

## **MODELO ACÚSTICO DE CABEZA Y TORSO MEDIANTE ANÁLISIS DE COMPONENTES PRINCIPALES**

Oscar A. Ramos<sup>a</sup>, Goncal Calvo<sup>b</sup>, Fabián C. Tommasini<sup>a</sup>

*<sup>a</sup>Centro de Investigación y Transferencia en Acústica, CINTRA. UA-CONICET UTN, Regional Córdoba.*

*Mtro. López esq. Cruz Roja Argentina, (5016) Córdoba, Argentina .oramos@scdt.frc.utn.edu.ar*

*<sup>b</sup>Universidad Politécnica de Cataluña Barcelona, España*

**Palabras clave:** funciones de transferencia de cabeza, realidad acústica virtual, análisis de componentes principales.

**Resumen.** La respuesta impulsiva en campo libre entre una fuente sonora y los tímpanos de un oyente (Head Related Impulse Response: HRIR), engloba todas las transformaciones asociadas con las claves determinantes para la percepción auditiva del entorno. Estas transformaciones son consecuencia de la interacción de la onda sonora con los hombros, la cabeza y los pabellones auriculares. Ha sido de interés para la Tecnología Binaural y la Simulación Acústica Virtual modelar *eficazmente* este comportamiento de manera de poder sintetizar estímulos sonoros, que escuchados mediante auriculares, evoquen en el usuario la sensación de presencia en el entorno modelado.

El propósito de este estudio es crear un modelo reducido que permita reconstruir una HRIR de un conjunto, mediante combinación lineal de unas pocas funciones básicas. Se consigue dicha reducción aplicando el método de análisis de Componentes Principales (PCA) a las Funciones de Transferencias de Cabeza (HRTF) complejas derivadas de un conjunto de HRIR (5050) -medidas en cinco individuos- por el Dr. Fred Wightman de la Universidad de Wisconsin.

Resultados preliminares de pruebas psicoacústicas demuestran que la utilización de pocas funciones básicas son suficientes para preservar las características fundamentales de fase y magnitud de las HRIR originales sin degradar la percepción acústica del entorno.

Esta reducción de los datos originales hace al método interesante para su aplicación en sistemas de tiempo real.

## 1 INTRODUCTION

El propósito de la Realidad Acústica Virtual es recrear la experiencia auditiva que una persona experimentaría en un entorno real, provocando en ella la sensación de presencia en dicho entorno.

El principio que sustenta la simulación acústica establece que: estímulos acústicos equivalentes, evocan sensaciones equivalentes (Blauert, 1997). En otras palabras, si mediante auriculares (binaural) o altoparlantes (transaural) se aplican a los tímpanos de un oyente las señales biológicamente correctas, será posible estimular en el mismo, la sensación de *inmersión* en el entorno modelado (Kleiner et. al., 1993).

En la simulación acústica tres subsistemas deben ser modelados: la fuente sonora, el recinto y el oyente. La fuente sonora se especifica por sus características de *directividad* y respuesta en frecuencia; el recinto, por la respuesta impulsiva entre una fuente sonora y un receptor (varía con la ubicación de la dupla fuente-receptor) y el oyente, por las Respuestas Impulsivas de Cabeza (HRIR, en ingles).

Una onda sonora, antes de alcanzar los tímpanos de un oyente, sufre una serie de transformaciones al interactuar con el torso, los hombros y especialmente con la cabeza y los pabellones auriculares. Son diferentes para cada oído y dependen del ángulo de incidencia de la onda. Las HRIR contienen dichas transformaciones.

Ha sido demostrado que aspectos bien definidos de las ondas sonoras que llegan a los tímpanos, llevan información que un ser humano utiliza para localizar con precisión una fuente sonora en el espacio (Blauert, 1997; Wightman and Kistler, 1989b).

Ahora bien, no hay resultados concluyentes sobre la verdadera naturaleza de tales transformaciones ni como el Sistema Auditivo las interpreta. Luego, no está claro lo que un modelo de dichas transformaciones debe contemplar.

Desde fines de los 80 hasta nuestros días, la implementación *eficaz* de modelos del oyente acaparó el interés de numerosos investigadores. Algazi y otros (2001b, 2002), emplean modelos teóricos tanto de la cabeza como de las orejas y modelan estos elementos con filtros paramétricos lineales. Han validado objetivamente los filtros implementados con mediciones de las ondas sonoras en cabezas y oídos artificiales. Otros autores, utilizan filtros de distintos tipos y orden (FIR y IIR) que son ajustados a mediciones de las HRIR (Wightman and Kistler, 1989a; Kulkarni, and Colburn, 2004; Ramos, et. al., 2006). Finalmente, un tercer abordaje, apela a deducir modelos funcionales o ecuaciones obtenidas a partir de HRIR medidas (Chen et. al. 1995) ó métodos estadísticos-matemáticos (Kistler and Wightman, 1992; Scarpaci and Colburn, 2005).

Es propósito del presente trabajo, es crear un modelo reducido del oyente mediante el método de Análisis de Componentes Principales (PCA), a partir de HRIR medidas a 5 sujetos en la Universidad de Wisconsin por el Dr. Fred Wightman. El modelo que se obtiene son aproximaciones o estimaciones de las HRIR reales.

El trabajo está organizado de la manera siguiente. En el punto 2 se profundiza el análisis de las HRIR; en el punto 3, se introduce al lector en las bases del Análisis de Componentes Principales y su aplicación a nuestro caso; en el punto 4 se realiza un análisis objetivo de los resultados y en el punto 5, se discuten los resultados de pruebas psicoacústicas preliminares.

## 2 LAS RESPUESTAS IMPULSIVAS DE CABEZA

El estudio de las alteraciones que se producen en la cabeza y el torso se remonta al siglo pasado cuando Lord Rayleigh enunció la teoría Duplex.

Firestone (1930) realizó un abordaje teórico experimental considerando a la cabeza como una esfera rígida y analizó los niveles de presión sonora en dos polos separados por 180

grados emulando la ubicación de los oídos. Concluyó que la diferencia de nivel entre ambos polos -originada por difracción provocada por la cabeza- es función de la frecuencia y de la dirección de la onda incidente. La diferencia de nivel entre el oído izquierdo y el derecho se conoce como Diferencia de Nivel Interaural (ILD, en inglés). Las ILD son despreciables por debajo de los 500 Hz (la longitud de onda es menor al diámetro de la cabeza) y crecen rápidamente hasta alcanzar su valor máximo alrededor de los 4000 Hz cuando la fuente se encuentra enfrentada a uno de los oídos.

