

Análisis objetivo y subjetivo de la compensación de salas mediante técnicas Wave-Field Synthesis

MSc. Laura Fúster

Instituto de Telecomunicaciones y Aplicaciones
Multimedia, iTEAM.
Universidad Politécnica de Valencia
Valencia, España
lfuster@iteam.upv.es

PhD. Alberto González

Instituto de Telecomunicaciones y Aplicaciones
Multimedia, iTEAM.
Universidad Politécnica de Valencia
Valencia, España
agonzal@dcom.upv.es

PhD. José Javier López

Instituto de Telecomunicaciones y Aplicaciones
Multimedia, iTEAM.
Universidad Politécnica de Valencia
Valencia, España
jjlopez[at]dcom.upv.es

PhD. Máximo Cobos

Instituto de Telecomunicaciones y Aplicaciones
Multimedia, iTEAM.
Universidad Politécnica de Valencia
Valencia, España
mcbos[at]iteam.upv.es

Resumen— En el campo de la reproducción de audio espacial, uno de los sistemas más prometedores hoy en día es el sistema de reproducción *Wave-Field Synthesis*. Dicho sistema proporciona con gran precisión espacial un campo sonoro en una amplia zona de escucha. Sin embargo, en la reproducción de sonido dentro de una sala, la señal originada sufre los efectos de reverberación producidos por las paredes, suelo y techo de la sala, que reducen el efecto espacial buscado. En este trabajo se presenta un método de compensación de salas basado en los sistemas de inversión multicanal que corrige los efectos producidos por la sala en los puntos seleccionados dentro del área de escucha. Dichos puntos han sido seleccionados a fin de permitir una zona de escucha amplia, y con una separación suficiente para poder cubrir la distancia entre ellos. Para calcular el banco de filtros inversos, se han utilizado los métodos de deconvolución rápida en el dominio de la frecuencia. Dichos filtros han sido calculados en una etapa previa a la compensación, ya que las reflexiones principales de la sala se han considerado invariables en el tiempo. No obstante, se ha elegido el algoritmo basado en el dominio de la frecuencia ya que tiene menor coste computacional para una posible implementación en tiempo real del sistema de compensación. Para validar el método propuesto se han realizado diversas medidas de laboratorio, tanto mediante simulaciones como reales en laboratorio. Dichas simulaciones han proporcionado interesantes resultados. Finalmente, se presenta una evaluación subjetiva inicial del sistema de compensación implementado, a fin de comparar los efectos del sistema de corrección en función de la ca-

lidad percibida y de la localización de la fuente percibida por los oyentes o jueces.

Palabras clave— *Wave-Field Synthesis*, compensación de salas, sistemas MIMO, análisis subjetivo.

Abstract— In the field of spatial audio reproduction, one of the most promising systems available today is *Wave-Field Synthesis*, which provides a precise spatial sound field in a wide area. However, the sound reproduction inside a room suffers from echoes from the walls, floor and ceiling that reduce its spatial effect. In this paper a room compensation method based on a multichannel inverse filter bank that corrects the room effects at selected points within the listening area is briefly explained. Points are selected in order to allow a huge sweet spot disregarding the quality of the room. For computing the bank of inverse filters, frequency domain algorithms have been used in a previous stage to the compensation, since the main room reflections have been considered invariant for each specific room. Nevertheless, the frequency domain algorithm has been chosen in case that a real time implementation of the compensation system was needed. Moreover, a carefully designed laboratory experiment to validate the method is also reported. This method is evaluated with interesting results using a simulated field. Finally, an initial subjective evaluation of the compensation system is presented, in order to compare the effects of the correction system in terms of the perceived quality and source localization by the subjects.

Keywords: wave-field synthesis, room compensation, MIMO systems, subjective analysis.

