

DISEÑO Y EVALUACIÓN OBJETIVA DE FILTROS BINAURALES

Oscar A. Ramos^{a,b}, Mariano Araneda^a y Fabian Tommasini^{a,b}

^a*Centro de Investigación y Transferencia en Acústica, CINTRA, Unidad Asociada del CONICET y Universidad Tecnológica Nacional, Regional Córdoba. Maestro Lopez esq. Cruz Roja Argentina, 5016 Córdoba, Argentina, oramos@scdt.frc.utn.edu.ar.*

^b*Miembro del Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET)*

Keywords: HRIR, filtros digitales, percepción auditiva, realidad acústica virtual

Resumen: Las ondas sonoras que llegan a los tímpanos de un oyente, contienen las claves fundamentales que un ser humano utiliza para localizar una fuente sonora en el espacio. Estas claves son: la diferencia de tiempo y de amplitud de las ondas entre ambos oídos y sus contenidos espectrales. Varían con la posición de la fuente y son diferentes para cada oído. Las Respuestas Impulsivas de Cabeza (HRIR) –medidas entre una fuente sonora y los tímpanos- contienen las claves mencionadas originadas por la interacción de las ondas con la cabeza, el torso y las orejas.

En el contexto de un Sistema de Realidad Acústica Virtual (SRAV), es crítico modelar las HRIR de manera de poder sintetizar sonidos –que escuchados con auriculares- evoquen la sensación que provocaría una fuente sonora en campo libre. Más aún, estos modelos deben ser de orden bajo de manera que el costo computacionales sea compatible con sistemas en tiempo real.

Este trabajo propone modelar las HRIR mediante filtros IIR, diseñados a partir de un procedimiento iterativo propuesto por Steiglitz-McBrid. Con este método es posible aproximar tanto la magnitud como la fase de las HRIR, dentro de valores perceptualmente aceptables.

Se utiliza un conjunto de HRIR medidas a 35 sujetos en el Center for Image Processing and Integrated Computing (CIPIC) de la Universidad de California.

Para evaluar objetivamente el modelo, se analizan las claves mencionadas y se utiliza como criterio de aceptabilidad las “mínimas diferencias notables” (jnd) derivadas de estudios psicoacústicos disponibles en la literatura. El estudio muestra también, la variabilidad entre las HRIR medidas y las modeladas de toda la población de sujetos.

1 INTRODUCCIÓN

Se puede considerar a las HRIR como las salidas de un sistema lineal e invariante en el tiempo (LTI). En teoría, las HRIRs o las Funciones de Transferencia de Cabeza (HRTF, en el dominio de la frecuencia) pueden ser representadas por filtros no recursivos de respuestas finitas (FIR). Si bien es posible utilizar técnicas tradicionales de diseños de filtros, se debe notar que el problema no es trivial. Las preguntas a las que nos enfrentamos al modelar estas respuestas son: ¿que es importante que reflejen los modelos de las HRTF?; ¿cuales son las limitaciones y los requerimientos de magnitud y fase en el rango de frecuencia audibles?, entre otras. La diferencia de tiempo interaural (ITD) tiene un rol dominante en bajas frecuencias (debajo de 1.5 KHz), mientras que la diferencia interaural de niveles (ILD) hace lo propio para altas frecuencias (Blauert, 1997). En lugares del espacio en que el ITD y el ILD son mínimos o nulos (plano vertical) o ambiguos (plano sagital paralelo al plano vertical, conocido como “cono de confusión”) los perfiles espectrales cumplen un papel fundamental en todo el rango de frecuencias (Wightman and Kislser, 1989). Ahora, no está claro cuan exactas estas claves deben ser modeladas para producir una convincente representación del entorno acústico. La fidelidad o autenticidad de la *imagen auditiva* creada por una fuente virtual no sólo se puede evaluar a partir de la habilidad de los sujetos para localizar con mayor o menor precisión una fuente en el espacio. Otros perceptos deben ser considerados, como por ejemplo, que la fuente sea percibida fuera de la cabeza (*out-of head*); compacta y bien definida. Hay muy pocos estudios que intentaron dilucidar cuales son las características físicas de los estímulos sonoros responsables de estas sensaciones (Hartman and Wittenberg, 1996). Una de las causas probables es que las HRTF’s son diferentes para cada sujeto debido a las diferencias antropométricas (cabeza, torso, hombros y orejas). Cuando un sujeto utiliza HRTF medidas a otro sujeto, la percepción del entorno acústico se degrada notablemente (Wenzel et al, 1993).