Continuando con el modelo esférico de la cabeza, Kuhn (1977) realizó estudios teóricos, que validó con mediciones, sobre la diferencia de los tiempos de arribo de la onda sonora a ambos oídos. Dedujo que para todas las frecuencias audibles dependían solo del ángulo de incidencia, pero diferían según la región de frecuencias considerada: eran mayores para bajas frecuencias (hasta 500 Hz); mínimos para frecuencias medias (entre 1000 y 1500 Hz) y aumentaban para frecuencias altas (> a 3000Hz). Esta diferencia se conoce como Diferencia de Tiempo Interaural: ITD.

Además de las dos evidencias mencionadas, las ondas sonoras se dispersan, refractan y cancelan en las protuberancias y cavidades de los pabellones auriculares (orejas). Esto da origen una serie de picos y valles en la zona de frecuencias medias y altas -entre los 4000 y 12000 Hz- que se desplazan dependiendo del ángulo de incidencia de la onda sonora. Las reflexiones y difracciones producidas en el torso y hombros afectan también esta zona del espectro (Shaw, 1974; Kuhn, 1977). Las variaciones en el espectro de las ondas que alcanzan ambos tímpanos pueden ser descritas como un proceso de filtrado natural que depende de la dirección de incidencia y de la anatomía del individuo.

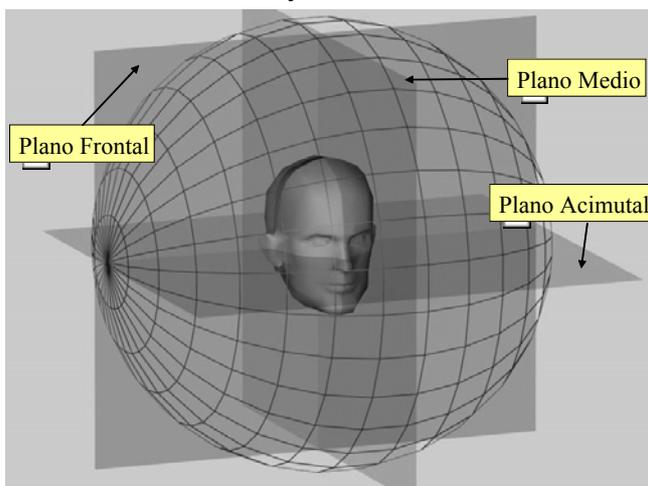


Figura 1: Referencias del sistema de coordenadas

La ILD, ITD y las transformaciones espectrales son las características de las ondas que el ser humano utiliza para localizar una fuente sonora en el espacio y se las denomina claves perceptuales.

Las HRIR se obtienen mediante la medición en cámara anecoica de la respuesta impulsiva entre una fuente sonora y ambos tímpanos de un oyente. Se determinan para diferentes posiciones discretas de la fuente ubicada sobre una superficie esférica de 1,5 metros de radio. En la proximidad de los tímpanos del sujeto -situado en el centro de la esfera- se

colocan micrófonos miniaturas; como fuente, se emplean parlantes y como señales impulsos angostos, series de máxima longitud (Rife et. al. 1989) ó códigos de Golay (Zhou et. al. 1992).

Para definir la ubicación de la fuente se utiliza un sistema de coordenadas esféricas. El origen es la intersección de tres planos: el plano horizontal o acimutal, que contiene una línea imaginaria que une ambos oídos; el plano medio, que divide la cabeza simétricamente y el plano frontal, perpendicular a los dos anteriores (Figura 1). Cada posición esta definida por los ángulos de acimut  $\varphi$  y elevación  $\Phi$ . Si  $\varphi=0$  y  $\Phi=0$ , la fuente está ubicada al frente del sujeto. Si  $\varphi$  es negativo la fuente está a la izquierda del plano medio y si  $\varphi$  es positivo a la derecha. De igual manera, si  $\Phi$  es negativo la fuente está por debajo del plano horizontal y si

$\Phi$  es positivo por arriba. A partir de ahora definiremos la posición de la fuente por el par (acimut, elevación), esto es, por ejemplo,  $(-90,0)$ .

Como se vio, el plano medio divide la cabeza del sujeto en dos hemisferios simétricos. Al oído que está en el mismo hemisferio que la fuente se lo denomina ipsilateral, mientras que al oído opuesto: contralateral.

Las transformadas de Fourier de las HRIR, se conocen como Funciones de Transferencia de Cabeza (HRTF). En la [Figura 2](#) se muestran las respuestas impulsivas y las magnitudes espectrales de ambos oídos para una posición en particular.

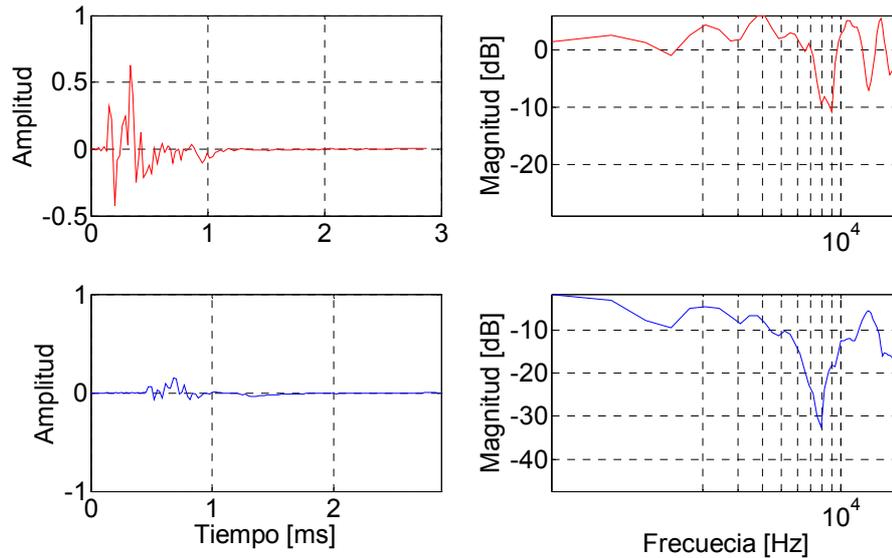


Figura 2: Respuesta impulsiva y magnitud del oído derecho (rojo) y del oído izquierdo (azul) para una fuente sonora ubicada en  $\varphi=45$  y  $\Phi=0$ .