I. INTRODUCCIÓN

Wave-Field Synthesis (WFS) [1] es uno de los sistemas de reproducción de audio espacial más prometedores en la actualidad. El campo sonoro sintetizado mediante dicho sistema permite abarcar una amplia área de escucha por medio de una disposición lineal (arrays) de altavoces. La ventaja de este sistema es que WFS proporciona una zona de escucha mucho mayor que otros sistemas, tan grande como la zona encerrada por los altavoces. Sin embargo, el sistema WFS está basado en la propagación de ondas en campo libre, por lo que uno de los mayores problemas a la hora de implementar un sistema WFS consiste en la interacción del array de altavoces con las paredes de la sala de escucha. La sala de escucha introduce nuevos ecos o rebotes que no forman parte de la señal que se va a reproducir, pues altera la síntesis del campo sonoro original, reduce el efecto espacial y, por tanto, el potencial de este sistema de reproducción.

En la última década, se han desarrollado y estudiado diferentes algoritmos para poder compensar dichos efectos de reflexión de la sala, utilizando para realizar la compensación los propios altavoces que se usan en el sistema de reproducción mediante WFS. Principalmente se han propuesto dos métodos de compensación: el primero de ellos está basado en el dominio de onda mediante la descomposición del campo sonoro en una serie de ondas planas, [2] y [3]; y el segundo de ellos basado en los sistemas de múltiples entradas y múltiples salidas, sistemas MIMO (Multiple-Input Multiple-Output). Este último tipo de métodos consideran el sistema de audio como un problema donde las múltiples entradas corresponden a los altavoces y las múltiples salidas a los micrófonos, [4]. En ambos métodos propuestos, para poder compensar los efectos de la sala, se debe introducir un banco de filtros como etapa previa a la reproducción de las fuentes por parte de los altavoces del sistema WFS.

La parte II introduce brevemente la teoría relacionada con *Wave-Field Synthesis* y la inversión multicanal. En la parte III se detalla el experimento desarrollado en el laboratorio y se muestran los resultados obtenidos para validar este método de compensación.

II. FUNDAMENTOS TEÓRICOS

La notación que se va a emplear en este artículo para elementos vectoriales es: para matrices, negrita y mayúscula (**A**); para vectores, negrita y minúscula (**a**); y para elementos escalares, minúscula (a).

A. *Wave-Field Synthesis*

El concepto de *Wave-Field Synthesis* (WFS) fue propuesto, por vez primera, en 1988, por A.J.Berkhout en la Universidad de Delft, [1]. Al partir de un enfoque físico basado en la teoría de ondas, se puede reproducir, con ayuda de un array de altavoces, un campo sonoro específico en un ambiente de reproducción concreto. La motivación de este nuevo enfoque para reproducción multicanal surge de la limitación de los sistemas de reproducción existentes hasta entonces. Entre estas limitaciones se encuentra la fuerte dependencia de la posición del oyente, por lo que estos sistemas son inadecuados para zonas de escucha extensas; o en la reproducción con auriculares, el efecto de escucha desde dentro de la cabeza. Este método WFS ofrece un área de escucha amplia con propiedades temporales y espaciales de alta calidad.

Wave-Field Synthesis se basa en el principio de propagación de Huygens, [1] y [5]. Según este principio, un frente de ondas radiado por una fuente, se puede sintetizar a partir de la contribución de una distribución de fuentes secundarias que se encuentren situadas sobre dicho frente de ondas. La presión sonora obtenida en un punto R interior al área rodeada por los altavoces se puede aproximar mediante:

$$P(r_R) = \sum_{l=1}^L Q_l(w) \frac{e^{-jk\Delta r_l}}{\Delta r_l} \quad (1),$$

donde L es el número de altavoces, k es el número de onda, Δr_l es la distancia entre la fuente secundaria y el punto R y $Q_l(w)$ describe la contribución de cada una de las L fuentes secundarias, o altavoces, para la frecuencia angular w [6]. Nótese que $e^{-jk\Delta r_l}/\Delta r_l$ representa la propagación en campo libre entre la fuente secundaria y el punto R .

