

ESTUDIO COMPARATIVO DE DIFERENTES TÉCNICAS DE INVERSIÓN DE LA ACÚSTICA DE RECINTOS EN REPRODUCCIÓN MULTICANAL DE SONIDO

José Javier López, Alberto González, Vicenç Almenar

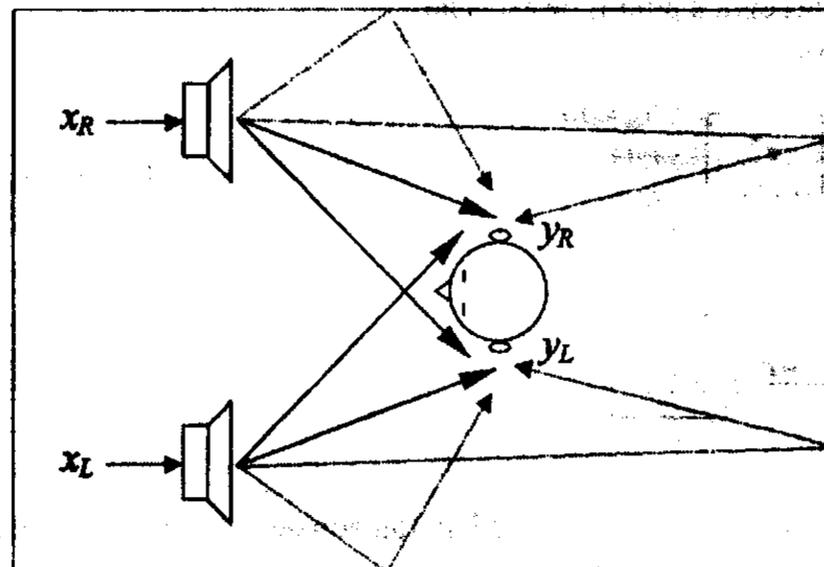
Departamento de Comunicaciones. Escuela Universitaria de Gandia
Ctra. Nazaret-Oliva s/n, 46730 GRAO DE GANDIA, Spain
Tel: +34 96 2849300, Fax: +34 96 2849313
Email : jjlopez@dcom.upv.es, agonzal@dcom.upv.es

SUMMARY

In sound reproduction systems signals are distorted by reverberation, also cross-talk effect is produced in stereo systems. This paper compares different techniques for avoid this problems based on placing a bank of filters before sources than minimizes the error between the desired signal at the listening point. LSE criteria, MINT method and Fast Deconvolution using regularization are studied under same conditions. HRTF are used as acoustic paths between speakers and listening points.

1. INTRODUCCIÓN

En los sistemas de reproducción de sonido en salas, las señales de voz o música sufren el efecto de reverberación producido por rebotes de la señal en las paredes que rodean al oyente. Además si se trata de un sistema de reproducción estéreo, se produce el efecto de *cross-talk*, parte de la señal que emite el altavoz izquierdo también llega al oído derecho y viceversa.



El conjunto de los efectos que se producen en la reproducción del sonido, se pueden modelar por una función de transferencia (RTF) entre cada una de las fuentes de sonido y cada uno de los sensores de recepción. Esta función de transferencia de características complejas varía dependiendo de la posición de las fuentes y del oyente.