## 2.1 Características de las HRIR

Las HRIR medidas pueden ser asimiladas como la respuesta al impulso de un sistema LTI (Lineal Time Invariant). Tales sistemas pueden ser descompuestos en dos: un sistema de fase-mínima y un sistema “all-pass” (Oppenheim, 1989). La magnitud del sistema resultante es igual a la magnitud del sistema de fase-mínima (la magnitud del “all-pass” es igual a 1) mientras que la fase es igual a la del sistema de fase-mínima más la del sistema “all-pass”.

$$H_{med} = |H_{med}| e^{j\vartheta_{med}} = H_{min} H_{ap} = |H_{min}| e^{j(\vartheta_{min} + \vartheta_{ap})} \quad (1)$$

$$\vartheta_{med} = \vartheta_{min} + \vartheta_{ap} \quad (2)$$

Los subíndices significan: *med*: medidas; *min*: fase-mínima y *ap*: “all-pass”.

En un sistema de fase-mínima, el logaritmo de la magnitud del espectro y su fase, están relacionadas por la transformada de Hilbert:

$$\vartheta_{min} = \text{imag}\{H[-\text{Ln}(|H_{med}|)]\} \quad (3)$$

La respuesta impulsiva de fase-mínima se reconstruye utilizando la ecuación (3) y la magnitud del espectro de las HRTF medidas (igual a la magnitud de fase-mínima). La componente de fase del *ap* se calcula despejándola de la ecuación (2).

En la [Figura 3](#) se ven las componentes de fase de las HRTF de la [Figura 2](#), en donde se observa que la fase *ap* es casi lineal para todo el rango de frecuencias.

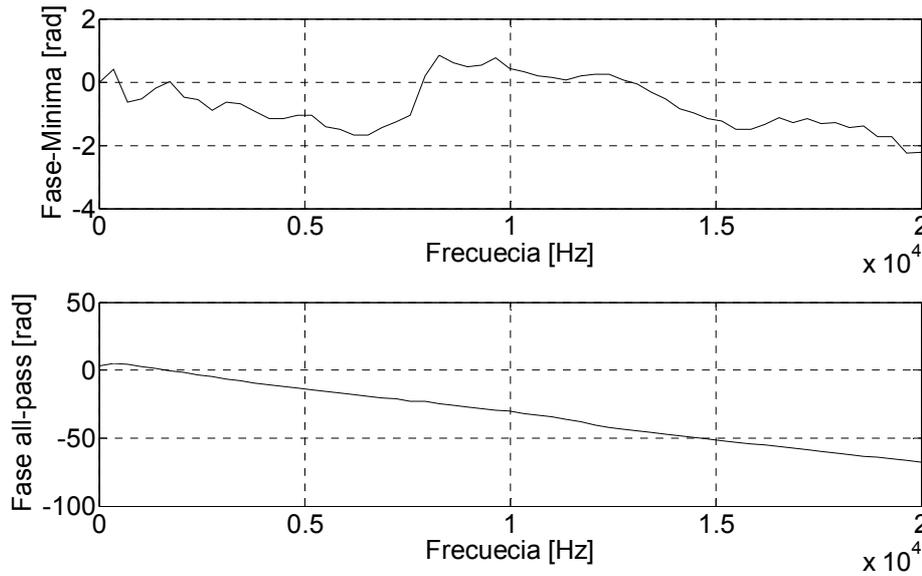


Figura 3: Componentes de fase de la HRTF del oído derecho para una fuente sonora ubicada en  $\varphi = 45$  y  $\Phi = 0$ .

Esta evidencia, anticipada ya por los estudios de Kunh (modelo esférico de la cabeza) más el hecho de que las HRTF son funciones de fase-mínima (Mehrgardt and Mellert, 1977), permitieron desarrollar un modelo simplificado de las HRTF conocido como: fase-mínima-más-fase-lineal.

La simplificación consiste en descartar la componente no lineal del *ap* y suponer que la fase del sistema *ap* es sólo lineal. Entonces, la fase *ap* será de la forma:

$$\varphi_{ap} \approx -\omega\tau_{ap} \quad (4)$$

En donde  $\tau_{ap}$  es un retardo constante, independiente de la frecuencia.

En la práctica, en lugar de calcular este retardo para cada oído, se determina el retardo entre ambos oídos que se agrega a la respuesta impulsiva de fase-mínima del oído que corresponda. Resumiendo, una HRIR medida, se puede reconstruir desde las HRIR de fase mínima (que en general son secuencias más pequeñas) más los retardos interaurales para cada posición.

Estudios psicoacústicos (Kistler and Wighthman, 1992; Kulkarni, at. al 1999) han corroborado que asumir que las HRTF responden a un sistema de fase-mínima más un retardo constante para todo el rango de frecuencia es, en general, aceptable para aplicaciones no críticas.

Estudios más recientes demuestran que la fase de las HRTF para bajas frecuencias (hasta 1000 Hz) cumplen un rol dominante y que asumir que es lineal, sobre todo para el oído contralateral, conduce a errores insalvables en aplicaciones críticas (Scarpaci at. al 2005; Zahorik at. al. 2006).

El modelo propuesto en este trabajo tiene en cuenta tanto la magnitud como la fase de las HRTF medidas y a su vez permite reducir la dimensión de los datos requeridos.

### 3 ANÁLISIS DE COMPONENTES PRINCIPALES (PCA)

Matemáticamente hablando, PCA, es una transformación lineal que permite, por un lado, reducir la dimensión de un conjunto de datos empíricos u observaciones, y por el otro, extraer aspectos comunes del conjunto original.

Consiste en proyectar el conjunto de datos originales sobre un conjunto de vectores ortonormales o funciones básicas de dimensión menor en el sentido en que la covarianza común es máxima.