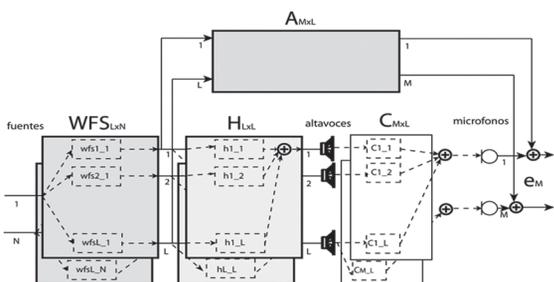
Después de tres décadas, dicho sistema ha sido ampliamente estudiado describiendo las grandes ventajas que supone, así como los problemas que

presenta: errores de amplitud debido a la simplificación del sistema de 3 dimensiones a un plano situado a la altura del oyente; la frecuencia de aliasing o frecuencia de reproducción máxima libre de errores debida al muestreo espacial; o efecto de fin de array. Existen varias técnicas para paliar estos defectos, por lo que en este artículo no vamos a entrar en detalle sobre ellas. Una descripción general de todas estas técnicas se puede encontrar en [6].

B. Aplicación de sistemas MIMO para la compensación en salas en WFS

Los sistemas con múltiples entradas y salidas linealmente relacionadas se les conoce comúnmente como sistemas MIMO. Donde la relación entre cada entrada y cada salida puede describirse por un sistema lineal e invariante en el tiempo (LTI), y si es así, por su correspondiente respuesta al impulso. Por lo tanto, un sistema MIMO de L entradas y M salidas se compone de $M \times L$ respuestas impulsionales. El problema de filtrado inverso en sistemas de reproducción de audio multicanal consiste básicamente en el diseño de una matriz \mathbf{H} de filtros digitales de duración finita (cada columna de \mathbf{H} representa un filtro diferente para cada señal que va a ser reproducida), cuya convolución con los canales de transmisión de la señal (\mathbf{C}), o matriz del sistema electroacústico, mejor se aproxime a la respuesta deseada (\mathbf{A}), [7].

FIG. 1. COMPENSACIÓN DE SALAS MULTICANAL APLICADO A WFS



Fuente: Los autores

La Fig. 1 representa el diagrama de bloques del filtrado inverso multicanal aplicado a WFS. En esta aplicación, el número de entradas y salidas al banco de filtros \mathbf{H} es el mismo e igual al número de altavoces en el array, L .

Las señales de entrada a la etapa de compensación no corresponden con el sonido original de la

señal virtual, sino que son las señales de excitación de las fuentes secundarias calculadas por el sistema de reproducción WFS (**WFS**). La matriz de filtros \mathbf{H} se calcula usando las respuestas de los canales de transmisión (\mathbf{C}) y las señales deseadas en los puntos de escucha calculadas en condiciones de campo libre.

El cálculo de los filtros inversos se puede realizar en una etapa previa a la reproducción, ya que se considera que las reflexiones principales de la respuesta son invariantes para una sala dada. No obstante, en el caso en que se utilizara un sistema de filtrado inverso adaptativo a las condiciones cambiantes de la sala en tiempo real, esto se podría realizar utilizando las señales de error \mathbf{e}_M que se calculan en función de la diferencia entre las señales recibidas en los puntos de control y las señales deseadas en cada uno de dichos puntos. Dichas señales de error realimentarían los filtros de \mathbf{H} para poder cancelar el error en la zona de escucha.

Se han propuesto diferentes métodos para obtener el banco de filtros de compensación. El método basado en una solución en el dominio del tiempo garantiza un error cuadrático mínimo [8], mientras que otros métodos calculan una solución aproximada en el dominio de la frecuencia usando FFT, [9].

En [10], los autores realizaron un estudio comparativo de ambos métodos y los resultados obtenidos mostraron que con un filtro de longitud suficiente los resultados alcanzados en el dominio de la frecuencia eran comparables a los obtenidos con el algoritmo basado en el dominio del tiempo. Y puesto que el tiempo necesario para calcular el primer tipo de filtros hace posible su implementación en tiempo real, para este experimento se han utilizado los filtros inversos calculados con el sistema de deconvolución rápida con regularización en el dominio de la frecuencia.

III. EXPERIMENTOS Y RESULTADOS

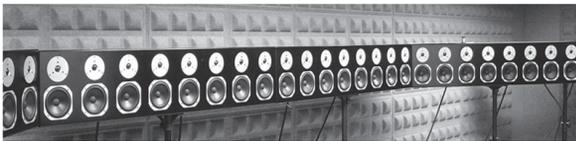
El propósito de los experimentos realizados en el laboratorio es el de validar la inversión multicanal de sistemas MIMO como una solución práctica a la compensación de salas en un área extensa con un sistema de reproducción WFS. A continuación se detallan los experimentos realizados, tanto objetivos como subjetivos, utilizando a oyentes; así como los resultados obtenidos en los mismos.

