

Aplicación de un arreglo adaptativo de micrófonos para el realce de voz

Adaptive microphone array for speech enhancement

William Andrés Cadavid*, Jhon Alexander Peñafiel*, Jorge Iván Marín**

Recibido: Febrero 2 de 2009

Aceptado: Junio 1 de 2009

Correspondencia: Programa de Ingeniería Electrónica, Universidad del Quindío, Avenida Bolívar Calle 12 norte Armenia Quindío. Email: gdsproc@uniquindio.edu.co

RESUMEN

En este trabajo se describe el diseño de un arreglo adaptativo de micrófonos para el realce de voz en un ambiente tipo teleconferencia usando filtrado espacial. Se muestran las simulaciones del sistema en MATLAB cuyos resultados se usaron para determinar los algoritmos a utilizar en la implementación en tiempo real, asimismo se describe en detalle el esquema de la implementación, las optimizaciones necesarias y las pruebas de desempeño del sistema implementado. Este último fue realizado sobre un procesador digital de señales TMS320C6701 de Texas Instruments al que se le conectó un arreglo de 2 sensores. Las señales recogidas por los sensores se pasan por un ADC que digitaliza las señales que posteriormente son procesadas para: calcular el número de fuentes presentes en el entorno usando el criterio MDL; calcular la dirección de llegada con el algoritmo RootMUSIC y finalmente enfocar el arreglo hacia la dirección desde donde se origina la señal deseada.

Palabras Clave: Formación de haz adaptativa, Arreglo de micrófonos, MDL, MUSIC, RootMUSIC

ABSTRACT

This paper describes the design of an adaptive microphone array for speech enhancement in teleconference environments by using spatial filtering. Matlab simulations of the system allowed determining the most suitable algorithms for the final real-time implementation. This implementation was made on a Texas Instruments digital signal processor TMS320C6701 attached to a sensor array with two elements. Implementation details, optimizations and performance tests of the implemented system are presented. The signals captured by the sensors are thereafter processed by the following algorithms: computation of the number of sources present in the environment using the MDL criterion; computation of the direction of arrival by means of a RootMUSIC algorithm, and finally, the beamforming toward the source of the desired signal.

Keywords: Adaptive Beamforming, Microphone Array, MDL, MUSIC, RootMUSIC

INTRODUCCIÓN

Para realzar la voz de un hablante en un ambiente tipo conferencia se plantean típicamente soluciones que consideran únicamente las características físicas del sensor, por ejemplo, el ser un micrófono alámbrico o inalámbrico o su directividad, y su disposición, como es el caso de los sistemas de retención en la cabeza más conocidos como manos libres. Pero una de las técnicas más recientes para el procesamiento de señales acústicas lo constituye los

arreglos de micrófonos (1), que asociados con el procesamiento digital de señales, conforman los sistemas multisensor inteligentes, que han probado tener mejor desempeño que aquellos que hacen uso de un único sensor (2). Los mayores avances en torno a los sistemas multisensor se encuentran en el campo de las telecomunicaciones, en el desarrollo de las denominadas antenas inteligentes (*Smart Antennas*).

La solución planteada en este trabajo para el realce de voz consiste en la implementación de un filtro espacial

*Joven investigador Grupo de Procesamiento Digital de Señales y Procesadores – GDSPROC. Universidad del Quindío.

** Profesor Programa Ingeniería Electrónica. Universidad del Quindío. e-mail: jorgemarin@uniquindio.edu.co

en el sistema de recepción. Aunque el filtrado espacial se puede realizar con un único sensor de alta directividad, la manera más adecuada de hacerlo es utilizando arreglos de sensores, ya que un único sensor requeriría de movimiento mecánico, sería incapaz de adaptarse a las interferencias, debido a que sus lóbulos laterales son fijos; no podría estar orientado a varios lugares al mismo tiempo y no sería posible conmutarlo a alta velocidad. Entre las aplicaciones que utilizan arreglos de sensores (3), (4), se encuentran: el radar, el sonar, la sismología, la biomedicina, las comunicaciones, la astronomía.

El procesamiento de arreglos se puede dividir de manera muy general en dos clases (5): la detección del número de fuentes; y la estimación, que permite aislar y analizar las señales producidas por cada una de las fuentes para finalmente llevar a cabo el proceso de formación de haz, más conocido como *beamforming*.