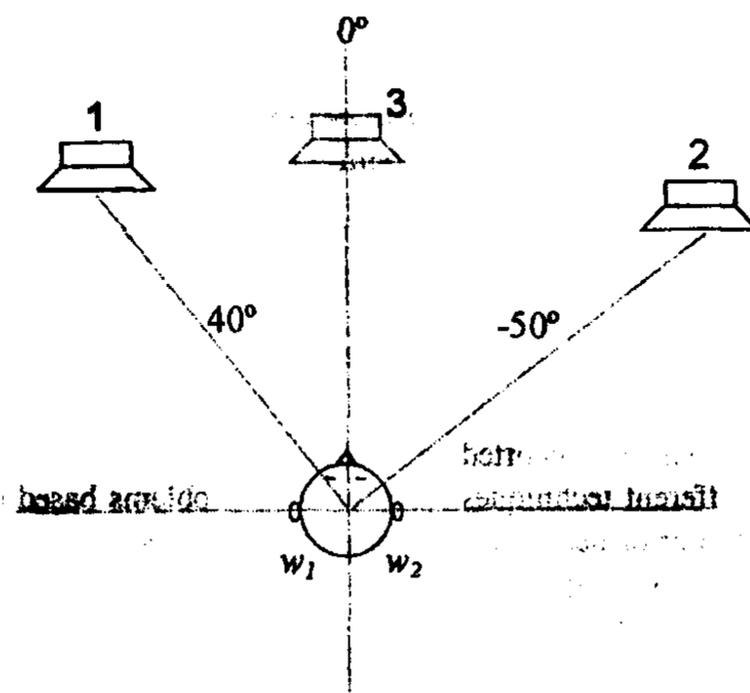
Generalmente se utiliza como solución un filtro de ecualización situado entre la fuente de sonido y el amplificador de potencia para ajustar la respuesta en frecuencia de la RTF. La solución más común es utilizar un ecualizador gráfico compuesto de múltiples filtros paso banda. Sin embargo solo se ecualiza la amplitud de la RTF en este caso, ignorando la fase, por lo que no se elimina la reverberación ni tampoco el *cross-talk*.

Si conocemos de forma exacta las respuestas impulsionales, podemos aplicar técnicas de inversión de la misma mucho más sofisticadas y aplicarlas a un filtro previo a los altavoces, para que el sonido llegue tal y como se desea.

El objeto de este trabajo es la comparación de distintos métodos que aparecen en la literatura sobre este tema, aportando métodos eficientes computacionalmente y viables para ser implementados en tiempo real.

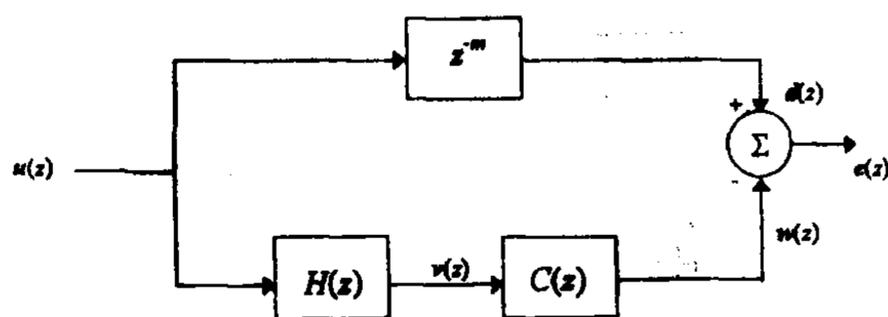
2. CONFIGURACIÓN DE LOS EXPERIMENTOS

Para la realización de los experimentos hemos supuesto los altavoces emplazados de forma no simétrica. El altavoz izquierdo se encuentra a 40° de la frontal del oyente, mientras que el altavoz derecho se encuentra situado a -50° . Un tercer altavoz emplazado justo en la frontal se empleará solo en el método MINT de inversión de la acústica (el cual necesita un número de fuentes superior al de puntos de escucha). La disposición no simétrica ayuda a evitar posibles polos comunes que afecten a la bondad de los algoritmos.



Denominaremos a las señales de la forma :

- $u(z)$ vector de señales a reproducir (2 en el caso de estéreo)
- $v(z)$ vector de señales de excitación a los altavoces
- $w(z)$ vector de señales en los puntos de escucha
- $d(z)$ vector de señales deseadas en los puntos de escucha
- $e(z)$ vector de la señal de error



$H(z)$ banco de filtros ecualizadores
 $C(z)$ caminos acústicos entre las fuentes y los puntos de escucha

Los caminos acústicos para el oyente de la figura, se han tomado de la base de datos de medidas de la HRTF del maniquí para medidas psicoacústicas KEMAR que ha realizado el MIT y que son de dominio público [4]. Estas respuestas son incluyen ningún tipo de reverberación por lo que este estudio se centrará en la eliminación del cross-talk. Los altavoces se encuentran en el plano horizontal. La respuestas impulsionales tienen en todos los casos una longitud de 512 muestras y han sido tomadas con una $f_s = 44.1$ KHz.