C. Experimentos

Como prototipo hemos utilizado un sistema de altavoces formado por 4 *arrays* de 8 altavoces cada uno. Cada uno de los altavoces corresponderá a una fuente secundaria; por lo que tendremos 32 altavoces o fuentes secundarias. El sistema se ha montado formando una U abierta, es decir, los dos *arrays* de los extremos se han abierto con un ángulo de 45 grados respecto a los dos *arrays* frontales. La separación entre los distintos altavoces que forman el sistema es constante de 18 cm, por lo que tendremos una frecuencia de *aliasing* de unos 950 Hz.

Dicho sistema se ha instalado en una sala de 4x7x2.5 m que ha sido acondicionada acústicamente para reducir las reflexiones en paredes y techo de la misma, como se puede apreciar en la Fig. 2, por lo que presenta poca reverberación en comparación a otras salas convencionales.

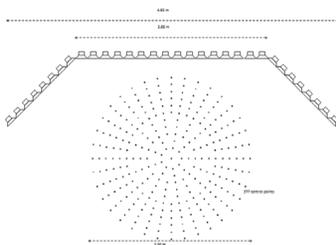
FIG. 2. SISTEMA DE REPRODUCCIÓN WFS DENTRO DE LA SALA DE AUDIO



Fuente: Los autores

Para poder realizar el análisis de la respuesta de la sala nos hemos centrado en una superficie enfrentada al sistema formado por el array de altavoces y cubriendo un círculo de 1.15 m de radio. Para hacerlo hemos utilizado un conjunto lineal de 12 micrófonos situado sobre una plataforma circular, que era manejada directamente por el software de medida de la respuesta de la sala. La Fig. 3 muestra la distribución de los altavoces y la de los puntos de la zona de control. El punto central del mismo es donde se ha realizado el giro para cubrir todo el área. De esta forma, se ha podido medir automáticamente en 577 puntos de forma distribuida sobre dicho círculo.

FIG. 3. SISTEMA DE WFS DE 32 ALTAVOCES EN FORMA DE U ABIERTA Y PUNTOS DE CONTROL



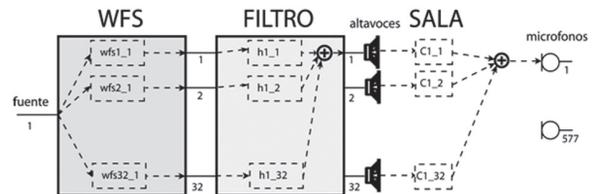
Fuente: Autores

Con esta distribución hemos obtenidos un sistema de 32 x 577 respuestas al impulso de la sala (RIR).

Es decir, la respuesta correspondiente entre cada altavoz o fuente secundaria y cada micrófono o punto de control. Según la frecuencia de *aliasing* para sistemas WFS, y que en este experimento es en concreto de 950 Hz, durante todo el experimento se ha trabajado con señales, y en este caso las RIR, remuestreadas a 8kHz.

El sistema para una única fuente primaria (por simplicidad), con 32 fuentes secundarias (altavoces) y 577 puntos de control (micrófonos) quedará como se muestra en la Fig. 4.

FIG. 4. SISTEMA DE REPRODUCCIÓN WFS CON EL FILTRO DE COMPENSACIÓN, PARA UNA FUENTE, 32 ALTAVOCES Y 577 MICRÓFONOS



Fuente: Los autores

Es decir, de la fuente primaria se calcularán las 32 respuestas de excitación de las fuentes secundarias para dicha fuente, que se reproducirán directamente por los altavoces. El cálculo de la etapa WFS viene representada por la matriz **WFS** (o vectores **wfs**), mientras que la respuesta de la sala desde cada altavoz a cada uno de los puntos de control, viene representada por los vectores **c**. Por simplicidad, en dicha figura se ha representado cómo sería la reproducción entre la fuente 1 y el punto de control 1. Pero en realidad, la matriz **C** estaría formada por 577 x 32 filtros.