Es necesario mencionar las consideraciones y supuestos para el desarrollo de las técnicas y algoritmos del procesamiento de las señales acústicas en un arreglo de micrófonos: a) para las señales acústicas se debe considerar el campo cercano; b) los elementos del arreglo son omnidireccionales; c) el número de elementos del arreglo M , es mayor que el número de fuentes acústicas P (entorno subdeterminado); d) el modelo de propagación de señales es en banda angosta; e) las señales de fuentes diferentes no son coherentes (estadísticamente independientes); f) el ruido térmico en el sistema tiene distribución gaussiana, media cero y no está correlacionado con las fuentes (6). El ruido de fondo no está correlacionado espacial ni temporalmente con las señales de interés; g) la reverberación y el eco son débiles debido al bajo coeficiente de reflexión de los muros del entorno, además los multitrayectos son pocos y bien distanciados de la posición del hablante, lo que permite que este tipo de señales se pueda considerar inmerso en la componente de ruido (ruido no direccional).

Modelo de la señal

Dado que se utiliza un arreglo lineal uniforme ULA (por sus siglas en inglés de *Uniform Linear Array*), la señal que llega al arreglo de sensores (micrófonos) en un instante de tiempo se puede expresar como

$$\mathbf{x}(n) = [x_1(n) x_2(n) \dots x_M(n)]^T \quad (1)$$

donde M es el número de micrófonos (4). $\mathbf{x}(n)$ se conoce también como *snapshot*. Esta señal también se puede expresar como:

$$\mathbf{x}(n) = \sqrt{M} \sum_{p=1}^P \mathbf{v}(\phi_p) s_p(n) + \mathbf{w}(n) \quad (2)$$

con $s_p(n)$ la señal que llega al sensor referencia en el instante n , P el número de fuentes, $\mathbf{w}(n)$ representa las componentes de ruido (ruido de fondo y ruido térmico) y ϕ es la dirección de arribo de la señal. Por su parte, el término

$$\mathbf{v}(\phi_p) = \frac{1}{\sqrt{M}} [1 \quad e^{-j2\pi \frac{d \sin(\phi_p)}{\lambda}} \quad \dots \quad e^{-j2\pi \frac{d \sin(\phi_p)(M-1)}{\lambda}}]^T \quad (3)$$

se conoce como vector de respuesta del arreglo, o vector guía (4). En esta ecuación, d es la separación entre los micrófonos, cuyo valor para evitar el aliasing espacial debe garantizar la relación $d \leq \lambda / 2$, conocida como criterio de Nyquist (7), y es similar al tratado en el dominio de la frecuencia.

Formador de Haz (Beamformer)

El formador de haz es un procesador que se usa en conjunto con un arreglo de sensores para llevar a cabo una forma versátil de filtrado espacial (3). En la Fig. 1 se muestra la estructura de un formador de banda angosta, cuya estructura es similar a un filtro FIR. Los formadores de haz de banda ancha no son estudiados en este trabajo, pero una revisión detallada puede encontrarse en (3) y (8).

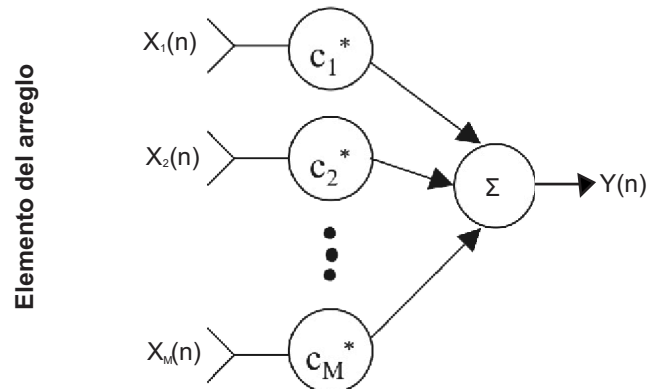


Figura 1: Formador de haz de banda angosta (narrow-band beamformer).

La salida del formador de haz de banda angosta está dada por

$$y(n) = \sum_{m=1}^M c_m^* x_m(n) \quad (4)$$

que se puede expresar en forma vectorial como

$$y(n) = \mathbf{c}^H(n) \mathbf{x}(n) \quad (5)$$

donde $\mathbf{c} = [c_1 \ c_2 \ \dots \ c_M]^T$ son los pesos. La cantidad

$$C(\phi) = |\mathbf{c}^H \mathbf{v}(\phi)| \quad (6)$$

se conoce como *patrón de haz* y no es más que la respuesta angular cuando se aplican los pesos a un conjunto de vectores de respuesta que barren todos los posibles ángulos entre -90° a 90° .

