

Obtención de la Respuesta en Frecuencia Auditiva Usando un Modelo Bio-mimético

O. Gutiérrez, G. Miramontes, E. González, I. de la Rosa, M. A. Araiza

Laboratorio de Procesamiento Digital de Señales,

Universidad Autónoma de Zacatecas, Zacatecas ZAC-98000.

correo-e: osvaldo_gtz_mt@hotmail.com, gmiram@ieee.org, gonzalez_efren@hotmail.com

Resumen — Se propone un modelo matemático con relación a la posición en el espacio de una fuente de sonido. El modelo matemático es función de los ángulos de azimut y elevación, partiendo de la medición de la Respuesta en Frecuencia Relativa a la Cabeza (HRTF) en un modelo biomimético. El modelo biomimético se construye con base en un maniquí (cabeza) que simula a una persona a la cual se le hacen incidir ondas sonoras desde diferentes posiciones en un espacio radial. Se observó que la posición de los polos y ceros del modelo matemático pueden dar información de la localización de la fuente.

Abstract— A mathematical model related to the position of an acoustic source is proposed. The mathematical model is a function of azimuth and elevation angles, starting from the measurement of the Head Related Transfer Function (HRTF) in a biomimetic model. Such a model was constructed based on a manikin (head) that simulates a person who is reached by sound waves from different positions in a radial space. It was observed that the position of zeroes and poles of the mathematical model can give information about the localization of the acoustic source.

Descriptores — *Diferencia de fase interaural (IPD), diferencia de nivel interaural*

(ILD), diferencia de tiempo interaural (ITD), función de transferencia relativa a la cabeza (HRTF).

I. INTRODUCCIÓN

IDENTIFICAR la dirección de un sonido producido por alguna fuente sonora ha sido un problema estudiado ampliamente [1]-[5], [7], [10]-[13].

Los humanos y la mayoría de los animales tienen dos sensores acústicos en su cabeza que son usados para obtener información acerca de las vibraciones del aire y su ubicación; ambos sensores son absolutamente necesarios, dado que lo importante para la localización de una fuente de sonido, es la diferencia entre los niveles percibidos [1]. Con base en ello, se han definido en la literatura, tres conceptos que hacen posible el sentido direccional: la diferencia de tiempo interaural (ITD, por sus siglas en inglés), que es la diferencia entre los tiempos de llegada de la señal a ambos oídos; la diferencia de fase interaural (IPD), que es la diferencia entre las fases con las que llega la señal a cada oído; y la diferencia de nivel interaural (ILD), que es la diferencia entre las amplitudes de la señal en ambos oídos.

Estos conceptos proporcionan información para la localización de una fuente sonora solamente en el plano horizontal. Para el caso de la localización vertical se recurre al análisis del filtrado espectral producido por la oreja (oído externo) y otros elementos de la anatomía humana. Este filtrado puede describirse por medio de una función compleja de respuesta en frecuencia llamada *Función de Transferencia Relativa a la Cabeza* (HRTF, por sus siglas en inglés). Las características espectrales en la HRTF debidas a la difracción y dispersión de la señal producidas por la oreja producen señales para determinar la elevación de la fuente, según Wightman y Kistler en 1997, citados en [5]. Tales características son particularmente notorias en la parte de altas frecuencias (> 5 kHz).

Este trabajo está organizado como sigue, en la sección II se presentan los conceptos básicos sobre la HRTF, en la Sección III se describe el arreglo experimental y el tipo de señal que ha sido utilizada como señal de excitación. En la sección IV se muestran los resultados y su análisis y finalmente en la sección V se dan algunas conclusiones.

II. LA FUNCIÓN DE TRANSFERENCIA RELATIVA A LA CABEZA

La (HRTF) describe cómo una onda sonora de entrada es filtrada debido a las propiedades de difracción y reflexión de la cabeza, oreja y torso, antes de que el sonido alcance el mecanismo transductor del oído interno.