Además, se ha representado para una única fuente, pero sería extrapolable a un número indeterminado de fuentes, difiriendo únicamente en la matriz WFS, que pasaría a ser de 32 x 1 a 32 x N filtros, siendo N el número de fuentes primarias (ver Fig. 1). Por otro lado, la etapa de transmisión (caracterizada por la matriz **C**) sería común a ambas. Es decir, la contribución de cada fuente primaria en cada fuente secundaria (altavoz) se sumaría como excitación que dicha fuente reproduciría, por lo que se haría previamente a su reproducción y no afectaría el resto del sistema. Una vez calculados los filtros inversos (matriz **H** de 32

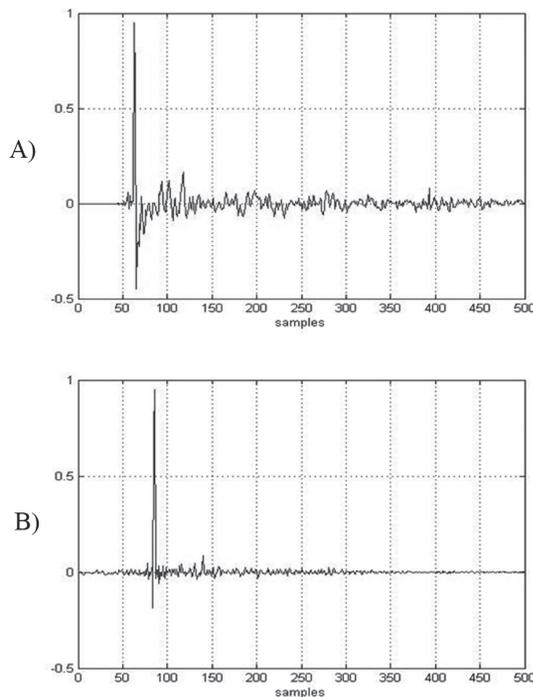
x 32 filtros) se introduce la etapa de compensación directamente entre las etapa de WFS y antes de su reproducción por las fuentes secundarias. Donde se ve claramente que, el sistema de inversión es independiente del número de fuentes primarias utilizadas y también de la posición de las mismas. Además, la matriz \mathbf{H} de filtros inversos calculados es común para todas las fuentes que se quieran reproducir.

D. Resultados

1) Análisis objetivo

Una vez medidas las respuestas de la sala y calculados los filtros inversos, se ha comprobado objetivamente si se consigue la compensación entre cada fuente primaria y cada punto de control. Para hacerlo, se ha calculado la respuesta entre cada fuente secundaria y cada punto de control, con y sin compensación. Dicha respuesta al impulso de la sala (RIR) contiene la señal principal o directa y las primeras reflexiones de la sala, Fig. 5A). Después de aplicar el algoritmo de compensación se obtiene la respuesta mostrada en la Fig. 5B), donde se puede apreciar que las reflexiones principales han sido canceladas.

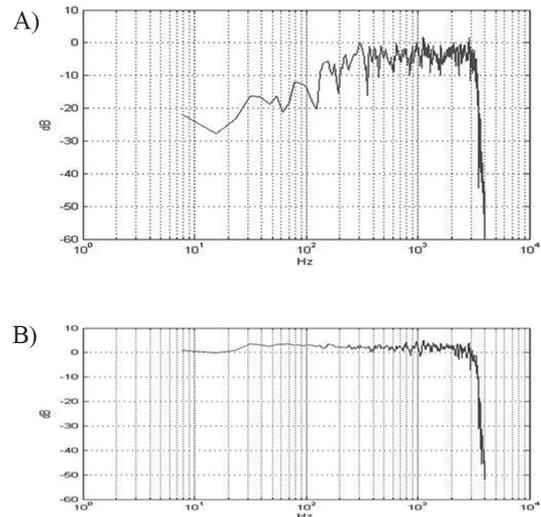
FIG. 5. RESPUESTA AL IMPULSO DEL SISTEMA ENTRE UNA FUENTE SECUNDARIA Y UN PUNTO DE CONTROL: A) ANTES DE LA COMPENSACIÓN Y B) DESPUÉS DE LA COMPENSACIÓN DE LA SALA



Fuente: Los autores

La compensación en el dominio de la frecuencia se puede observar en las Fig. 6A) y 6B). La gráfica para la señal compensada presenta una respuesta prácticamente plana hasta 1kHz y bastante plana a frecuencias mayores.