Clasificación de los Formadores de Haz

De acuerdo a la forma en que se seleccionen los pesos \mathbf{c} , la formación de haz se puede dividir en dos tipos: independiente de los datos y estadísticamente óptimo.

1) Formación de Haz Independiente de Datos: Se diseña para tener una respuesta deseada independiente de la estadística de los datos. Entre estos formadores se encuentran el formador clásico o convencional, también conocido como filtrado espacial adaptado (*matched*) o formador de retardo y suma DSBF (por sus siglas en inglés de *Delay and Sum Beam Former*), que se basa en la compensación de los retardos espaciales de la señal al llegar al arreglo. Lo anterior se expresa como $\mathbf{c}_m = \mathbf{v}(\phi_s)\phi_s$, donde ϕ_s es el ángulo al que se desea que apunte el arreglo. Otro tipo de formador independiente de los datos es el formador enventanado, en el que partiendo de la formación clásica, se aplica una función ventana al vector guía con el fin de reducir el nivel de los lóbulos laterales y/o controlar el ancho del haz (9).

2) Formación de Haz Estadísticamente Óptima: Los pesos se eligen para optimizar la salida del arreglo basándose en un conocimiento a priori de las estadísticas de los datos y en algún criterio de optimización. Generalmente se busca maximizar la SINR. Aunque existen diversas técnicas (8), en este trabajo sólo se consideró el filtro espacial óptimo que se busca mantener la señal deseada mientras se minimizan las contribuciones debidas a ruido e interferencia, es decir:

$$\min_{\mathbf{c}} \mathbf{c}^H \mathbf{R}_n \mathbf{c} \quad \text{restringido a} \quad \mathbf{v}^H(\phi) \mathbf{c} = 1 \quad (7)$$

donde \mathbf{R}_n es la matriz de covarianza correspondiente exclusivamente a la interferencia. La ecuación (7) se conoce como problema de mínima varianza de respuesta sin distorsión MVDR (por sus siglas en inglés de *Minimum Variance Distortionless Response*) cuya solución se encuentra utilizando el método de los multiplicadores de *Lagrange* (3) y está dada por

$$\mathbf{c} = \frac{\mathbf{R}_n^{-1} \mathbf{v}(\phi_s)}{\mathbf{v}(\phi_s)^H \mathbf{R}_n^{-1} \mathbf{v}(\phi_s)} \quad (8)$$

Teóricamente se busca reducir la varianza (o potencia) debida al ruido de interferencia, pero en la práctica no es posible estimar dicha matriz, así que se recurre a la matriz de autocorrelación de la señal captada por los micrófonos \mathbf{R}_x , que incluye tanto la señal deseada como las interferencias, es decir, $\mathbf{R}_x = \mathbf{R}_s + \mathbf{R}_n$, donde \mathbf{R}_s representa la matriz de autocorrelación de la señal deseada. Lo anterior se basa en la suposición que la señal y la interferencia son procesos estadísticamente independientes y de media cero. Por consiguiente, dado que minimizar la varianza o potencia debida al ruido de interferencia es equivalente a

minimizar la potencia de la señal de salida $\mathbf{c}^H \mathbf{R}_x \mathbf{c}$, expresión utilizada en la práctica para calcular los pesos del formador de haz, se obtiene al emplear \mathbf{R}_x en lugar de \mathbf{R}_n en la ecuación (8). Este último problema da origen a la formación adaptativa. En este trabajo se exploró únicamente la adaptación por bloques, que consiste en realizar un estimado de la matriz de correlación utilizando un bloque de N *snapshots*. Esta estimación se realiza con (3)

$$\hat{\mathbf{R}}_x = \frac{1}{N} \sum_{n=1}^N \mathbf{x}(n) \mathbf{x}^H(n) = \frac{1}{N} \mathbf{X} \mathbf{X}^H \quad (9)$$

siendo $\mathbf{X} = [\mathbf{x}(1) \mathbf{x}(2) \dots \mathbf{x}(N)]$

Estimación de la dirección de llegada (doa)