Generalmente, la HRTF para una posición dada de una fuente de sonido, suele calcularse directamente en el dominio de la frecuencia usando un barrido frecuencial de ondas seno, o bien, secuencias de longitud máxima (MLS) [6]. En este trabajo, a diferencia de la mayoría de los reportados en la literatura, se plantea

el uso de una señal multifrecuencia, la cual se describe más ampliamente en [8] y [9].

Para describir el espacio auditivo que envuelve a un oyente se usan tres planos normales, todos referidos al eje interaural, que es una línea imaginaria que une los dos oídos, el plano medio, el plano frontal, y el plano horizontal, ver Figura 1. La posición de una fuente dentro del espacio auditivo se describe usando un sistema de coordenadas de tres componentes, centrado en la cabeza del oyente, Figura 2. El ángulo de azimut θ es el desplazamiento angular desde el plano medio sobre el plano horizontal, con ángulos positivos hacia la derecha y ángulos negativos hacia la izquierda. El ángulo de elevación ϕ mide el desplazamiento vertical desde el plano horizontal, con ángulos positivos hacia arriba y negativos hacia abajo. El alcance r es la distancia desde el centro del eje interaural hasta la fuente [7].

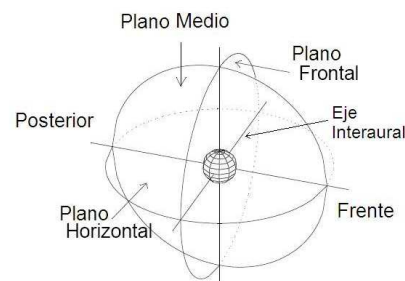


Figura 1. Los tres planos del espacio auditivo.

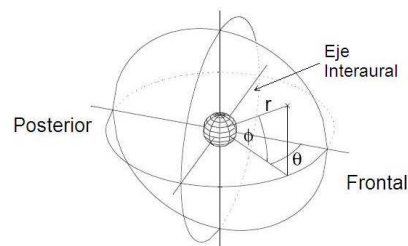


Figura 2. Sistema de coordenadas con azimut, elevación y alcance.

III. ARREGLO EXPERIMENTAL

El sistema consta de un modelo físico basado en un maniquí de unisel al cual se le ha puesto un recubrimiento de latex; su función es simular la cabeza de una persona que escucha. En la Figura 3 se pueden apreciar dos vistas del maniquí utilizado como modelo. Se le adaptaron unos moldes de la oreja contruidos del mismo material, tomando como referencia una oreja de una persona, con la intención de obtener una mayor semejanza con la anatomía humana. En el interior del maniquí se instalaron dos micrófonos, alineados con sus “oídos” los cuales, conectados a un equipo electrónico, permiten capturar la señal de audio para su posterior procesamiento. Para colocar los micrófonos se le hizo un corte al maniquí en la parte posterior de la cabeza. En la Figura 4 se pueden ver los micrófonos instalados. Este modelo se colocó en una estructura metálica que permite variar la posición de la fuente de sonido tanto en azimut como en elevación.



Figura 3. Vistas lateral y frontal del maniquí.

El sistema se colocó en el interior de una cabina de grabación (no profesional) para aislarlo de cualquier ruido distinto de la señal de audio multifrecuencia. En la Figura 5 se muestra la ubicación de todos estos componentes.

A. Señal multifrecuencia

A diferencia de otros trabajos en los que por lo general se utiliza una señal MLS, en



Figura 4. Colocación de los micrófonos.



Figura 5. Instalación del sistema dentro de la cabina.

este trabajo se utilizó una señal multifrecuencia, la cual se genera mediante un algoritmo de optimización [8]. Esta señal reduce el factor pico mediante la suma de M señales senoidales [9]. Los datos de entrada para este algoritmo son:

- N - Número de puntos de la señal producida.
- M - Número de frecuencias que conforman la señal.