La idea de modelar las HRTF utilizando filtros digitales data del año 1989 (Wightman and Kiler, 1989, a y b). En este estudio pionero, los autores comprobaron que los sujetos podían localizar fuentes sonoras virtuales (sintetizadas filtros FIR) usando auriculares con la misma precisión con que localizaban fuentes sonoras reales. Con anterioridad, Mehrgardt and Mellert, (1977) habían demostrado que las HRIR son secuencias de fase-mínima para la mayoría de las posiciones y que el *resto de fase*, esto es, la diferencia entre la fase total menos la fase-mínima, es casi lineal según la posición pero independiente de la frecuencia. El retardo asociado con el resto de fase corresponde al retardo interaural (ITD). Por lo tanto, las HRTFs pueden ser descompuestas en dos sistemas: un sistema de fase-mínima y un sistema “all-pass” (Oppenheim y Schafer, 1989). Estas evidencias, permitieron desarrollar un modelo simplificado de las HRTF conocido como: fase-mínima-más-retardo (minimum-phase-plus-delay). La simplificación consiste en suponer que la fase del “all-pass” es lineal e independiente de la frecuencia. Entonces, la fase “all-pass” será de la forma:

$$\phi_{ap} \approx -\omega\tau_{ap} \quad (1)$$

en donde τ_{ap} es un retardo constante, independiente de la frecuencia. Resumiendo: las HRTFs medidas se pueden aproximar desde las HRTFs de fase-mínima más los retardos interaurales para cada posición. Variantes de este modelo tienen que ver con el método de obtención de los retardos.

Estudios psicoacústicos han corroborado que la habilidad de los sujetos para localizar una fuente sonora utilizando HRTFs medidas no se degrada si se utilizan HRTFs generadas por este modelo (Kistler and Wigthman, 1992; Kulkarni, at. al 1999). Sin embargo, estudios más recientes demuestran que la fase de las HRTF para bajas frecuencias (hasta 1500 Hz) cumple

un rol dominante y que asumir que es lineal, sobre todo para el oído contralateral, conduce a errores insalvables en aplicaciones críticas (Scarpaci et al. 2005; Zahorik et al. 2006).

Es propósito de este trabajo, diseñar filtros IIR que contemplen tanto la magnitud como la fase de las HRTFs y que los errores sean perceptualmente aceptables

2 DESCRIPCIÓN DEL MODELO

2.1 El método iterativo de Steiglitz-McBrid

Si a un sistema con una función de transferencia $G(z)$ se le aplica una entrada:

$$X(z) = \sum_{j=-\infty}^{+\infty} x[j]z^{-j} \quad (2)$$

se obtendrá una respuesta:

$$Y(z) = \sum_{j=-\infty}^{+\infty} y[j]z^{-j} \quad (3)$$

El objetivo es, diseñar un filtro IIR de la forma:

$$H(z) = \frac{N(z)}{D(z)} = \frac{\sum_{k=1}^M b_k z^{-k}}{1 - \sum_{k=1}^N a_k z^{-k}} \quad (4)$$

Si se le aplica una entrada $X(z)$ se obtendrá una respuesta $W(z)$. Para encontrar los coeficientes a_k y b_k se debe minimizar el error cuadrático medio entre $W(z)$ y la respuesta original del sistema $Y(z)$.

$$\sum e_j^2 = \frac{1}{2\pi j} \oint \left| X(z) \frac{N(z)}{D(z)} - Y(z) \right|^2 \frac{dz}{z} \quad (5)$$

La minimización de la ecuación anterior implica un problema de regresión no lineal. El método de Steiglitz-McBride (Steiglitz and McBride, 1965) propone, en su lugar, minimizar la siguiente ecuación en sucesivas iteraciones:

$$\oint \left| X(z) \frac{N(z)_i}{D(z)_{i-1}} - Y(z) \frac{D(z)_i}{D(z)_{i-1}} \right|^2 \frac{dz}{z} \quad (6)$$

en donde i es el número de iteraciones $i = 1, 2, 3, \dots$. Se observa que si $D(z)_{i-1}$ converge a $D(z)_i$ la expresión (6) tiende a la (5).

