

ARRAY DE MICRÓFONOS CON DIRECTIVIDAD VARIABLE PARA AUDITORIOS

REFERENCIA PACS: 43.60.Bf

Guaus i Termens, Enric
Departament d'Acústica, Enginyeria La Salle, Universitat Ramon Llull
Pg. Bonanova nº 8
08022 Barcelona. España
Tel: 34 932 902 400 (ext. 202)
Fax: 34 932 902 416
E-Mail: eguaus@salleURL.edu

ABSTRACT

This work presents a manually variable directivity microphone array. This system, mounted on the roof over the speakers, must be a substitute of the typical microphones mounted on the table. So, the system must be real-time and must have a joystick for selecting the maximum captation direction all over the range between ± 60 degrees. This paper shows the basic concepts of microphone arrays and how the system is implemented in a DSP, using the Wavelet Transform.

RESUMEN

Este trabajo muestra la implementación de un array de micrófonos con directividad variable. El sistema debe ser montado en el techo de una sala de conferencias encima de los oradores y debe sustituir los micrófonos de mesa delante cada orador. El sistema debe funcionar a tiempo real con una variación de directividad entre $\pm 60^\circ$. Este artículo muestra los conceptos básicos de la agrupación de micrófonos y explica la implementación en DSP mediante la Transformada Wavelet.

1. INTRODUCCIÓN

Como punto de partida, recordaremos algunas de las expresiones básicas del comportamiento de las ondas, que nos serán necesarias para los análisis posteriores. La ecuación de onda en función del espacio y el tiempo, para una onda acústica de poca amplitud, viajando en campo libre a través de un medio isotrópico es

$$\nabla^2 p(\mathbf{x}, t) = \frac{1}{c^2} \frac{d^2 p(\mathbf{x}, t)}{dt^2} \quad (\text{Eq. 1})$$

donde c es la velocidad del sonido de la onda en el medio, t es el tiempo y \mathbf{x} es el espacio representado en el sistema de coordenadas correspondiente. Si suponemos que estamos en régimen estacionario senoidal,

$$p(\mathbf{x}, t) = p(\mathbf{x}) e^{j\omega t} \quad (\text{Eq. 2})$$

la ecuación de onda nos queda de la siguiente forma:

$$\nabla^2 p(\mathbf{x}) = -\left(\frac{\omega}{c}\right)^2 p(\mathbf{x}) \quad (\text{Eq. 3})$$

La solución a esta ecuación de onda dependerá del modelo de propagación que estemos estudiando. Así pues, para una propagación de ondas planas en la dirección positiva de ξ nos queda,

$$p(\mathbf{x}) = P e^{-j\left(\frac{\omega}{c}\right)\xi} \quad (\text{Eq. 4})$$

observando que la amplitud P es compleja y constante. Por otra parte, si suponemos un modelo de propagación esférica nos queda una expresión cuya amplitud disminuye con la distancia ξ :

$$p(\mathbf{x}) = P \left(\frac{1}{\xi}\right) e^{-j\left(\frac{\omega}{c}\right)\xi} \quad (\text{Eq. 5})$$

Pero en cualquiera de los dos casos siempre aparece el término $e^{j(\omega/c)\xi}$ que indica la fase con la que nos llega la señal.

2. AGRUPACIÓN DE TRANSDUCTORES

Cualquier transductor lleva asociado, según su estructura interna y forma de fabricación, una característica de directividad en concreto. La directividad se define como la variación de la sensibilidad en función del ángulo de incidencia, y se representa mediante diagramas polares. Normalmente, la directividad se normaliza respecto la dirección de máxima radiación o captación. Algunas características típicas son la omnidireccional, bidireccional o forma de ocho y cardioide.