En nuestro caso, las HRIR son secuencias de 256 puntos (muestreadas a 44100 Hz). Fueron medidas en 505 posiciones alrededor de una esfera de 1,5 metros de diámetro con una resolución de 10 grados. Las elevaciones varían entre -50 grados (debajo del plano horizontal) y +90 grados (encima de la cabeza) mientras que los acimuts varían entre 0 y -180 grados (a la izquierda del plano medio) y entre 10 y 170 grados (a la derecha del plano medio).

Los datos de entrada son las transformadas de Fourier de las HRIR de los 5 individuos. Debido a que el ancho de banda de interés es de 15000 Hz, se utilizaron las primeras 88 componentes en frecuencias. La dimensión de la matriz de cada sujeto es: 505 (posiciones) x 2 (oídos) = 1010 filas por 88 columnas de frecuencias. Por lo tanto, la matriz global con los datos de los 5 sujetos será: 5x1010 filas por 88 columnas = 444.400. Denominaremos a esta matriz  $H_{pf}$  de dimensión (p, f).

Para obtener las componentes principales, se llevó a cabo los pasos siguientes:

1. Se obtuvo la media global según  $p$  de la matriz  $H$ :

$$H_0(f) = \frac{1}{p} \sum_{i=1}^{i=p} H_i(f) \quad (5)$$

2. Se centró los datos respecto a su media

$$D_i(f) = H_i(f) - H_0(f) \quad (6)$$

3. Se calculó la covarianza de  $D$ :

$$C = \frac{1}{p} \sum_{i=1}^{i=p} D_i(f) D_i(f)^* \quad (7)$$

4. Se obtuvo los eigenvectores  $Q_i$  y los eigenvalores de la matriz  $C$  mediante métodos de álgebra lineal.

5. Se calculo los pesos  $W_i$  de las funciones básicas proyectando  $D$  en  $Q$ :

$$W_i = Q_i(f)' D(f) \quad (8)$$

Para reconstruir totalmente una HRTF en particular será:

$$\hat{H}(f) = H_0(f) + \sum_{i=1}^N Q_i(f) W_i \quad (9)$$

en donde  $N$  es la totalidad de las funciones básicas. En los puntos siguientes se determinará la cantidad  $L$  de funciones básicas menor a  $N$ , necesarias para reconstruir una HRTF perceptualmente indiferenciable de la original.

## 4 ANÁLISIS OBJETIVO DE RESULTADOS

Se evaluó la magnitud del espectro, la fase y la respuesta impulsiva reconstruidas a partir de las primeras 20 funciones básicas (99.86 % de la varianza total). Para la discusión se utiliza el sujeto denominado *SOS*, habiéndose comprobado comportamientos similares con los restantes.

### 4.1 Magnitud del Espectro

Para estimar el error entre las magnitudes de las HRTF's medidas y las reconstruidas por el modelo, se empleó la fórmula siguiente:

$$Error = \sqrt{\frac{1}{n} \sum_{n=1}^{n=88} [20 \log_{10} |H_m(f)| - 20 \log_{10} |H_r(f)]^2} \quad (10)$$

Esto es, la raíz cuadrada de la media de las diferencias de los logaritmos base 10 de las magnitudes medidas ( $H_m$ ) y las reconstruidas ( $H_r$ ) elevadas al cuadrado, para todo el rango de frecuencias (Kulkarni, 2004).

En la [Figura 4](#) se muestran los errores para ambos oídos según la ubicación de la fuente. Como un ejemplo ilustrativo se grafica también, las magnitudes medidas y las reconstruidas para la posición (90, 0).

Los errores son mayores para el oído derecho cuando la fuente está ubicada a la izquierda (acimuts negativos); lo contrario ocurre con el oído izquierdo. Crecen para elevaciones negativas (debajo del plano horizontal) y disminuye para elevaciones positivas. Esta tendencia sistemática del error se debe, por un lado, al fenómeno de difracción que produce la cabeza en el oído oculto; y por el otro, a la contribución que hacen las reflexiones que se producen en el torso y hombros, cuando la fuente está por debajo del plano horizontal. Ambos fenómenos hacen que el perfil espectral de las magnitudes sea más irregular para el oído contralateral degradando el ajuste. Esta tendencia de los errores fue observada en todos los sujetos.

Estas evidencias se manifiestan en el ejemplo (90, 0). El grado de ajuste de la magnitud del oído derecho enfrentado a la fuente, es notablemente superior al del oído izquierdo.

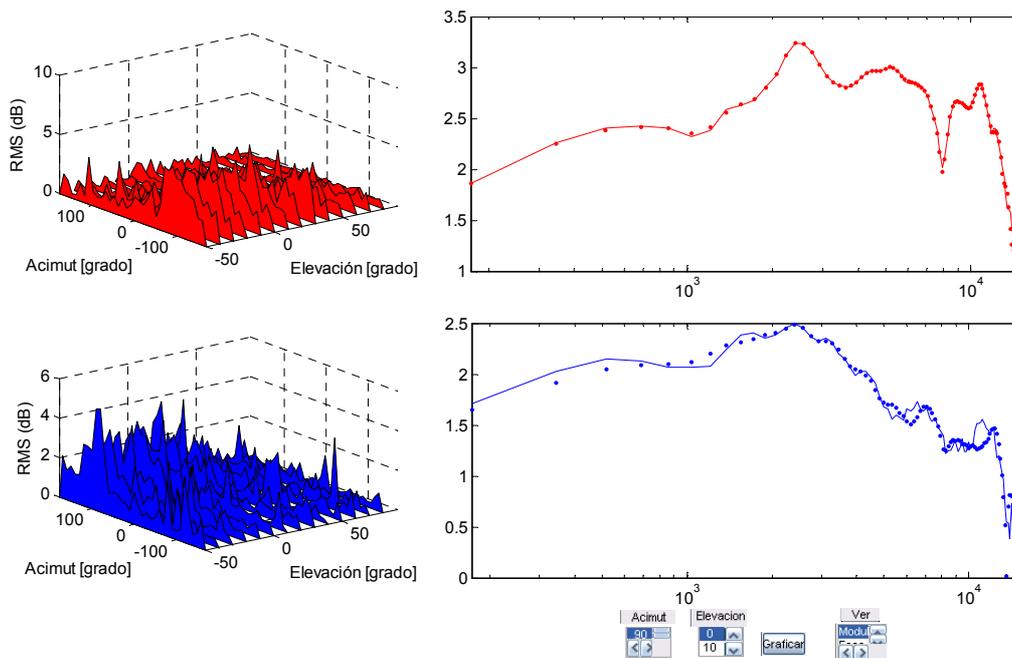


Figura 4: Errores de las magnitudes para el oído derecho (rojo) y para el oído izquierdo (azul). En la columna de la derecha se ve como el grado del ajuste entre la HRTF medida (línea continua) y la reconstruida (puntos) es superior para el oído derecho ( $\varphi = 90$  y  $\Phi = 0$ )