3. COMPARACION DE DIFERENTES TÉCNICAS DE INVERSIÓN

Para la disposición de elementos que se muestran en el punto anterior, se han calculado las respuestas impulsionales del banco de filtros $H(z)$ que debe de eliminar los efectos de cross-talk en la recepción del sonido. Este banco de filtros tendrá una estructura de 2×2 filtros excepto en el método MINT que será de 3×2 .

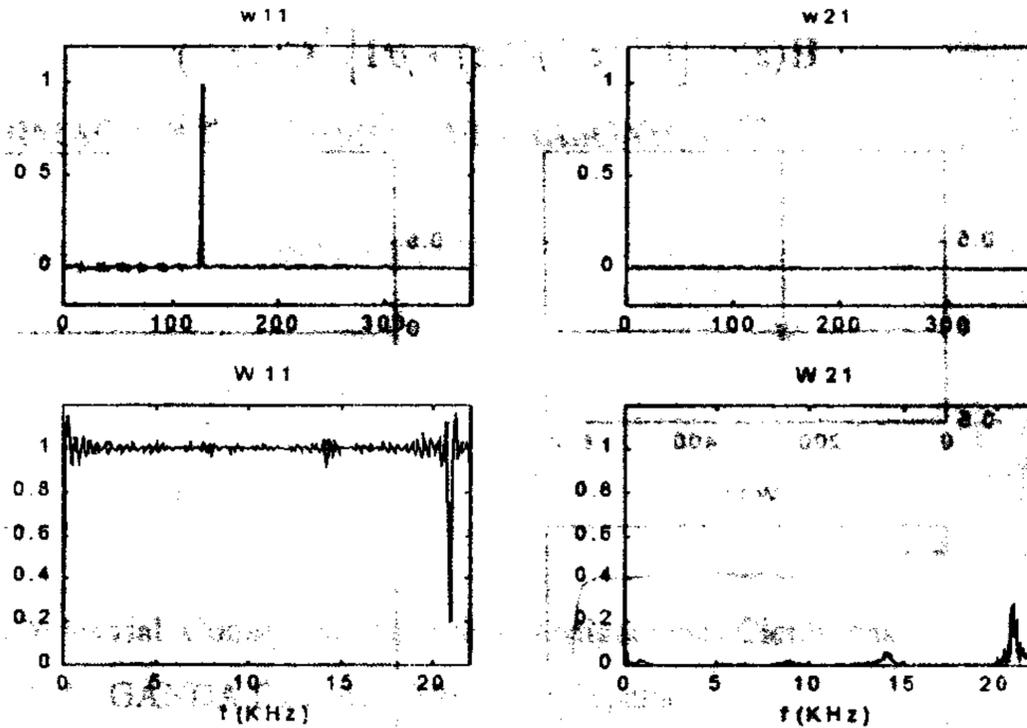
3.1 MINIMOS CUADRADOS (LSE)

La técnica de mínimos cuadrados (LSE) se puede emplear con éxito [1] [2] [3], para resolver el sistema de ecuaciones que aparece cuando expresamos la convolución en forma de producto matricial. Considerando M la longitud de las repuestas impulsionales de cada camino acústico y N la longitud de los filtros ecualizadores.

$$\begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$$

La matriz C no puede ser cuadrada y tiene unas dimensiones de $2(N+M-1) \times 2N$

Por lo tanto se recurre a resolver el sistema mediante mínimos cuadrados $C^T D = C^T C H$