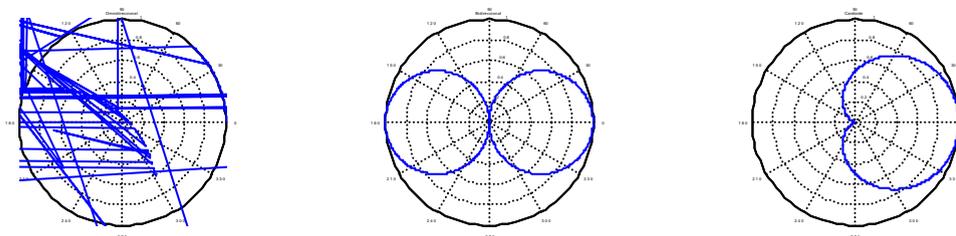


Fig. 1: Diagramas de directividad para transductores omnidireccionales, bidireccionales v cardioides.

En el caso de los micrófonos, la directividad favorece la captación de presión sonora en una dirección determinada en sacrificio de otras, de tal forma que a la dirección favorecida se le puede asignar una ganancia respecto la presión que captaría un micrófono omnidireccional. Se define el ángulo de cobertura de un micrófono, con una directividad fijada entorno a un eje de radiación principal, como el ángulo en que la función de directividad tiene una determinada caída. El ángulo de cobertura se puede definir para pérdidas de 3 dB, de 6 dB o hasta de 10 dB, en función del autor o fabricante.

En muchas aplicaciones, la directividad de un solo micrófono no es suficiente para focalizar una fuente, o para discernir esta fuente respecto el ruido de fondo de la sala. Es por eso que se recurre a la agrupación de micrófonos. Los objetivos de la agrupación de micrófonos son tres: ganancia mayor al eje principal de radiación, disminución del ángulo de cobertura y una mayor relación señal-ruido de la fuente focalizada respecto el resto de oradores o ruido de fondo. Es necesario aclarar que las características de sensibilidad y directividad de los distintos elementos que forman la agrupación no tienen que ser idénticas, pero para facilitar el estudio se toman micrófonos iguales. También debemos comentar que no

hablaremos de la respuesta en frecuencia de los micrófonos individualmente puesto que supondremos que tienen respuesta plana a la zona de interés, aunque el comportamiento de la agrupación sí que dependa de la frecuencia de trabajo.

El principio de funcionamiento de la agrupación de micrófonos se basa en la interferencia constructiva de la onda acústica que llega a los transductores para las direcciones en las que queremos máxima captación y en la interferencia destructiva en el caso contrario. Así pues, las características de directividad de la agrupación vendrán definidas por los siguientes parámetros: 1º. Configuración geométrica de la agrupación. 2º. Distancia relativa entre los elementos. 3º. Sensibilidad de cada elemento. 4º. Fase de cada elemento. 5º. Directividad de cada elemento. Como norma general, a mayor número de elementos, mayor índice de directividad se consigue, aunque suelen aparecer lóbulos secundarios en otras direcciones que perjudican nuestras intenciones. En nuestro caso y por motivos puramente funcionales, trabajaremos con una agrupación de 2 micrófonos, puesto que el DSP de que disponemos sólo tiene 2 entradas.

2.1. Agrupación de 2 Micrófonos

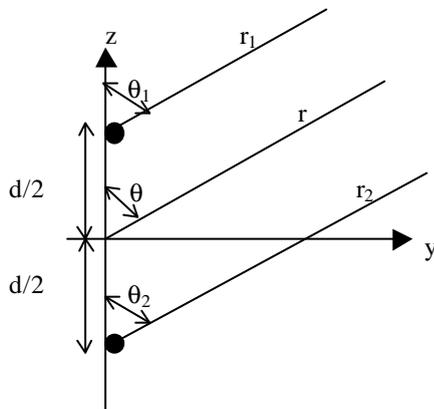


Fig. 2. Agrupación de 2 micrófonos

Hagamos el estudio detallado del comportamiento de una agrupación de 2 micrófonos separados una distancia d y situados sobre el eje z en un sistema de coordenadas cartesianas. Haremos el estudio en 2 dimensiones, puesto que la variación de la captación en la tercera dimensión (eje x) dependerá exclusivamente de la directividad de los elementos individualmente (la agrupación tiene simetría de revolución). Las expresiones de la presión acústica captada por cada uno de los micrófonos es:

$$p_1 = \frac{A}{r - \Delta r_1} e^{j\left(\omega t - \frac{\omega}{c}(r - \Delta r_1)\right)}$$

$$p_2 = \frac{A}{r + \Delta r_2} e^{j\left(\omega t - \frac{\omega}{c}(r + \Delta r_2)\right)}$$