En la sección anterior se supuso que la dirección de llegada de la señal deseada ϕ_s es conocida, por tanto, se requieren técnicas que permitan estimarla a partir de los datos. La matriz de correlación para el modelo de señal se puede escribir como (6)

$$\mathbf{R}_x = E \{ \mathbf{x}(n) \mathbf{x}^H(n) \} \quad (10)$$

Además \mathbf{R}_x se puede escribir en términos de su autodescomposición EVD (por sus siglas en inglés *Eigenvalue Decomposition*) como (6):

$$\mathbf{R}_x = \sum_{m=1}^M \lambda_m \mathbf{q}_m \mathbf{q}_m^H = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^H \quad (11)$$

donde λ_m son los autovalores de \mathbf{R}_x en orden descendente $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M$ y \mathbf{q}_m son los correspondientes autovectores. es una matriz diagonal con los autovalores en orden descendente, mientras que las columnas de \mathbf{Q} son los correspondientes autovectores. La autocorrelación se puede dividir en dos componentes (6):

$$\mathbf{R}_x = \mathbf{Q}_s \mathbf{\Lambda} \mathbf{Q}_s^H + \sigma_w^2 \mathbf{Q}_w \mathbf{Q}_w^H \quad (12)$$

donde, $\mathbf{Q}_s = [q_1 \ q_2 \ \dots \ q_P]$ y $\mathbf{Q}_w = [q_{P+1} \ q_{P+2} \ \dots \ q_M]$ con P el número de fuentes. Esto muestra que el espacio M -dimensional se puede dividir en dos subespacios: subespacio de señal (\mathbf{Q}_s) y subespacio de ruido (\mathbf{Q}_w). Estos dos subespacios son ortogonales y se puede mostrar que los autovalores de ruido satisfacen $\mathbf{v}^H(\phi_p) \mathbf{q}_m = 0$ para $m > P$ (4). Esta división en subespacio da origen a diversas técnicas que utilizan, ya sea el subespacio de señal o el subespacio de ruido (6); en este documento sólo se describen las técnicas relacionadas con el subespacio de ruido, debido a la alta resolución que mostraron en las simulaciones. Los métodos basados en subespacio de ruido se basan en la propiedad que los vectores guía son ortogonales a cualquier combinación

lineal de los autovectores de ruido y pueden ser de dos tipos: Métodos de búsqueda y algebraicos.

1) Métodos basados en Búsqueda: Se basan en el cálculo del *pseudoespectro* cuyos picos máximos indican la ubicación espacial de las señales. El prefijo *pseudo* se debe a que no brinda información acerca de la potencia de la señal en las direcciones de los picos máximos. Entre estos métodos, uno de los más populares es el Algoritmo de Clasificación de Múltiple Señal MUSIC (por sus siglas en inglés *Múltiple Signal Classification*). Este método usa todos los autovectores del subespacio de ruido, realizando una sumatoria para tener un mejor refinamiento (4). Este método no es tan sensible a errores en la estimación de los autovectores de ruido. El *pseudoespectro* se puede calcular a través de (4):

$$S(\phi) = \left(\sum_{m=P+1}^M |v^H(\phi)q_m|^2 \right)^{-1} \quad (13)$$

2) Métodos Algebraicos: El procedimiento de búsqueda que se realiza en los métodos anteriores se puede remplazar por un algoritmo de solución de raíces. En la versión algebraica del algoritmo MUSIC válida para arreglos lineales uniformes, también llamada RootMUSIC, se genera el polinomio $R(z)$, correspondiente a la transformada z-espacial del denominador del *pseudoespectro* MUSIC. Este polinomio es de orden $2M-2$ y tiene $M-1$ pares de raíces, donde cada par posee una de las raíces dentro del círculo unidad y una fuera de él. Las P raíces más cercanas al círculo unitario corresponden a las exponenciales con los respectivos ángulos, de tal forma que la solución para las DOA está dada por:

$$\phi_i = \text{sen}^{-1} \left(\frac{\lambda}{2\pi d} \angle z_i \right) \quad \text{para } i=1\dots P.$$

DETERMINACIÓN DEL NÚMERO DE FUENTES

Para realizar una exitosa separación de los dos subespacios, es necesario conocer el número de fuentes P . Los $M-P$ autovalores más pequeños de \mathbf{R}_x deben ser en teoría iguales, debido a que corresponden a los autovalores de ruido. Por tanto, determinar cuántos de los autovalores son iguales, resuelve el problema de encontrar el número de fuentes. Para llevar a cabo esta tarea se utilizan los métodos AIC (Akaike's Information Criterion) y el MDL (*Minimum Description Length Criterion*). Ambos minimizan un criterio sobre el número de fuentes detectable. En este trabajo se exploró el método MDL que se encuentra descrito en (5).

METODOLOGÍA

En esta sección se indican las técnicas y algoritmos que fueron necesarios para la construcción del sistema, que basado en arreglo de micrófonos, permite el realce de la voz de un hablante en tiempo real.