En cada iteración se filtra las señales $X(z)$ e $Y(z)$ con un filtro todo-polo con función de transferencia $\frac{1}{D(z)_{i-1}}$ y se obtiene $X_p(z) = \frac{X(z)}{D(z)_{i-1}}$ y $Y_p(z) = \frac{Y(z)}{D(z)_{i-1}}$. En la primera iteración es necesario contar con los coeficientes a_k . Estos se obtienen por cualquier método de diseño de filtros todo-polo.

En la práctica, el procedimiento se realiza en el dominio del tiempo. En cada iteración se

filtran las señales de entrada y salida $x[j]$ e $y[j]$ con un filtro todo-polo cuyos coeficientes a_k son calculados en la iteración anterior. Obtenidas las señales filtradas $x_p[j]$ e $y_p[j]$ se define la siguiente ecuación en diferencias lineal.

$$y_p[j] + a_1 y_p[j-1] + \dots + a_k y_p[j-k] = x_p[j] + b_1 x_p[j-1] + \dots + b_k x_p[j-k] \quad (7)$$

Si se despeja $y_p[j]$

$$y_p[j] = -a_1 y_p[j-1] - \dots - a_k y_p[j-k] + x_p[j] + b_1 x_p[j-1] + \dots + b_k x_p[j-k] \quad (8)$$

si se definen las matrices

$$\theta = [a_1 \quad \dots \quad a_k \quad b_1 \quad \dots \quad b_k]^T$$

$$\varphi[j] = [-y_p[j-1] \quad \dots \quad -y_p[j-k] \quad x_p[j-1] \quad \dots \quad x_p[j-k]]^T \quad (9)$$

$y_p[j]$ será:

$$y_p[j] = \varphi^T[j] \theta \quad (10)$$

Esta ecuación lineal nos permite, en cada iteración, obtener el vector θ , o sea, los coeficientes a_k y b_k a utilizar en la próxima iteración. Para obtener los valores iniciales de los coeficientes a_k se utilizó la función Levinson. Se obtuvieron resultados satisfactorios efectuando un número de cinco iteraciones.

2.2 Las HRIRs modeladas

Se utilizó una base de datos de HRIR medidas en el Center for Image Processing and Integrated Computing (CIPIC) de la Universidad de California (Algazi et al 2001). La base contiene mediciones realizadas a 47 sujetos en 1250 posiciones por sujeto y en ambos oídos. La ubicación de la fuente sonora se especifica por el ángulo de acimut θ (25) y el ángulo de elevación Φ (50) referenciados a un eje que pasa por ambos oídos denominado eje inter-aural (Figura 1).

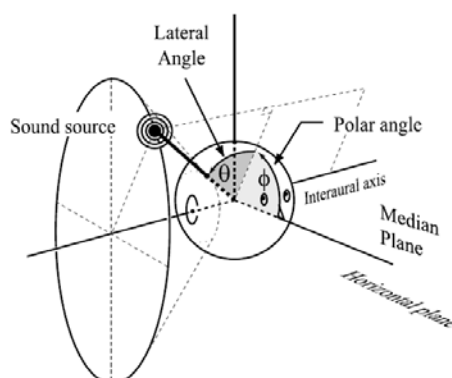


Figura 1: Sistema de coordenada polar inter-aural (extraída de Algazi et al 2001).

Las mediciones se realizaron para elevaciones que varían uniformemente en escalones de 5.625° entre -45° y $+230.625^\circ$. Los ángulos de acimut fueron: -80° , -65° , -55° , desde -45° a $+45^\circ$ en escalones de 5° , $+55^\circ$, $+65^\circ$ y $+80^\circ$. Acimut 0° y elevación 0° , corresponde al frente;

acimut 0° y elevación 180° detrás del sujeto; acimut negativos a la izquierda y positivos a la derecha del sujeto; elevaciones negativas y mayores a 180° debajo del eje inter-aural al frente y detrás respectivamente.