Con estos datos y dada una frecuencia de muestreo F_s tenemos que

$$F_{min} = \frac{F_s}{N} \quad F_{max} = \frac{F_s}{N}M \leq \frac{F_s}{2}.$$

Con $F_s=44100$ Hz y tomando $N=2048$

$$F_{min} = \frac{44100}{2048} = 21,5Hz$$

$$F_{max} = (21,5)M = 20000Hz \Rightarrow M \approx 930$$

Así pues, la señal que se produce como resultado cubre muy aproximadamente todo el

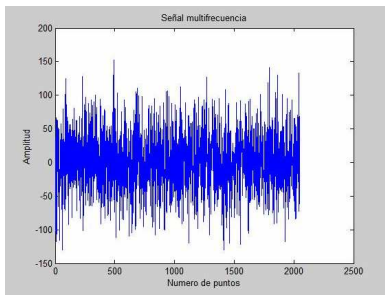


Figura 6. Señal multifrecuencia con 2048 pts. y 930 frecs.

intervalo de frecuencias audibles, y tiene una apariencia como la que se muestra en la Figura 6. Una ventaja importante de esta señal es que tiene una respuesta en frecuencia prácticamente plana en el intervalo de frecuencias de interés.

Una vez generada la señal, el arreglo que contiene los datos que la producen, se convierte en un archivo de audio (.wav) para poder reproducirlo mediante la tarjeta de sonido de la computadora a través de la bocina.

IV. ANÁLISIS Y RESULTADOS

A. Presentación de resultados

Se consideraron un total de 29 posiciones de la fuente sonora en incrementos de 30° , desde un ángulo de azimut de 90° derecha hasta 90° izquierda; y desde un ángulo de elevación de -30° hasta 90° (parte superior de la cabeza). En cada posición se registraron 3 capturas de la señal, que posteriormente fueron promediadas con el fin de reducir el error introducido por el ruido propio del sistema. Se realizó el mismo número de mediciones en dos ocasiones, una con los micrófonos dentro de la cabeza y otra con los micrófonos sin la cabeza, para un total de 174 capturas.

En las figuras 7 a 13 se presenta, en tres dimensiones, el grupo de espectros correspondientes a cada posición de ángulo de azimut. En ellas se puede observar una atenuación de

la señal en la parte de los 10 kHz, que gráficamente se traduce en una “ranura” la cual, conforme aumenta la elevación de la fuente, tiende a disminuir y desaparecer.

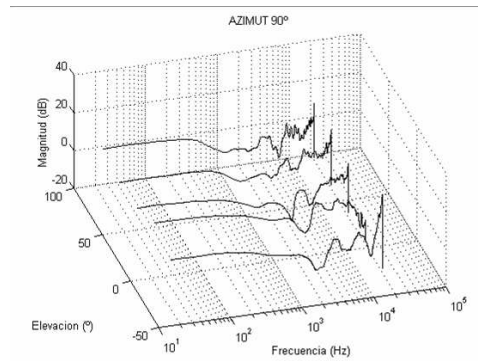


Figura 7. Variación del espectro respecto a la elevación, $az=90^\circ$.

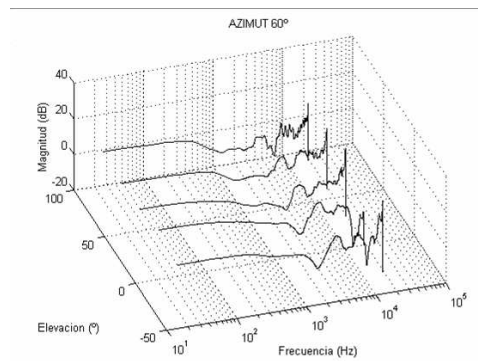


Figura 8. Variación del espectro respecto a la elevación, $az=60^\circ$.