HARDWARE

De acuerdo con el esquema de la Fig. 2, el sistema consta de un arreglo de micrófonos, y un DSP. La frecuencia de muestreo para el caso de señales de voz sin pérdida de información es 8kHz. El procesador digital de señales (DSP) con el que se realizan las pruebas es el TMS320C6701 que permite trabajar en punto flotante. Este DSP consta de una entrada de micrófono estéreo que puede partirse en dos entradas monofónicas con las que se realizan las pruebas. Para el propósito de implementación el número de micrófonos es $M=2$.

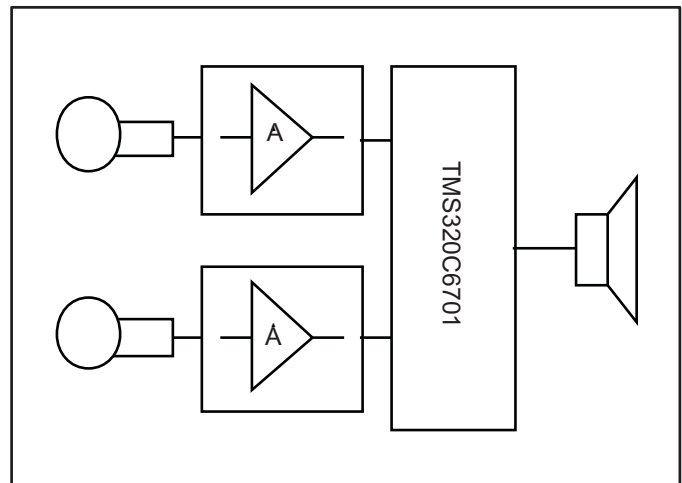


Figura 2: Diagrama de bloques del sistema.

SOFTWARE

El lenguaje utilizado para codificar los algoritmos es C y no ensamblador, debido en parte a la simplicidad y a que el compilador de C del entorno de desarrollo CodeComposer permite obtener un código altamente optimizado con un desempeño comparable al generado a partir de ensamblador.

En la Fig. 3 se muestran los diagramas de flujo para realizar la formación de haz clásica (o formación enventanada) y óptima. De acuerdo a la figura, el sistema funciona bajo la técnica de DMA (Acceso directo a memoria) de doble buffer para la adquisición de la señal. La Fig. 3b, muestra la rutina de servicio de interrupción asociada al fin de captura por DMA. En dicha rutina, cada vez que se lee un bloque, éste se procesa con los pesos actuales, y solo cuando la bandera Band es igual a cero se almacena el bloque de datos capturado en un arreglo temporal, que es usado por la rutina principal (main) para actualizar los pesos. En la rutina principal (Fig. 3a), se indica como la actualización de los pesos se realiza una vez haya un bloque listo (la bandera Band está en 1). Este esquema de codificar el sistema permite la ejecución en tiempo real, pues paralelo a la captura de los datos se realiza la actualización de los pesos.

