal con la salida del filtro adaptativo, se obtendrá una estimación de la señal deseada libre de interferencia.

Al evaluar la performance del beamformer frente a las simulaciones, se notó que se lograba una considerable mejora al incrementar le frecuencia de muestreo de la señal capturada por los micrófonos. El diagrama polar de captura del beamformer reveló que esta mejora, apreciable como un estrechamiento en el lóbulo principal como se aprecia en la Figura 3, se debe principalmente a el hecho que los retardos entre las señales de los distintos micrófonos pueden ser implementados con mucha mayor precisión, ya que una muestra representa una unidad de tiempo considerablemente menor.

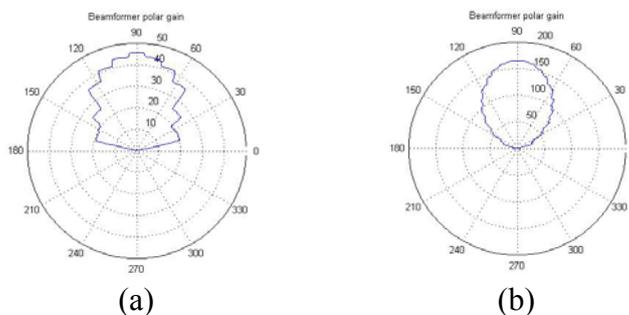


Figura 3. Selectividad del beamformer a (a)  $f_s = 48$  kHz y (b)  $f_s = 192$  kHz

Se puede observar que a medida que se incrementa el valor de la frecuencia de muestreo, el valor del retardo se vuelve más grande. Si se considera que cuanto más grande es el retardo con mejor precisión se van a poder separar dos casos en donde solamente difiera el  $\theta$ , una mayor frecuencia de muestreo es una gran ventaja. Sin embargo, es una desventaja al momento de implementar el algoritmo adaptativo (LMS), ya que implica vectores y filtros de mucho mayor orden sin ofrecer una ventaja significativa. Por esto se busco implementar retardos fraccionarios.

La solución más sencilla para implementar dichos retardos fue usando un filtro interpolador para lograr un retardo fraccionario sin la necesidad de usar estructuras multibanda (con varias frecuencias de muestreo. Al usar este modelo, los resultados fueron aceptables, sin embargo, al incrementar la complejidad de la señal de entrada (al usar señales reales de audio) muchas veces el filtro no proveía el retardo solicitado correctamente. Esto estaba dado por el hecho de que, el orden del filtro  $N$  debe estar cerca de  $2 \cdot D$  (el retardo), pero para simplificar las operaciones todos los retardos de los micrófonos fueron implementados con filtros del mismo orden, para obtener retardos de grupo comparables en todos los casos. Este factor, sumado a la respuesta en frecuencia no ideal del filtro de Lagrange reveló que la implementación de sistemas polifásico es una alternativa más atractiva.

Para el sistema polifásico, se utilizó un filtro interpolador (compuesto de un bloque intercalador de ceros y un filtro pasabajos FIR en  $\pi/l$ , siendo  $l$  el factor de sobremuestreo [9]), un retardo entero y un filtro decimador (con un pasabajos similar al del filtro interpolador, pero en alta frecuencia [10]), representando de esta forma el sistema polifásico más sencillo. Cabe aclarar que este sistema es válido debido a que las entradas son todas limitadas en banda, y las operaciones que se ejecutan en alta frecuencia no alteran la característica espectral de dichas señales, por lo que se puede ejecutar la decimación a la frecuencia de muestreo original sin incurrir en la aparición de alias ni pérdida de datos en el filtro decimador debido a componentes de alta frecuencia, un problema común en los sistemas polifásicos.

Al implementar este sistema, se debió elegir una relación entre la frecuencia de muestreo original optima para el algoritmo adaptativo, y el sobremuestreo a utilizar para lograr un buen resultado del *beamformer* tradicional. Esta relación se encontró óptima para el caso de una frecuencia de muestreo de 16 kHz, con un *upsampling* de 50x o 100x. Este último factor es enteramente dependiente de las capacidades del procesador donde se realicen las operaciones.