Otro efecto importante que se puede apreciar es una atenuación significativa en el intervalo de frecuencias de 5 a 6 kHz conforme la fuente pasa al lado opuesto del oído de referencia, y se vuelve mayor conforme la elevación aumenta. Las gráficas muestran además, que en el intervalo de frecuencias de 2 a 3 kHz se presenta otro pequeño nivel de atenuación de la señal para elevaciones bajas, y conforme la elevación aumenta, este “valle” desaparece también.

B. Análisis de los datos

Observando los diagramas correspondientes de polos y ceros de cada una de las funciones

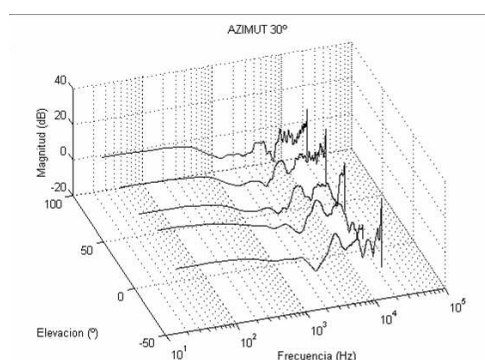


Figura 9. Variación del espectro respecto a la elevación, az=30.

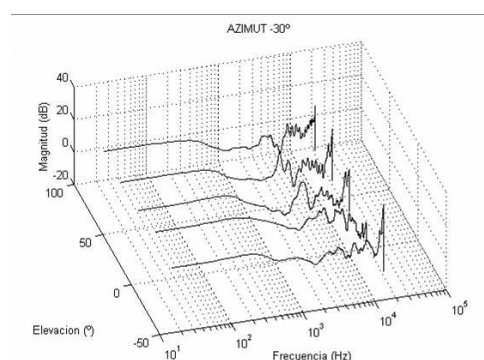


Figura 11. Variación del espectro respecto a la elevación, az=-30.

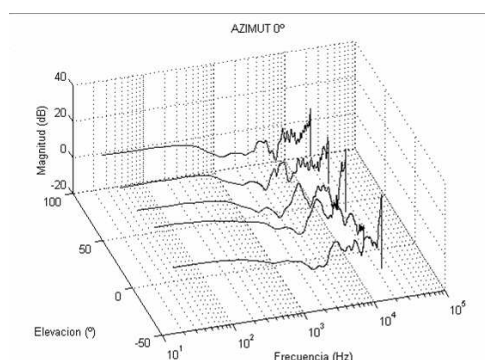


Figura 10. Variación del espectro respecto a la elevación, az=0.

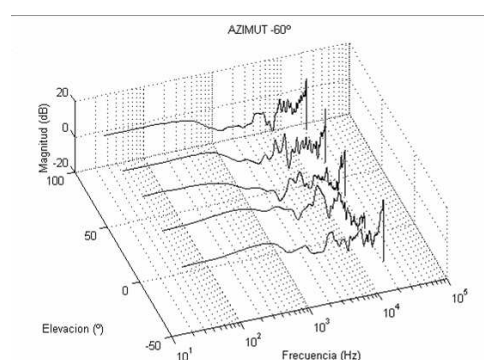


Figura 12. Variación del espectro respecto a la elevación, az=-60.

de transferencia, se puede establecer la dependencia que ellos tienen con respecto a la posición en elevación de la fuente.

En la Figura 14 se puede apreciar un modelo estimado de la respuesta en frecuencia producida por el sistema para una posición de la fuente de $\theta = 30^\circ$ y $\phi = 0^\circ$, y en la Figura 15 se muestra su diagrama correspondiente de polos y ceros, resultado de un modelo en el espacio de estados de la forma:

$$x(t + Ts) = Ax(t) + Bu(t) + Ke(t)$$

$$y(t) = Cx(t) + Du(t) + e(t)$$

Variando el ángulo de elevación, se observa un cambio en la forma del espectro; dicho cambio se traduce en una ubicación distinta de los polos y ceros del modelo matemático.