FIG. 6. RESPUESTA EN FRECUENCIA DE LA SALA ENTRE UNA FUENTE SECUNDARIA Y UN PUNTO DE CONTROL: A) ANTES DE LA COMPENSACIÓN Y B) DESPUÉS DE LA COMPENSACIÓN DE LA SALA



Fuente: Los autores

2) Análisis subjetivo.

Por otra parte, para constatar los resultados obtenidos en las gráficas anteriores, se ha realizado también un análisis subjetivo, para el que se ha tomado una voz cantada como fuente primaria. Dicha voz cantada permitirá distinguir fácilmente la reverberación producida por la sala en la zona de escucha. Dicha señal se ha remuestreado a 8kHz para poder convolucionarla primero con los filtros calculados. En la reproducción de WFS sin compensar los efectos de la sala, se ha convolucionado la respuesta de la fuente primaria con los filtros \mathbf{wfs} y se ha remuestreado la señal de excitación de cada uno de los altavoces a 44100 Hz para poder ser reproducida por las tarjetas de audio del sistema. Para el sistema de compensación, primero se han convolucionado los filtros \mathbf{wfs} con los \mathbf{h} , y estos ya, con la señal a reproducir. Dichas señales obtenidas, se pueden reproducir directamente por los altavoces.

A la hora de realizar el análisis subjetivo, se ha desarrollado un programa que permite, tras seleccionar el número de altavoces del sistema, reprodu-

cir por cada uno de ellos sin retardo la excitación correspondiente a cada una de las fuentes secundarias correspondientes a la fuente primaria. Por tanto, se puede situar al oyente delante del conjunto de altavoces y pedirle que compare las dos señales reproducidas, sin filtros de compensación h y con ellos.

En este experimento se pidió a 10 oyentes entrenados en el campo de reproducción espacial de sonido que caminasen por la zona de reproducción. Inicialmente por la zona comprendida por los puntos de control, es decir, la circunferencia de 2.3 m de diámetro y posteriormente que lo hicieran por el resto del área enfrentada a los 32 altavoces.

El objetivo del test es doble. Por un lado, se pretende evaluar si la localización de la fuente virtual se mejora en el área de control tras la compensación de la respuesta. Por otro lado, la respuesta no está controlada fuera de esta área, por lo que cabe pensar que un sujeto podría tener dificultades para percibir la dirección de la fuente virtual si queda fuera de la zona de compensación. Por tanto, la evaluación de este aspecto también resulta necesaria.

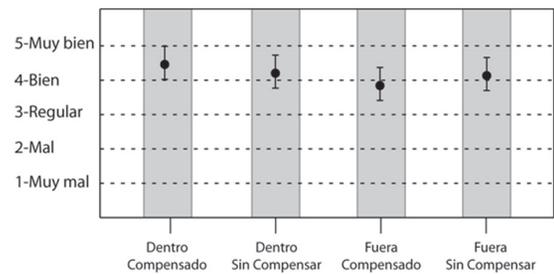
El atributo espacial escogido para evaluar esta percepción es el llamado "locatedness", el cual cuantifica la facilidad con la que un sujeto es capaz de percibir una fuente sonora en una determinada posición espacial, [11]. De esta forma, a los sujetos se les hizo las siguientes preguntas, para evaluar sus respuestas mediante una escala de 5 niveles:

¿Con qué facilidad consigue usted localizar la fuente sonora? ¿Cómo de bien percibe una posición clara de la fuente?

- 1-Muy mal
- 2-Mal
- 3-Regular
- 4-Bien
- 5-Muy bien

La Fig. 7 muestra las medias y los intervalos de confianza al 95% de las respuestas de los sujetos para los cuatro casos de interés, sujeto en el área de escucha con compensación, sin compensación, con compensación y sujeto fuera del área de escucha sin compensación.