(Eq. 6)

donde

$$\Delta r_1 = \frac{d}{2} \sin \theta_1$$

$$\Delta r_2 = \frac{d}{2} \sin \theta_2$$

(Eq. 7)

Considerando que la fuente está a una distancia suficientemente grande como para considerar $r \gg d$ (aproximación de campo lejano), el módulo de las dos expresiones se puede aproximar por A/r , puesto que $\Delta r_1 \ll r$ y $\Delta r_2 \ll r$, o lo que es lo mismo, $r_1 \approx r_2 \approx r$. En cuanto la fase, aunque podamos aproximar que $\theta_1 \approx \theta_2 \approx \theta$ y que por tanto $\Delta r_1 \approx \Delta r_2 \approx \Delta r$, no podemos despreciar Δr frente r porque este término nos proporciona información de vital importancia: interferencia constructiva o destructiva. Con estas aproximaciones, la presión total captada por la agrupación de micrófonos es:

$$p = p_1 + p_2 = \frac{A}{r} \left(e^{j\frac{\omega d}{c} \sin \theta} + e^{-j\frac{\omega d}{c} \sin \theta} \right) e^{-j\frac{\omega}{c} r} e^{j\omega t}$$

(Eq. 8)

que aplicando Euler y simplificando obtenemos:

$$p = \underbrace{\frac{A}{r}}_{\text{Amplitud elemento individual}} \underbrace{2 \cos \left[\frac{1}{2} \left(\frac{w}{c} d \cos q + b \right) \right]}_{\text{Factor de Array (AF)}} \underbrace{e^{jw t}}_{\text{Variacion temporal senoidal}} \quad (\text{Eq. 9})$$

donde β es la diferencia de fase de la excitación de los dos elementos, y éste es el parámetro importante con el que trabajaremos para definir la directividad variable.

Como conclusión, diremos que la característica total de la agrupación de micrófonos es la multiplicación de características del elemento individual por el Factor de Array, suponiendo elementos idénticos. De este modo, en función del diseño de la agrupación, tendremos unas características de directividad u otras.

2.2. Directividad Variable

Hemos visto en el apartado anterior que la directividad de la agrupación depende del Factor de Array, y que éste depende de la frecuencia (w), distancia de separación entre elementos (d), ángulo de incidencia (θ) y diferencia de fases entre las señales captadas por los micrófonos (β). Nuestro objetivo es diseñar un sistema que proporcione un máximo de la función de directividad en un ángulo θ definible por el usuario, y con un ancho de banda suficientemente grande. La idea es que el sistema sea capaz de modificar la fase de la señal que captamos de los dos micrófonos, y solo puede hacerse o bien modificando físicamente la distancia entre micrófonos, o bien modificando la señal eléctrica captada, retardándola más o menos en el tiempo. La segunda opción es la implementada en nuestro sistema.

Para una frecuencia fija w y una distancia entre micrófonos d , podemos calcular el retardo en tiempo entre las señales captadas por los micrófonos, para conseguir un máximo de directividad a un ángulo θ , igualando el Factor de Array a 1:

$$\cos \left[\frac{1}{2} \left(\frac{w}{c} d \cos q + b \right) \right] = 1 \quad \longrightarrow \quad b = -\frac{w}{c} d \cos q \quad (\text{Eq. 10})$$