## 4.2 Fase

Los errores de fase fueron calculados de igual manera que los errores de las magnitudes. La fase fue calculada como la tangente inversa de la componente imaginaria dividida por la componente real.

La [Figure 5](#) muestra los resultados. Se observa que los errores están concentrados en las elevaciones y en los acimuts negativos para el oído derecho y en los acimuts positivos para el oído izquierdo. Como ya se habrá advertido, los errores siguen la misma tendencia que el de las magnitudes, con la particularidad de que su variación con la posición es abrupta y no gradual como lo es con las magnitudes. Si se presta atención al ejemplo en la columna de la derecha, se puede comprobar el excelente ajuste de la fase para el oído derecho (ipsilateral) mientras que para el izquierdo (contralateral) el ajuste es bueno en bajas y medias frecuencias.

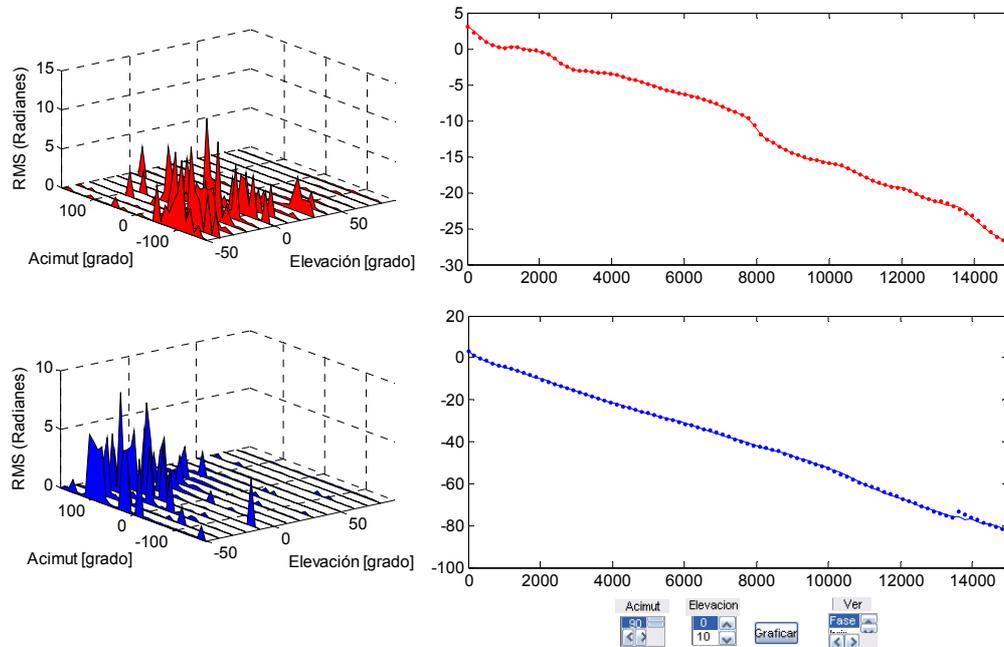


Figura 5: Error de fase para el oído derecho (rojo) y para el oído izquierdo (azul). En la columna de la derecha se ve el ajuste excelente en bajas y medias frecuencias ( $\varphi = 90$  y  $\Phi = 0$ )

### 4.3 Respuesta Impulsiva

Para evaluar las respuestas impulsivas calculamos la correlación cruzada normalizada  $c$ , entre las HRIR medidas y las reconstruidas por el modelo. Este índice de similitud o coherencia expresan los diferentes grados de correlación entre dos señales. Los valores de  $c$  están comprendidos entre 0 y 1; se dice que dos señales son coherentes o idénticas cuando  $c = 1$ ; en caso contrario si  $c = 0$ , se habla de señales incoherentes o diferentes. De alguna manera este índice engloba tanto las diferencias de magnitud como las de fase vistas más arriba.

En la [Figura 6](#) se muestran los resultados. Como era de esperar, el índice de coherencia tiene el mismo comportamiento que los errores de fase y magnitud: es apenas menor que 1, para el oído contralateral y por debajo del plano horizontal y muy próximo a 1 en las posiciones restantes.