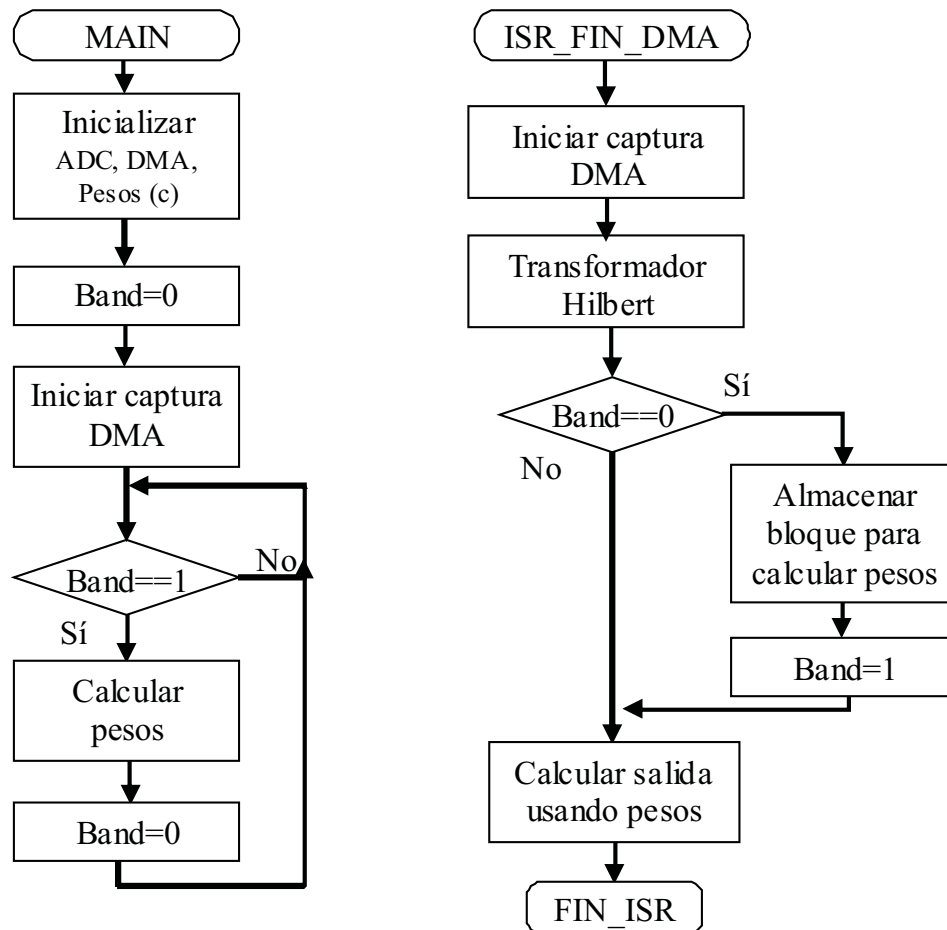


Figura 3: Diagrama de flujo. a)Programa Principal b)Rutina de Servicio de Interrupción por finalización de captura de un bloque.

La rutina cuyo tiempo de ejecución es crítico, es la función para el cálculo de la salida. En la rutina principal, el bloque denominado “calcular pesos” realiza el procedimiento de la Fig. 4. Los números que se muestran junto con los bloques en cada paso del cálculo de los pesos óptimos son los tiempos máximos en segundos que se tarda dicha rutina al ejecutarse en el DSP, para un bloque de 2000 muestras por cada canal. El tiempo máximo total es $T_t=79.1\text{ms}$. Como la captura de las 2000 muestras se tarda 250ms, es posible recalculer los pesos en cada bloque y por tanto se puede garantizar el funcionamiento del sistema en tiempo real.

El bloque NOS de la Fig. 4 calcula el número de fuentes utilizando el criterio MDL, mientras que el bloque FEST calcula los ángulos de llegada empleando el algoritmo Root

MUSIC, y a partir del ángulo de interés se genera el vector de respuesta en la dirección deseada con la función GENERAR V, que hace uso de los algoritmos indicados en la sección 2.3.

RESULTADOS Y DISCUSIÓN

Los resultados de simulaciones realizadas en *Matlab* permitieron escoger los parámetros más acordes para la implementación: 2000 *snapshots*, el criterio MDL, el algoritmo RootMUSIC. Para el caso de dos micrófonos ($M=2$) el número de fuentes máximo que puede detectar el algoritmo es $P=1$. En la tabla 1 se muestra el número de éxitos del criterio MDL para un conjunto de 100 pruebas utilizando en cada una 2000 *snapshots* y una fuente a diversas SNR.