En la Figura 16 se muestra una nueva estimación. Su correspondiente diagrama de polos y ceros se muestra en la figura 17.

Comparando los diagramas de polos y ceros se puede ver cómo los pares de ceros complejos conjugados aparecen cerca del círculo unitario cuando las atenuaciones están apenas por debajo de 0 dB; sin embargo, aparecen más alejados cuando la magnitud de la atenuación de la respuesta es más significativa, del orden de 5 a 10 dB. Inclusive los pares de ceros complejos más alejados aparecen hacia el lado de alta frecuencia, que es donde en el espectro se presentan gráficamente las ranuras, efecto de la oreja.

Conviene mencionar que los modelos estimados no son exactos; existe bastante margen de diferencia entre la respuesta del sistema

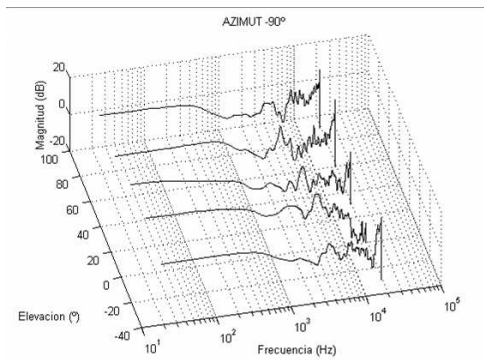


Figura 13. Variación del espectro respecto a la elevación, az=-90.

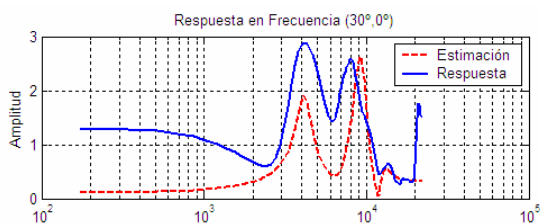


Figura 14. Estimación de modelo matemático $\theta = 30^\circ$ y $\phi = 0^\circ$.

y la respuesta estimada. Aún así, los modelos estimados son bastante complejos, según se observa en los diagramas de polos y ceros correspondientes, lo que se traduce en 10 variables de estado.

V. CONCLUSIONES

Con la realización de este experimento se pudo verificar que, mediante el uso de una cabeza artificial, a la cual se le adaptaron

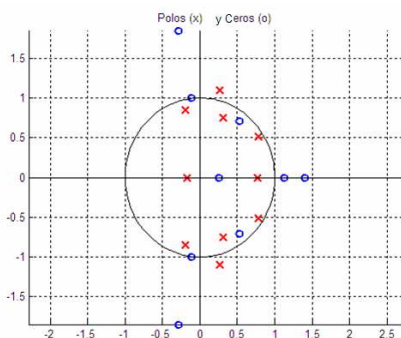


Figura 15. Ubicación de los polos y ceros para el modelo estimado de la Figura 14.

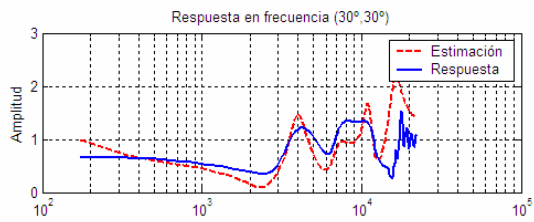


Figura 16. Estimación de modelo matemático $\theta = 30^\circ$ y $\phi = 30^\circ$.

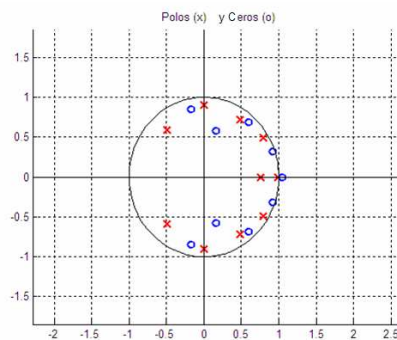


Figura 17. Ubicación de los polos y ceros para el modelo estimado de la Figura 16.

micrófonos en su interior y moldes del oído externo en su superficie, fue posible obtener información sobre la localización del sonido en una distancia corta, y con ello se propuso un modelo matemático para la caracterización del sistema.