Los valores de β , en radianes, debemos convertirlos a tiempo (Eq. 11), y finalmente, si el sistema implementado es digital, obtendremos el nº de muestras (Eq. 12):

$$t[\text{seg}] = \frac{b[\text{rad}]}{w \left[\frac{\text{rad}}{\text{seg}} \right]} \quad (\text{Eq. 11}) \quad n^\circ \text{muestras} = t[\text{seg}] \cdot f_m \left[\frac{\text{muestras}}{\text{seg}} \right] \quad (\text{Eq. 12})$$

Veamos un ejemplo: para $w=2\pi \cdot 1000$ rad/s, $d=0.5$ m y $\theta_{\max}=30^\circ$ corresponde una $\beta=-7.886$, y conseguimos un diagrama polar como el de la Fig. 4. El signo de β indica que el retardo puede ser positivo o negativo en el tiempo, que en este tipo de montajes con 2 micrófonos, no tiene mucho sentido. En nuestro ejemplo, el retardo entre los dos canales es de $\tau=1.255$ ms que muestreando a $f_m=44100$ Hz equivale a 55.34 muestras.

Pero qué ocurre con este modelo si la señal que nos llega es de $f=2$ KHz? Con la β calculada para $w=2\pi \cdot 1000$ rad/s, $d=0.5$ m y $\theta_{\max}=30^\circ$ pero con la nueva frecuencia nos quedará un diagrama de directividad como el de la Fig. 5, que nada tiene que ver con lo que nosotros

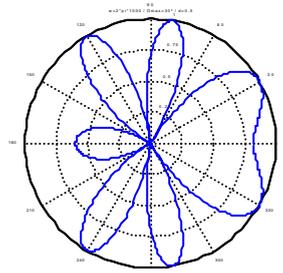


Fig. 4. $w=2\pi \cdot 1000$ rad/s / $d=0.5$ m / $\theta_{\max}=30^\circ$ / $\beta=-7.886$

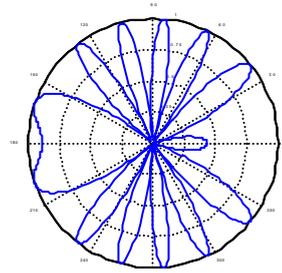


Fig. 5. $w=2\pi \cdot 2000$ rad/s / $d=0.5$ m / $\theta_{\max}=30^\circ$ / $\beta=-7.886$

deseamos.

Está visto que la Función de Array depende de la frecuencia y que es prácticamente imposible conseguir un máximo de directividad para un ángulo en concreto aplicando un mismo retardo τ a toda la banda de frecuencias.

3. LA TRANSFORMADA WAVELET: CODIFICACIÓN EN SUBBANDAS

No entraremos aquí en la explicación profunda de la Transformada Wavelet y de su implementación rápida mediante el Análisis Multiresolución (véase la bibliografía), pero sí que comentaremos que la Transformada Wavelet surgió para resolver los problemas de resolución constante de la Short Time Fourier Transform (Principio de Incertidumbre, o desigualdad de Heisenberg) y que aprovecha la propiedad de escalaje de la Transformada de Fourier:

$$y(t) \longleftrightarrow \Psi(\Omega) \quad \frac{1}{\sqrt{a}} y\left(\frac{t}{a}\right) \longleftrightarrow \sqrt{a} \Psi(a\Omega) \quad (\text{Eq. 13})$$

para definir, a partir de una Wavelet Madre $\psi(t)$, una Familia de Wavelets $\psi_{a,b}(t)$ que son las versiones escaladas y trasladadas de la Wavelet Madre

$$y_{a,b}(t) = \frac{1}{\sqrt{a}} y\left(\frac{t-b}{a}\right) \quad (\text{Eq. 14})$$

Podemos definir la Transformada Continua de Wavelet como

$$W(a,b) = \int_{-\infty}^{\infty} f(t) y_{a,b}(t) dt \quad (\text{Eq. 15})$$

De este modo, conseguimos un plano tiempo-escala (hablaremos de escalas y no de frecuencias) con celdas de distintas resoluciones para distintas escalas, de tal modo que desde el punto de vista frecuencial se puede interpretar como un banco de filtros de porcentaje constante, y no como un banco de filtros de banda constante que es lo que nos proporciona un espectrograma mediante la STFT.