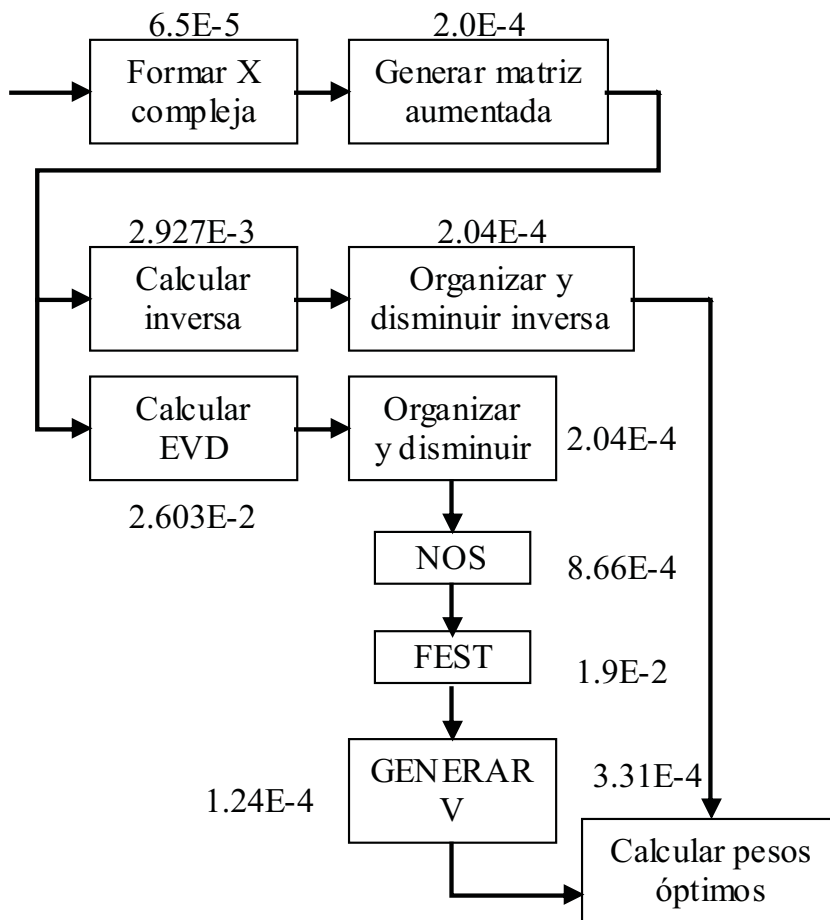


Figura 4: Cálculo en el DSP de los pesos óptimos.

Tabla 1: Resultados del MDL

SNR dB	Número de éxitos
0	64%
10	100%
20	100%

Los resultados de la implementación para este algoritmo se probaron utilizando una fuente acústica, mostrando que para un nivel de señal de entrada mayor que 15dB en un ambiente con poco ruido el algoritmo identifica la presencia de dicha fuente.

Los resultados de la DOA utilizando el RootMUSIC empleando la división subbanda mostrada en (1) se muestran en la tabla 2. Para el caso de la implementación, sí se tiene un nivel de señal de entrada mayor que 15dB en un ambiente con poco ruido, los ángulos tienen una desviación de 25° alrededor del ángulo real.

Tabla 2: resultados del RootMUSIC

SNR	ÁNGULO REAL								
	-80°	-60°	-40°	-20°	0°	20°	40°	60°	80°
0dB	-30.2	-68.3	-32.5	-18.5	1.2	21.4	38.7	58.1	25.0
10dB	-80.4	-60.1	-40.0	-20.0	0.0	20.0	39.9	39.8	80.6
20dB	-80.0	-60.0	-39.9	-20.0	0.0	19.9	40.0	60.0	80.0

En la Fig. 5 se muestra el patrón de haz del formador óptimo para una simulación que consistió en ubicar la señal de interés a -15° con SNR=20dB y una interferencia en la dirección 30° a SNR=50dB. Cada una de las gráficas corresponde a una diferente región de frecuencia (descomposición subbanda). Se encuentra que el formador de haz es capaz de atenuar la interferencia directiva ubicada a 30° en las bandas de más alta frecuencia, y solo para las dos bandas más altas, el formador es capaz de direccionar el lóbulo principal hacia la dirección de interés.

Las pruebas realizadas con el DSP muestran que al utilizar el formador enventanado se puede eliminar casi cualquier interferencia que se encuentre a 50° desde la dirección de la señal deseada, incluso con potencias similares a la de la señal

de interés. Los resultados son muy alentadores, ya que a pesar de realizar la implementación utilizando tan sólo dos micrófonos se obtiene selectividad espacial y el consecuente realce de voz.