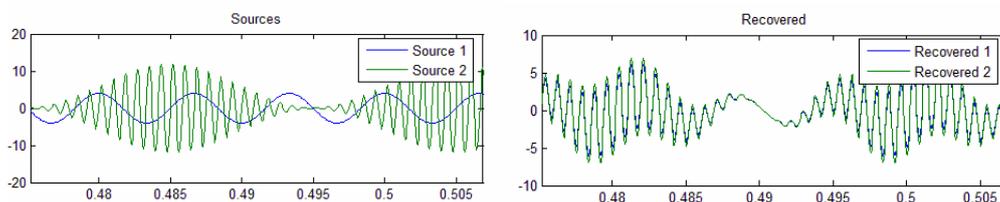
### 3.5 Simulaciones

Para verificar la performance del algoritmo de separación desarrollado, se procedió a realizar simulaciones del mismo. Debido a que las señales de audio son complejas, se comenzó realizando simulaciones sencillas, es decir, fuentes representadas por señales sinusoidales (tonos) o señales del tipo AM (amplitud modulada).

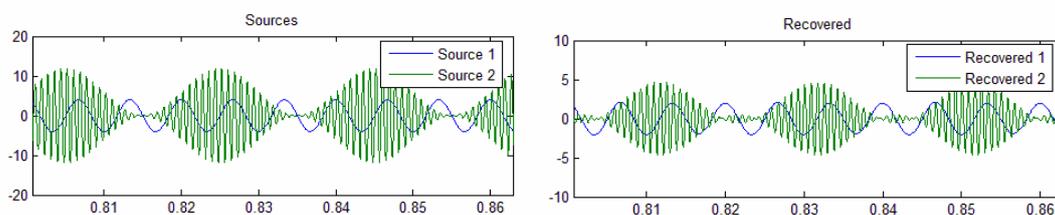
Cabe señalar que la frecuencia de muestreo utilizada fue de 16kHz. Además, cabe destacar que se utilizó un *oversampling* de 100x, es decir, retardos fraccionarios de 1/100. Además, la distancia entre micrófonos del arreglo es de 5cm. En las figuras 4 y 5 se pueden observar los resultados de estas simulaciones con dos fuentes de sonido (una en el ángulo  $70^\circ$  y otra en  $110^\circ$ ) y un arreglo de tres micrófonos.

En la figura 4 se puede observar que la señal no puede recuperarse correctamente, mientras que cuando se utilizó el algoritmo adaptativo (figura 5), dichas señales fueron separadas sin distorsión apreciable. Además se pudo constatar que la simulación realizada utilizando un arreglo de siete micrófonos resultó mucho menos eficiente que la realizada con el algoritmo adaptativo y un arreglo de tres micrófonos.

Por lo tanto se concluye que el bloque adaptativo implementado con el algoritmo LMS, incrementa notablemente la eficiencia del sistema.



**Figura 4.** Separación convencional



**Figura 5.** Separación adaptativa

### 3.5.1 Simulaciones con audio

Para verificar la eficiencia en señales reales se utilizaron fragmentos de audio grabados en condiciones de estudio, los cuales se posicionaron virtualmente en dos fuentes distribuidas en un plano y al cual se le agregó un ruido blanco (SNR aproximadamente  $-35\text{dB}$ ) en los micrófonos para simular condiciones reales de captura. En el sistema de simulación se consideró la propagación de ondas de sonido de manera realista y se tomó en cuenta el decaimiento de la intensidad sonora directamente proporcional a la distancia de las fuentes; esto permite una simulación de la captura de audio por micrófonos reales. Este audio luego se paso por la etapa de detección y, con estos resultados, se simulo el *beamformer*.

Para esta simulación, las fuentes se encontraban a 2,2 m. de un arreglo de cinco micrófonos omnidireccionales separados 5 cm cada uno; las fuentes se ubicaron a  $22,5^\circ$  de la recta normal del arreglo de micrófonos (separadas  $45^\circ$  entre ellas). Esta disposición intenta recrear una condición realista de dos interlocutores que pueden superponerse, una aplicación donde sería útil la detección y separación de las fuentes. La señal de audio fue capturada por cada micrófono a 16 kHz, y los retardos fraccionarios se hicieron con un *oversampling* de 100x.