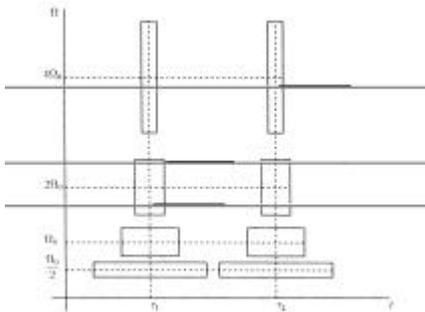


Fig. 6. Plano tiempo-escala para la TW

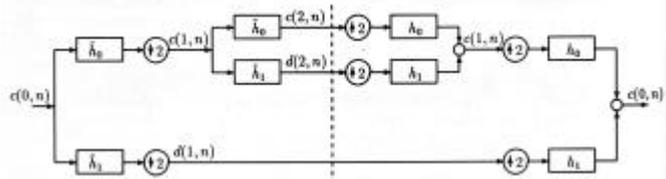


Fig. 7. Diagrama de bloques para un AMR

La Transformada Wavelet aquí definida es redundante e imposible de implementar a la práctica. Si bien una opción para solucionar el problema es discretizar los parámetros a y b , definiendo entonces la DWT (Discrete Wavelet Transform) nos será más útil trabajar con el AMR (Análisis Multiresolución) que se basa en el uso de filtros PR-QMF (Perfect Reconstruction Quadrature Mirror Filters) para la descomposición de una señal en distintas subbandas, permitiendo una recomposición perfecta de la señal original.

3.1. Procesado de las Señales de cada Micrófono

En nuestro caso escogeremos una decodificación en 4 subbandas mediante wavelets de Daubechies de $N=5$, de tal modo que nos queden tres de ellas centradas a las frecuencias de 2 KHz, 1 KHz y 500 Hz, y una última banda que no utilizaremos con la aproximación de baja frecuencia. De este modo, tendremos la señal de cada micrófono dividida en tres bandas de octava (500, 1K y 2K), donde se concentra la energía de la voz humana. Ahora es necesario

repetir el proceso explicado al apartado 2.2. para determinar cuál es el retardo necesario τ en cada octava para conseguir un máximo en la función de directividad (Factor de Array) en la dirección deseada θ_{max} . Debido a los operadores de delmación mostrados a la Fig. 7, al pasar el retardo de segundos a muestras, deberemos tener en cuenta diferentes frecuencias de muestreo para cada subbanda.

4. IMPLEMENTACIÓN A TIEMPO REAL

Para la implementación a tiempo real hemos utilizado el DSP de Analog Devices ADSP-21061 que nos proporciona 2 canales de entrada y un conversor A/D y D/A con frecuencia de muestreo múltiple. La versatilidad y facilidad de programación nos han ayudado a la elección. Los micrófonos utilizados han sido del tipo Electret y se han conectado, después del circuito de polarización, a las entradas del DSP.

A cada una de las señales se les ha aplicado la descomposición AMR. Aunque el retardo τ calculado en función del ángulo sólo se aplique a una de las dos señales, la descomposición se aplicará a los dos canales, por problemas de sincronización: la descomposición y reconstrucción aplica, ya de por sí, un retardo cuyo valor estará en función de los filtros utilizados y del nivel de descomposición. Se aplicará un retardo distinto para cada banda de octava en función de la β calculada. Una vez aplicado el retardo, se procederá a la reconstrucción de la señal y, finalmente, a la suma de los dos canales. Llegados a este punto, ya tenemos la señal procedente de los dos micrófonos y cuyo diagrama de directividad tiene un máximo en la dirección escogida θ_{max} .