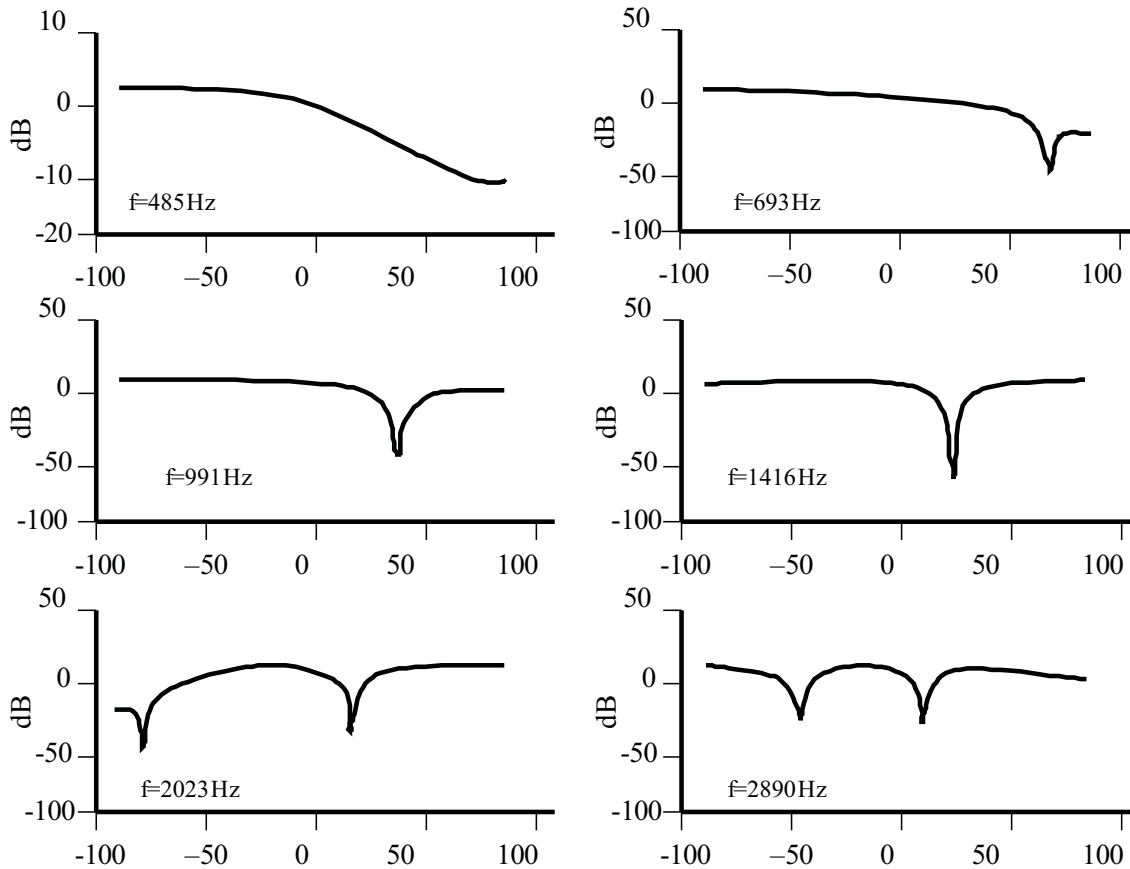


Figura 5: Formación de haz óptima en subbanda para $M=2$.

CONCLUSIONES

Los resultados de la simulación indican que la separación entre micrófonos con mejor desempeño para las señales acústicas bajo prueba y los entornos planteados es $d=0.11\text{m}$. Se encontró que el cálculo de la DOA está afectado por la condición de campo cercano, efectos de banda ancha y errores de cuantización, convirtiéndose éste último en el parámetro más crítico del sistema. Al comparar el tiempo de

procesamiento para el cálculo de los pesos óptimos, para un bloque capturado de 2000 muestras, con el tiempo de captura de dicho bloque, se identificó que es posible recalcular los pesos sin desechar información y garantizar el funcionamiento del sistema en tiempo real. Se mostró que un sistema basado en arreglo de micrófonos permite realzar la voz, obteniendo un incremento significativo de la SINR respecto a la que se obtiene si se emplea un único sensor con similares características a los utilizados por el arreglo.

AGRADECIMIENTOS

Los autores agradecen a la emisora cultural de la Universidad del Quindío, la UFM Estereo, por el préstamo de los estudios de grabación durante las fases de prueba y ajuste del sistema desarrollado.

BIBLIOGRAFÍA

- (1) Campbell D. K. Adaptive Beamforming Using a Microphone Array for Hands-Free Telephony, PhD thesis, Virginia Polytechnic Institute and State University; 1999.
- (2) Applebaum S. P. Adaptive arrays. IEEE Trans. on Antennas and Propagation. 1976;24: 585–598.
- (3) Van Veen B, Buckley K. M. Beamforming Techniques for Spatial Filtering. En: Madisetti V. K, Williams D, editores. The Digital Signal Processing Handbook. 11th ed. New York: CRC Press; 1999.
- (4) Vinay K. I, Manolakis D. G, Stephen M. K. Statistical and Adaptive Signal Processing. McGraw Hill Education; 2000.
- (5) Williams B. D. Detection: Determining the Number of Sources. En: Madisetti V. K, Williams D, editores. The Digital Signal Processing Handbook. 11 ed. New York: CRC Press; 1999.
- (6) Gonen E, Mendel J. M. Subspace-Based Direction Finding Methods. En: Madisetti V. K, Williams D, editores. The Digital Signal Processing Handbook. 11th ed. New York: CRC Press; 1999.
- (7) Proakis J. G. Digital Signal Processing. Principles, Algorithms, and Applications. New Jersey: Prentice Hall; 1996.
- (8) Van Veen B, Buckley K. M. Beamforming: A versatile approach to spatial filtering. IEEE ASSP Magazine. 1988;5:4-24.
- (9) Harris F.J, Buckley K. M. On the use of windows for harmonic analysis with the discrete Fourier transform. Proc. IEEE, 66; January 1978.