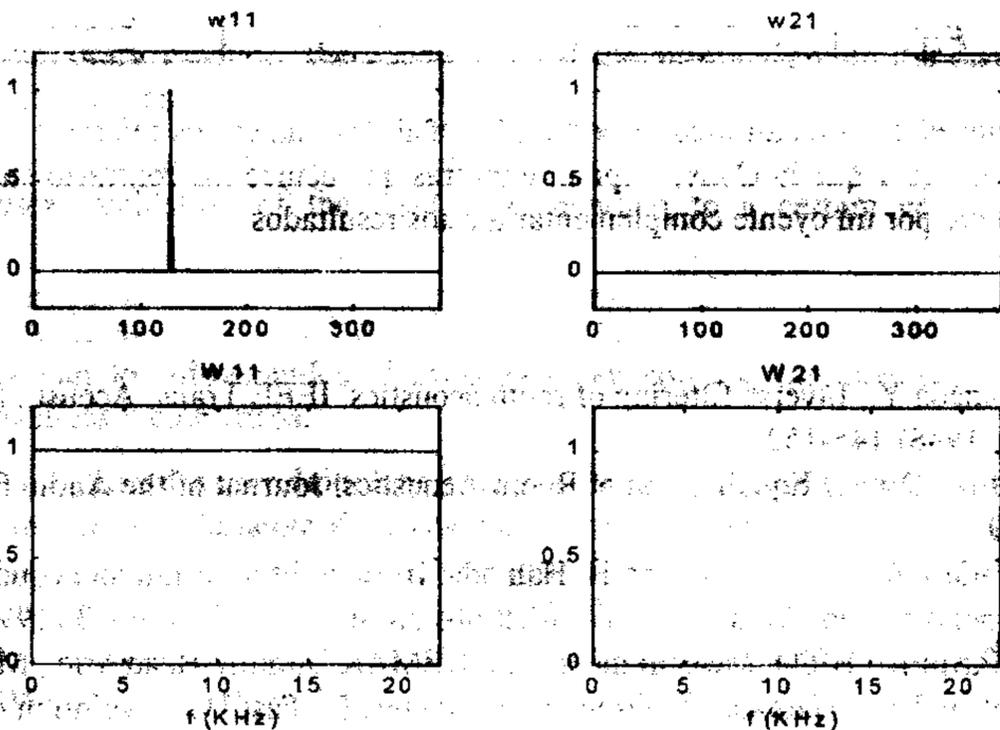


3.2 MÉTODO MINT

El método MINT propuesto por Miyoshi y Kaneda [1], propone utilizar un número de fuentes superior (por lo menos en 1) al número de puntos de recepción. En nuestro experimento hemos empleado 3 altavoces dispuestos como se indica en el punto 2. El sistema de ecuaciones queda así:

$$\begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \\ H_{31} & H_{32} \end{bmatrix}$$

La matriz C tiene unas dimensiones de $2(N+M-1) \times 3N$. Para conseguir que sea cuadrada debemos de elegir un filtro ecualizador con una longitud $N = 2M - 2$. Para estas condiciones y si la matriz no es singular o está mal condicionada, se obtiene una solución teóricamente perfecta como muestra la figura.