5. RESULTADOS

Sigamos con el ejemplo anterior. Para $d=0.5m$ y $\theta_{max}=30^\circ$, calculemos los 3 valores de β , τ y n° de muestras para las tres bandas de 500 Hz, 1 KHz y 2 KHz:

frec. (Hz)	β (rad)	τ (mseg)	f_m (Hz)	Muestras
500	-3.943	1.255	1414.2	1.774
1000	-7.886	1.255	2828.4	3.549
2000	-15.772	1.255	5656.8	7.099

Observamos que tenemos retardos digitales de fracción de muestra, imposibles de implementar. Ante este problema, tenemos 3 opciones distintas: Modificar las frecuencias de muestreo aplicando un Oversampling, interpolar el valor de la muestra resultante con los datos almacenados en el buffer o redondear hasta el número entero más cercano. En nuestro caso, por motivos de simplicidad, redondearemos hasta el valor entero más cercano. Los gráficos teóricos obtenidos con estos valores son los mostrados a la Fig. 8

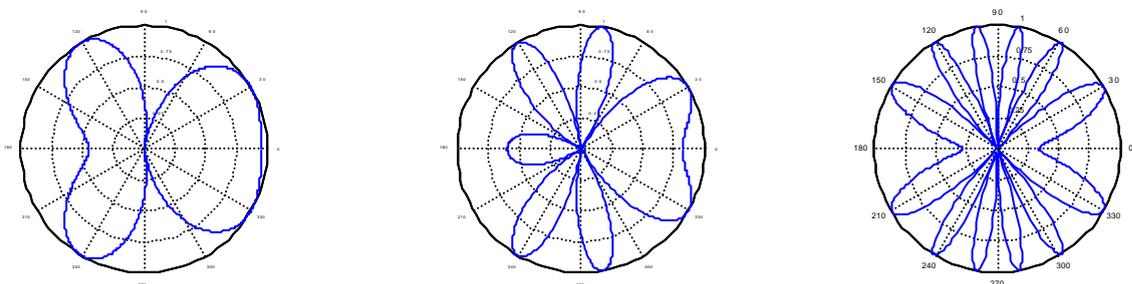


Fig. 8. Directividades con máximo a 30° para $f=500$ Hz, 1 KHz y 2 KHz

Lamentablemente aun no disponemos de resultados prácticos del sistema formado por la agrupación de 2 micrófonos conectados al DSP y midiendo la directividad en cámara anecoica. Una vez realizadas las pruebas, estarán disponibles en la dirección de internet <http://www.salleURL.edu/~eguaus>.

6. CONCLUSIONES

Hemos diseñado una agrupación de 2 micrófonos con directividad variable y a tiempo real cuyos resultados son, aunque no espectaculares en cuanto a directividad, el principio de un sistema mayor con una agrupación de micrófonos en dos dimensiones y con un ancho de banda superior. El uso de la Transformada Wavelet y Análisis Multiresolución nos ha sido de gran ayuda para poder descomponer la señal en distintas bandas y definir así la directividad en concreto para cada una de ellas.

7. BIBLIOGRAFIA

- COMPUTER-STEERED MICROPHONE ARRAYS FOR SOUND TRANSDUCTION IN LARGE ROOMS; J. L. Flanagan, J. D. Johnston, R. Zahn, G. W. Elko; j. Acoust. Soc. Am. 78 (5), November 1985
- MEJORA DE VOZ CON ARRAYS DE MICRÓFONOS MEDIANTE DESCOMPOSICIÓN EN COMPONENTES DE FASE MÍNIMA Y PASO TODO; Joaquín González Rodríguez, José Luis Sánchez Bote, Javier Ortega Garcia; Comunicaciones Tecniacústica'99
- MICROFONOS (tecnología y aplicaciones); John Borwick; Escuela de Cine y Vídeo
- FUNDAMENTALS OF ACOUSTICS; Lawrence E. Kinsler & Austin R. Frey; John Wiley & Sons, INC
- MULTIREOLUTION SIGNAL DESCOMPOSITION; Ali N. Akansu & Richard A. Haddad; Academic Press
- ANTENNA THEORY, ANALYSIS AND DESIGN; Constantine A. Balanis, John Wiley & sons