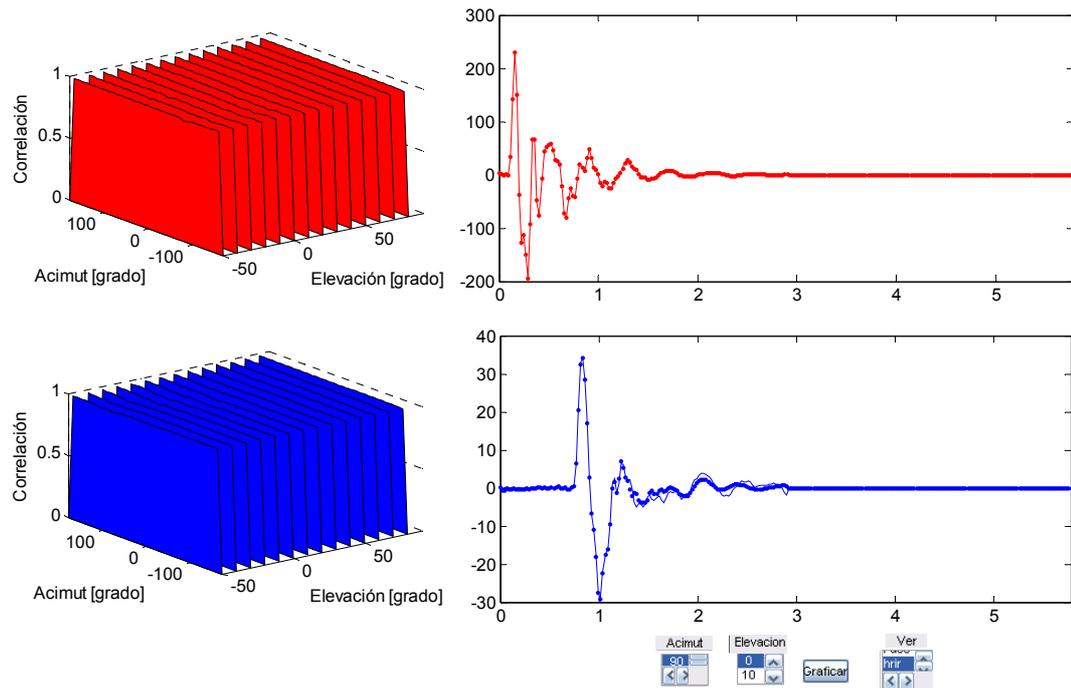


Figura 6: Correlación cruzada entre las HRIR medidas y las reconstruidas por el modelo. En la columna de la derecha se observa un ejemplo para la fuente ubicada en  $\varphi = 90$  y  $\Phi = 0$

A modo de conclusión de este punto, diremos que: se observó en los cinco sujetos una tendencia similar de los errores calculados, esto es: son mayores cuando la fuente sonora esta ubicada debajo del plano horizontal (elevaciones negativas) y en el oído contralateral. Para cada sujeto, los valores del error del oído derecho son diferentes a los del izquierdo debido a la asimetría propia de cada persona respecto al plano medio (Figura 2). Estas asimetrías tienen que ver con que la cabeza dista de ser esférica; la entrada a los conductos auditivos no están separados 180 grados; diferencias de tamaños de los pabellones auriculares, entre otras.

De igual manera, los valores de los errores difieren entre sujetos. Se comprobó que estas diferencias, en algunos casos, son perceptualmente detectables y que los resultados de las pruebas psicoacústicas difieren según las HRIR del sujeto que se utilice.

En la Figura 7, se grafica los errores de magnitud de ambos oídos versus el orden del modelo (número de funciones básicas utilizadas en la reconstrucción) para la posición (20,0) y en la Figura 8 para la posición (90,0). Estas dos posiciones tiene la particularidad de que, el oído derecho (ipsilateral) siempre está enfrenteado a la fuente, mientras que el oído izquierdo (contralateral) siempre está oculto por la cabeza, pero en grados diferente: la sombra es máxima en la posición (90,0) y mínima en la posición (20,0).

El error es pequeño en ambos casos y para el oído derecho, disminuyendo rápidamente con el orden. Por el contrario, el error para el oído izquierdo es sensiblemente mayor para la posición (90,0) (la *sombra* es máxima) y algo menor para (20,0) (la *sombra* es parcial). Los errores se estabilizan según la posición: para la (20,0) en modelos de orden 20 y para la posición (90,0), en modelos de orden 30. Observando las curvas de las diferencias (trazos en negro) se infiere que los errores del oído contralateral, oculto por la cabeza, gobierna el orden del modelo.

Desde el punto de vista perceptual, no hay acuerdo en la literatura sobre lo que podríamos llamar un *error aceptable*. En el punto siguiente, se avanzará en ese sentido.

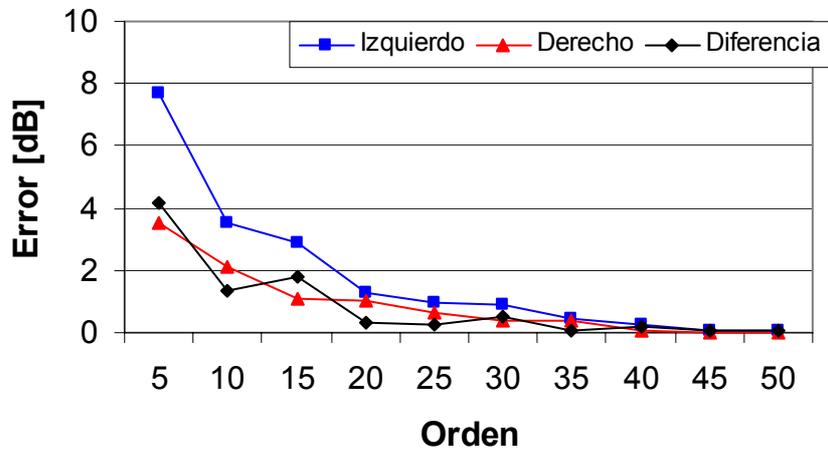


Figura 7: Error entre la magnitud de la HRTF medida y la reconstruida, vs. el orden del modelo para  $\varphi = 20$  y  $\Phi = 0$

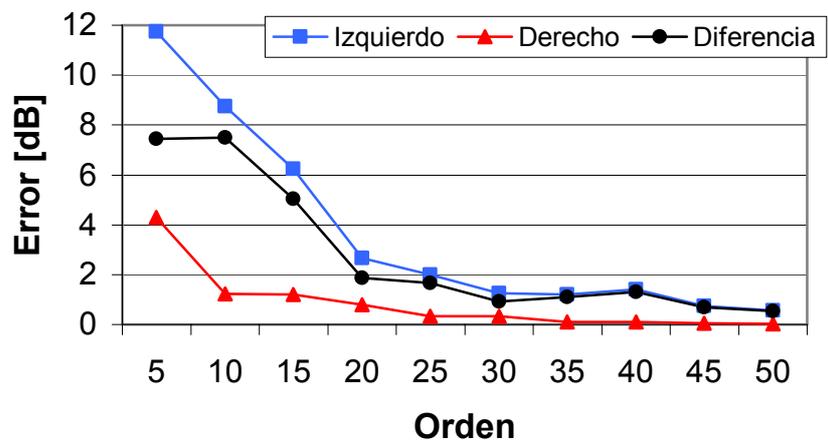


Figura 8: Error entre la magnitud de la HRTF medida y la reconstruida, vs. el orden del modelo para  $\varphi = 90$  y  $\Phi = 0$

## 5 ANÁLISIS PSICOFÍSICOS DE LOS RESULTADOS

El modelo fue examinado en una serie de experimentos psicofísicos. Los resultados que se presentan corresponden a pruebas pilotos, realizadas a un sujeto normal-oyente con entrenamiento musical.