En relación al modelo matemático, queda mucho por analizar todavía. Se presentó una propuesta preliminar que puede servir de base para profundizar en el análisis. Se mostró la importancia de la HRTF, ya que ésta contiene la información sobre la posición de la fuente (tanto en azimut como en elevación) y de ahí se obtiene la propuesta del modelo matemático.

RECONOCIMIENTOS

Los autores desean agradecer el apoyo de PROMEP, O. Gutiérrez agradece el apoyo del grupo de PDS de la UAZ, el apoyo de PROMEP con clave P/CA-68 2006-33-20, y a

las autoridades del Instituto Tecnológico Superior de Fresnillo por las facilidades otorgadas para el desarrollo del proyecto.

REFERENCIAS

- [1] A. Kushleyev, V. Vohra, "Sound Localization", Institute for Systems Research, University of Maryland, Summer 2004. Disponible en http://isr.umd.edu/research/REU_reports/2004_PPTs/Kushleyev_Vohra_paper.pdf
- [2] Miller, Joel D., "Modeling Interaural Time Difference Assuming a Spherical Head", Music 150, Musical Acoustics, Stanford University, December 2001. Disponible en <http://home.earthlink.net/~joeld-miller/SLAB/MUS150Paper.pdf>
- [3] V. R. Algazi, R. O. Duda, "Approximating the HRTF using simple geometric models of the head and torso", *J. Acoust. Soc. Am.* 112 (5), Nov. 2002, pp 2053-2064.
- [4] Miramontes et al, "Localización de Fuente Acústica por ERT", Encuentro de Investigación en Ingeniería Eléctrica 2005, Zacatecas, Zac., Marzo 17-18.
- [5] V. C. Raykar, R. Duraiswami, L. Davis, "Extracting significant features from the HRTF", International Conference on Auditory Display, Boston, MA, July 2003.
- [6] <http://en.wikipedia.org/wiki/HRTF>
- [7] P. T. Calamia, "Three Dimensional Localization of a close range acoustic source using binaural cues", Master Thesis, The University of Texas, Austin, Texas, May 1998.
- [8] Farden D., Miramontes G., Tallman D., "DSP-based Instrumentation for Electrochemical Impedance Spectroscopy", 195th Meeting of the Electrochemical Society, Seattle, WA (USA) Vol. 99-5, pp 98-108, Feb 1998.
- [9] G. Miramontes de León, F. Gómez Aguilera, E. García Domínguez, "A DSP56002-based Frequency Response Network Analyzer", 14th International Conference on Electronics, Communications and Computers (CONIELECOMP 2004), 16-18 February 2004, Veracruz, México. pp. 82-86, IEEE Computer Society 2004, ISBN 0-7695-2074-X.
- [10] V. R. Algazi, C. Avendano, R. O. Duda, "Elevation localization and head-related transfer function analysis at low frequencies", *J. Acoust. Soc. Am.* 109 (3), March 2001, pp. 1110-1122.
- [11] V. R. Algazi, R. O. Duda, D. M. Thompson, "The use of head-and-torso models for improved spatial sound synthesis", Presentado en la 113^a Convención de la Audio Engineering Society, Los Angeles, California, October 2002.
- [12] I. Rossell, R. Barti, D. Sansegundo, "Diseño e implantación de un sistema de sonido tridimensional basado en la HRTF", Departamento de Acústica Ingeniería La Salle, Universidad Ramón Llull, Barcelona, España.
- [13] V. R. Algazi, R. O. Duda, D. M. Thompson, "The CIPIC HRTF database", IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2001.