FIG. 7. ANÁLISIS SUBJETIVO: MEDIAS E INTERVALOS DE CONFIANZA PARA LAS CUATRO SITUACIONES



Fuente: Los autores

Los resultados obtenidos muestran que la percepción de localización de la señal reproducida es buena en todos los casos. Aunque ésta es ligeramente superior en el caso del área compensada, no parece serlo de forma significativa, lo que hace pensar que al aplicar la compensación, la mejora en la respuesta subjetiva del oyente no es apreciable. Probablemente, esto se debe al entorno donde se ha llevado a cabo el experimento, pues se trata de una sala con acústica controlada donde el efecto de las reflexiones no es tan importante como en otros entornos más reverberantes.

CONCLUSIONES

En este artículo se ha pretendido presentar y dar solución de forma viable al problema de la compensación de salas en los sistemas de reproducción de audio espacial *Wave-Field Synthesis*.

En las simulaciones realizadas en este experimento y en experimentos anteriores, y tras los cálculos de los filtros inversos, se obtuvieron claramente mejoras en el resultado final a la hora de compensar el sistema respecto a su reproducción directa en el interior de salas. No obstante, el análisis inicial subjetivo realizado muestra que en dichas salas, que presentaban reflexiones no muy pronunciadas, los efectos de la sala son prácticamente imperceptibles por el oyente. Por tanto, la comparación de la reproducción mediante el sistema sin compensar respecto al sistema compensado no muestra diferencias claras.

Como se ha comentado, se debe tener en cuenta que este experimento se ha realizado en una sala tratada acústicamente para absorber la mayoría de las reverberaciones, por lo que el sistema de WFS sin compensación funciona considerablemente bien. En la actualidad, se trabaja en otro

tipo de salas, como podría ser una habitación común, donde las reverberaciones son más fácilmente identificables. Por otro lado también se prueba el sistema para un conjunto con un mayor número de altavoces que cierren todo el sistema y permitan abarcar una mayor zona de escucha y también para una zona de control mayor.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el Programa Nacional de I+D+I del Ministerio de Educación y Ciencia con el proyecto TEC2006-13883-C04-01, por la Generalitat Valenciana a través de PROMETEO/2009/013 y por fondos FEDER.

REFERENCIAS

- [1] A.J. Berkhout. "A holographic approach to acoustic control," *Journal of the Audio Engineering Society*, vol.36, pp. 977–995, 1988
- [2] E. Hulsebos, D. de Vries, and E. Bourdillat, "Improved microphone array configurations for auralization of sound fields by wave field synthesis," in *110th Audio Engineering Society Convention*. May 2001, p. Prepr.5337
- [3] S. Spors, H. Buchner, and R. Rabenstein, "Efficient active listening room compensation for wave field synthesis," in *116th Audio Engineering Society Convention*, 2004
- [4] E. Corteel, "Equalization in an extended area using multichannel inversion and wave field synthesis," *J. of the Audio Engineering Society*, vol. 54, no.12, December 2006
- [5] A.J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *Journal of Acoustic Society of America*, vol. 93, no.5, pp. 2764–2778, 1993.
- [6] E. Verheijen, Sound Reproduction by Wave Field Synthesis. PhD thesis, Delft University of Technology, 1997
- [7] P. A. Nelson, F. Orduna-Bustamante, and H. Hamada, "Multichannel signal processing techniques in the reproduction of sound," *Journal of Acoustic Society of America*, vol. 44, pp. 973–989, 1996
- [8] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Transactions Acoustics, Speech and Signal Processing*, vol. 36, pp. 145–152, 1988
- [9] O. Kirkeby, P.A. Nelson, H. Hamada, and F. Orduna-Bustamante, "Fast deconvolution of multichannel systems using regularization," *IEEE Transactions Speech and Audio Processing*, vol. 6, no.2, pp. 189–194, 1998
- [10] L. Fuster, J. Lopez, A. Gonzalez, and P. Faus, "Time and frequency domain room compensation applied to wave field synthesis," in *8th International Conference on Digital audio Effects (DAFx'05)*, Septiembre 2005
- [11] H. Wittek, F. Rumsey and G. Theile, "Perceptual Enhancement of Wavefield Synthesis by Stereophonic Means", *Journal of the Audio Engineering Society*, vol.55, no.9, pp.723-251, September 2007