Se utilizó un paradigma de discriminación entre pulsos de ruido gaussiano filtrados con las HRIR medidas y las reconstruidas por el modelo. El paradigma empleado fue de elección forzada de 4-intervalos y 2-alternativas (4I-2AFC). Al sujeto se le presentó, mediante auriculares, una secuencia de cuatros sonidos o estímulos de 200 milisegundos de duración,

separados por 250 milisegundos de silencio. Tres de los cuatro sonidos, se filtró con las HRIR medidas. El cuarto, que al azar ocupaba el segundo o el tercer lugar, se procesó con la HRIR reconstruida. La tarea que se encomendó al sujeto fue: dilucidar la ubicación que ocupaba el sonido filtrado con la HRIR reconstruida.

Los sonidos se reprodujeron mediante auriculares (Philips HE036) a un nivel confortable. Al comienzo y al final del ruido se les aplicó rampas cosenos de 25 ms cada una. Para evitar que el sujeto se familiarizara con el estímulo, la muestra del ruido era diferente para cada ensayo. Las posiciones estudiadas fueron (20,0) y (90,0) y el orden de los modelos se varió entre 10 y 35 en incrementos de 5. Cada posición se repitió 10 veces; por lo tanto, una prueba constaba de 20 ensayos (2 posiciones x 10 repeticiones). Se realizaron pruebas en condiciones monoaurales (un solo oído) y binaurales, con la finalidad de discernir la importancia de los errores calculados en ambas posiciones.

En los experimentos monoaurales, el sujeto escuchaba con sólo un auricular (el derecho o el izquierdo). Por el contrario, en las binaurales, el sujeto escuchaba por ambos auriculares. En la [Figura 9](#) y la [Figura 10](#) se presentan los resultados para la posición (20,0) y (90,0) respectivamente.

El eje de absisas está graduado en porcentajes de respuestas correctas o aciertos. Un porcentaje del 100 % significa que el sujeto pudo discriminar en la totalidad de los ensayos (diez), el estímulo procesado con la HRIR reconstruida. Un porcentaje de respuestas correctas del 50% o menos, significa que el sujeto no pudo discriminar entre un estímulo procesado con la HRIR medida de uno procesado con la reconstruida (el 50% de dos opciones corresponden al azar).

Se puede inferir una marcada similitud entre los gráficos de los errores (Figuras 7 y 8) con los gráficos de las pruebas psicofísicas. Se recuerda que en ambas posiciones, el oído derecho (ipsilateral) siempre está enfrentado a la fuente, mientras que el oído izquierdo (contralateral) siempre está oculto por la cabeza, pero en grados diferente.

En las pruebas monoaurales realizadas con el oído derecho, revelan que el orden del modelo requerido (respuestas próximas al 50%) es entre 10 y 15. Observando los gráficos de los errores de las Figuras 7 y 8, se ve que corresponden a 2 decibeles o menos. Las pruebas realizadas con el oído izquierdo, muestran comportamientos diferentes según la posición. Para la posición (90,0) los sujetos pudieron discriminar en el 100% de los casos para ordenes del modelo entre 10 y 15; el 60% para orden 20 y por debajo del umbral del azar para modelos mayores a 25. Se comprueba que los errores correspondientes a esos órdenes, son menores a 2 decibeles. Para la posición (20,0) los porcentajes de discriminación son menores para todos los órdenes examinados. El umbral del azar se alcanza para modelos de orden 20.

Los resultados de las pruebas binaurales están determinados por los resultados monoaurales para el oído izquierdo. Para la posición (20,0) y para modelos de orden 20, el porcentaje de aciertos sigue las variaciones del oído izquierdo, incluso ligeramente por encima del umbral del azar. El mismo comportamiento se deduce de la Figura 10 para la posición (90,0).

A partir de los resultados de estas pruebas preliminares se puede afirmar que el orden adecuado para reconstruir una HRIR medida está comprendido entre 20 y 25. En consecuencia, la reducción de la dimensionalidad de los datos originales es considerable. Se debe recordar que los datos originales, para cada sujeto, se componen de 505 respuestas impulsivas de 256 puntos para cada oído:  $505 \times 256 \times 2 = 258560$ ; mientras que el modelo de orden 20 requiere de 20 funciones básicas de 88 puntos:  $20 \times 88 = 1760$ ; más 20 coeficientes por posición y oído:  $20 \times 505 \times 2 = 20200$ , dando un total de 21960 elementos. La reducción es mayor a 10 veces.

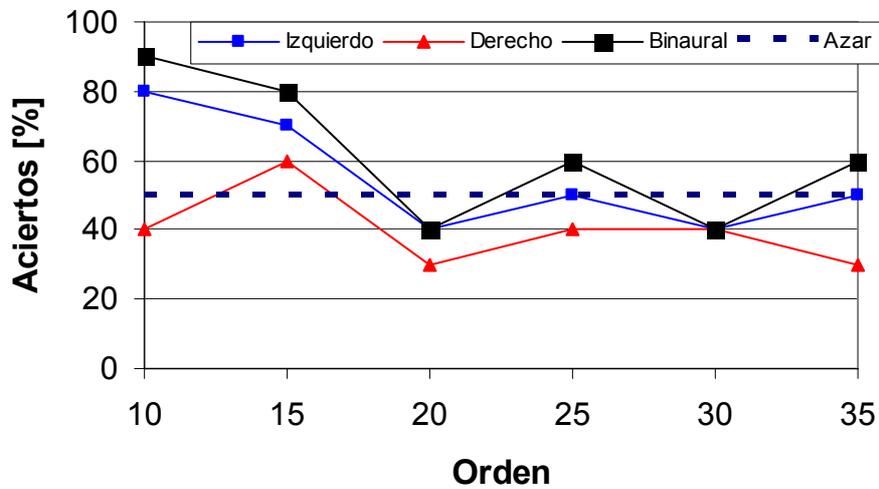


Figura 9: Porcentajes de aciertos vs. el orden del modelo para la posición  $\varphi = 20$  y  $\Phi = 0$