El algoritmo de detección, en las condiciones descritas anteriormente, resulto sumamente efectivo, con un error menor al 0.5% en la detección de los ángulos de las fuentes. Con estos resultados se alimentó el bloque de *beamformer*.

Al escuchar las señales recuperadas, el audio de cada una de las fuentes es muy inteligible, a diferencia de la señal capturada en los micrófonos, donde la mezcla de ambas fuentes con prácticamente la misma intensidad hace que sea muy poco inteligible. La fuente a la cual no apunta el *beamformer* se encuentra presente en el audio recuperado, pero la atenuación es muy notoria y dicha fuente es muy complicada de discernir del ruido de fondo presente, y no molesta a la inteligibilidad de la fuente deseada en la mayoría de los casos. Una

excepción a ello es cuando hay un silencio prolongado en una de las fuentes, en cuyo caso la fuente no deseada se destaca más por el simple hecho que no hay una fuente de mayor intensidad. Este *crosstalk* en general no afecta la inteligibilidad del audio para un oyente humano, y esta condición en general mejora al incrementarse la separación de las fuentes o al agregar micrófonos al arreglo.

Como caso de interés, se investigó la posibilidad de separar dos fuentes no equidistantes al arreglo, en particular con una fuente cuya intensidad sea considerablemente menor a la otra, de forma que la mezcla directa de fuentes en los micrófonos del arreglo hagan casi imposible su comprensión. Para esto, se conservaron las condiciones de la simulación anterior pero se dispuso de las fuentes una más lejana que la otra. Los resultados revelaron que, si bien la señal más lejana se vio perjudicada, la inteligibilidad del audio recuperado continuó siendo aceptable.

### 3.5.2 Resultados

Analizando objetivamente los resultados obtenidos, se puede concluir que los objetivos fijados en un principio, fueron alcanzados satisfactoriamente.

Sin embargo, cabe destacar que la performance obtenida por nuestro sistema puede ser mejorada sustancialmente si se aplican nuevas técnicas que actúen directamente sobre las falencias y limitaciones del mismo.

Para perfeccionar el algoritmo de separación de fuentes se debe hacer hincapié en su principal limitación, la cual reside en su incapacidad para distinguir fuentes que se encuentren separadas menos de  $20^\circ$ . Por este motivo, se cree que la investigación debería orientarse a mejorar la selectividad del haz formado por el *beamforming*. Una posible implementación que podría ser analizada en un futuro, sería aumentar las dimensiones del arreglo de micrófonos. Se cree que esto, en principio, resultaría en una notable mejora del *beamforming* y, como ventaja adicional, también ayudaría a optimizar la detección.

### Referencias

- [1] Ralph O.Schmidt, (1986), "Multiple Emitter Location And Signal Parameter Estimation" .
- [2] Richard Roy and Thomas Kailath, (1989), "ESPRIT - Estimation Of Signal Parameters Via Rotational Invariance Techniques".
- [3] Ilan Ziskind and Mati Wax, (1988), "Maximum Likelihood Localization Of Multiple Sources By Alternating Projection".
- [4] R. Alzate, G. Castellanos, G. Acevedo, M. Sanchez, (2004), "Precisión En Medida De Localizacion Para Una Fuente Acústica Utilizando Tdoa".
- [5] Nahdi Nezefat, Mostafa Kaveh and Wenyuan Xu, (2004), "Estimation Of The Number Of Sources Based On The Eigenvectors Of The Covariance Matrix".
- [6] Jian Li and Petre Stoica, (2006), "Robust Adaptive Beamforming" (John Proakis, Series Editor). Editorial Wiley Interscience. New Jersey. EEUU
- [7] Jeffrey Connor, (2005), "A Study of Despread-Respread Multitarget Adaptive Algorithms in AWN Channel", Thesis submitted to the Department of Electrical and Computer Engineering for the Degree of Master of Science.
- [8] Simon Haykin, (1996), "Adaptive Filter Theory". Editorial Prentice Hall. Tercera Edición. EEUU.
- [9] The MathWorks Inc., "mfilt.firinterp - FIR filter-based interpolator", Filter Design Toolbox.
- [10] The MathWorks Inc., "mfilt.firdecim - Direct-form FIR polyphase decimator", Filter Design Toolbox.