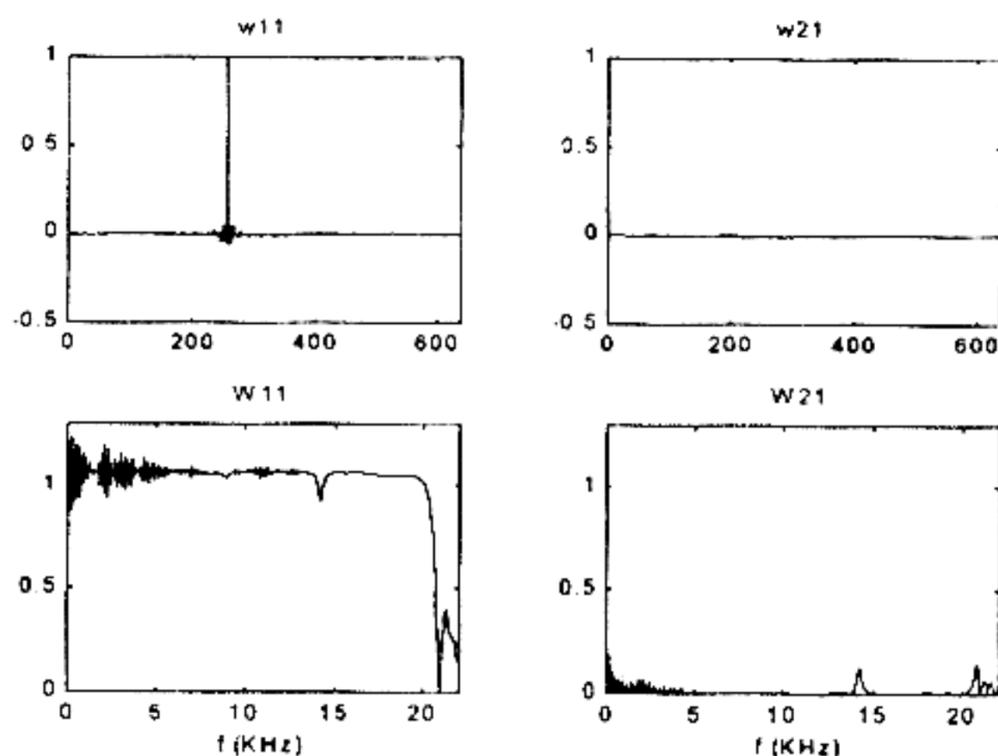
3.3 DECONVOLUCIÓN RÁPIDA USANDO REGULARIZACIÓN

La utilización de la transformada discreta de Fourier (DFT) para deconvolucionar una/s respuesta/s impulsionales, invirtiendo cada uno de los coeficientes frecuenciales de la DFT presenta el problema del aliasing, ya que aparece el efecto de convolución circular.

Sin embargo utilizando una longitud para el filtro inverso mucho mayor que la del filtro que se pretende invertir y usando un factor de regularización se pueden obtener buenos resultados [5]. El empleo de la FFT, hace que ésta sea una forma computacionalmente muy eficiente de calcular el banco de filtros $H(z)$.

La expresión para el banco de filtros queda así, donde β es el coeficiente de regularización, que se calcula de forma empírica.

$$H(z) = [C^T(z^{-1})C(z) + \beta I]^{-1} C^T(z^{-1})$$



4. ANÁLISIS DE RESULTADOS Y CONCLUSIONES

En cada caso hemos representado la repuesta al impulso del sistema así como la respuesta en frecuencia para la señal 1 y los micrófonos 1 y 2. Para una reproducción ideal w_{11} debería de ser el impulso unidad y w_{22} cero. Observamos que el método MINT nos proporciona la solución ideal y que el LSE proporciona una respuesta frecuencial con muchos picos, ligeramente peor que la Regularización que habría que valorar subjetivamente.

	LSE	MINT	Conv. Rápida
$\overline{e_{11}^2}$	0.0080	≈ 0	0.043
$\overline{e_{21}^2}$	0.0012	≈ 0	0.00078

En la tabla observamos el error cuadrático respecto a la señal descada en los micros 1 y 2. El LSE proporciona un mejor error en el micro 1 que la Conv. Rápida, mientras que ocurre lo contrario en el micro 2. Un estudio valorado subjetivamente por un oyente complementaría estos resultados.

REFERENCIAS

- [1] Miyoshi M., Kaneda Y., Inverse filtering of room acoustics, IEEE Trans. Acoust., Speech, Signal Processing, 36 (1988) 145-152
- [2] Mourjopoulos J.N., Digital Equalization of Room Acoustics. Journal of the Audio Engineering Society, 42 (1994) 884-900
- [3] P.A. Nelson, F. Orduña-Bustamante and H. Hamada, Inverse filter design and equalization zones in multichannel sound reproduction, IEEE Transactions on Speech and Audio, 3 (1995), 185-192
- [4] Gardner W., Martin K., HRTF Measurements of a KEMAR Dummy-Head Mic., MIT Int. Report, (1994)
- [5] Kirkeby O., Nelson P.A., Hamada H., Orduña-Bustamante F., Fast Deconvolution of Multichannel Systems Using Regularization