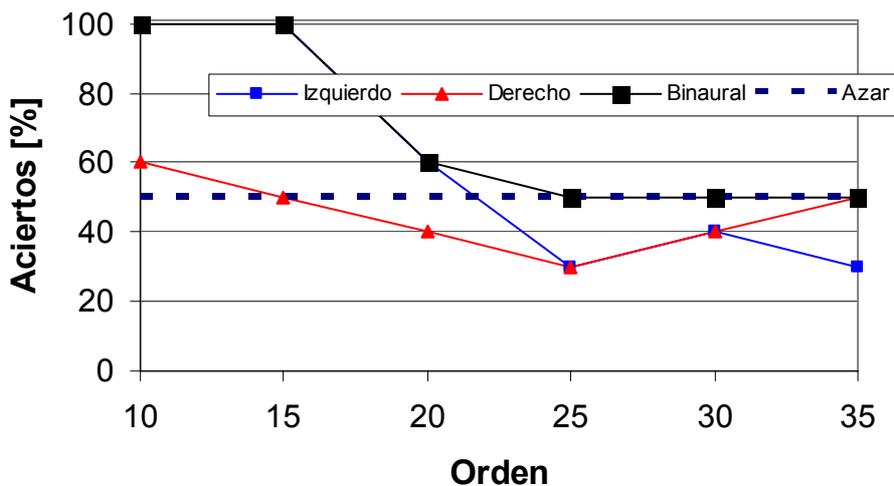


Figura 10: Porcentajes de aciertos vs. el orden del modelo para la posición  $\varphi = 90$  y  $\Phi = 0$

## 6 CONCLUSIONES

La aplicación del método de las Componentes Principales a un conjunto de HRIR medidas nos ha permitido desarrollar un modelo reducido de cabeza y torso que fue validado con pruebas psicoacústicas.

Las pruebas psicoacústicas realizadas determinaron que la utilización de entre 20 y 25 funciones básicas son suficientes para reconstruir una HRIR que no puede distinguirse de la original. El estudio de los errores RMS de las magnitudes para esos órdenes, reveló que el error es menor a 2 dB. Además, se consiguió reducir en 10 veces la cantidad de datos necesarios para reconstruir una HRIR, disminuyendo considerablemente la capacidad de almacenamiento requerida.

Se piensa que estudios complementarios deben ser realizados para mejorar el ajuste del oído contralateral ya que el orden necesario y suficiente de los modelos está determinado por dicho oído. Esto posibilitaría reducir aún más, el orden requerido de los modelos.

## REFERENCIAS

- V. Algazi, R. Duda, D. Thompson and R. Morrison. Structural Composition and decomposition of HRTF, *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Mohonk Mountain House, New Paltz, NY, 103–106, 2001b.
- V. Algazi, R. Duda, R. Duraiswami, N. Gumerov and Z. Tang. Approximating the head-related transfer function using simple geometric models of the head and torso,” *J. Acoust. Soc. Am.* 112, 2053–2064, 2002.
- J. Blauert. *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge, MA. The MIT Press, 1997.
- F. Firestone. The phase difference and amplitude ratio at the ears due to a source of pure tone. *J. Acoust. Soc. Am.*, 2, 260-268, 1930.
- J. Chen, B. D. Van Veen, K. E. Hecox. A spatial feature extraction and regularization model for the head-related transfer function. *J. Acoust. Soc. Am.* 97 439–452, 1995.
- M. Kleiner, B. Dalenbäck and P. Svensson. Auralization -an overview. *J. Audio Eng. Soc.*, 41 (11), 861–875, 1993.
- A. Kulkarni, K. Isabelle and S. Colburn S. Sensitivity of human subjects to head-related transfer-function phase spectra. *J. Acoust. Soc. Am.*, 105, 5, 2821-2840, 1999.
- A. Kulkarni, and S. Colburn. Infinite-impulse-response models of the head-related transfer function. *J. Acoust. Soc. Am.*, 115(4), 1714-1728, 2004.
- G. Kuhn. Model for the interaural time differences in the azimuthal plane. *J. Acoust. Soc. Am.*, 62(1), 157-167, 1977
- S. Mehrgardt and V. Mellert. Transformation characteristics the external human ear, *J. Acoust. Soc. Am.* 61, 1567–1576, 1977.
- A. Oppenheim and R. Schaffer. *Discreet-Time Signal Processing*. Prentice-Hall Inc. New Jersey, USA, 1989.
- O. Ramos, V. Jasá, J. Gorgas, D. Céspedes Daza. Simulación Binaural de Espacios Acústico. Memorias del V Congreso Iberoamericano de Acústica. Santiago de Chile. Chile, 2006.
- D. Rife and J. Vanderkooy. Transfer-function measurement with maximum-length sequences. *J. Audio Eng. Soc.*, 37, 419-443, 1989.
- J. Scarpaci and S. Colburn J. Principal Components Analysis Interpolation of HRTF's Using Locally Chosen Basis Functions. *Proceedings of 11 Meeting of the International Conference on Auditory Display*. Limerick, Irlanda, 2005.
- E. Shaw. Transformation of sound pressure level from the free field to the eardrum in the horizontal plane, *J. Acoust. Soc. Am.* 56, 1848–1861, 1974.
- F. Wightman and D. Kistler. Headphone simulation of free-field listening I: Stimulus synthesis, *J. Acoust. Soc. Am.* 85, 858-867, 1989a.
- F. Wightman and D. Kistler. Headphone simulation of free-field listening II: Psychophysical validation, *J. Acoust. Soc. Am.* 85, 868–878, 1989b.
- F. Wightman and D. Kistler. A model of head-related transfer functions based on principal components analysis minimum-phase reconstruction. *J. Acoust. Soc. Am.*, 91, 3, 1637-1647, 1992.
- P. Zahorik y otros. Perceptual recalibration in human sound localization: Learning to remediate front-back reversals. *J. Acoust. Soc. Am* 120, 1, 343-359, 2006.
- B. Zhou, D. Green and J. Middlebrooks. Characterization of external ear impulse responses using Golay codes. *J. Acoust. Soc. Am.*, 92, 1169-1171, 1992.