Las HRIR fueron medidas entre la fuente sonora y micrófonos miniaturas colocados a las entradas de los conductos auditivos. La longitud de cada HRIR es de 200 muestras tomadas a 44.100 Khz. Mas detalles se pueden consultar en [Algazi et. al \(2001\)](#).

3 EVALUACIÓN DEL MODELO

Previo a la aplicación del método, se quitó al conjunto de HRIR los retardos iniciales. Estos retardos se deben, por un lado, a la distancia entre la fuente y los micrófonos (común a ambos oídos) y por el otro, a las diferencias de tiempos inter-aural. A continuación, se calcularon los coeficientes de los filtros por el método propuesto y se restituyó los retardos a las respuestas impulsivas de los filtros y a las HRIRs medidas.

Para evaluar objetivamente los errores de la magnitud y de la fase, se utilizó la fórmula siguiente:

$$SD = \sqrt{\frac{1}{N} \sum_{i=1}^{i=N} (20 \log |H_e(\omega_i)| - 20 \log |H_a(\omega_i)|)^2} \quad (11)$$

$$PD = \sqrt{\frac{1}{N} \sum_{i=1}^{i=N} (\angle |H_e(\omega_i)| - \angle |H_a(\omega_i)|)^2} \quad (12)$$

En donde SD significa ‘distorsión espectral’ y PD ‘distorsión de fase; el subíndice *e* hace referencia a las HRTF empírica o medida y el *a* a la aproximada; N es el número de frecuencias discretas que cubren el rango de interés (0-15000 Hz)

3.1 Distribución del error según la ubicación de la fuente.

En la figura 2 se observa las *SD* para un sujeto en particular en función de la posición y para filtros de orden 20 (20 polos y 20 ceros). Note que un azimut de $+80^\circ$ significa que la fuente está casi enfrentada al oído derecho y un azimut de -80° al oído izquierdo. Al oído que está del mismo lado que la fuente, se lo llama *ipsilateral* mientras que al oído oculto por la cabeza, se lo llama *contralateral*.

El error varía sistemáticamente con la posición de la fuente. En general, es mayor para el oído contralateral en el plano horizontal (elevación = 0°) y por debajo del eje inter-aural en el plano vertical.

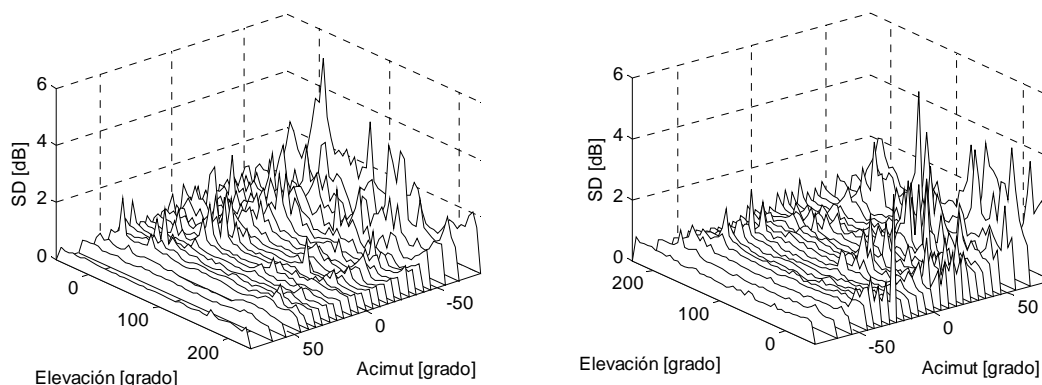


Figura 2: Distribución de la SD del oído derecho (columna de la izquierda) y del oído izquierdo (columna de la derecha) del sujeto 003 de la base

El mejor ajuste se consigue para el oído ipsilateral y es mínimo en el plano vertical (azimut=0°) para una elevación=+90° (justo encima de la cabeza).

Este comportamiento sistemático del error fue similar para todos los sujetos de la base. Más aún, es independiente del modelo que se utilice. Resultados análogos se obtienen cuando se utiliza el modelo de Componentes Principales (Ramos et al., 2008). No se muestra la distribución del error de fase debido a que presenta el mismo comportamiento que el de magnitud.

A modo de síntesis podemos decir que: los errores mayores se producen en el oído que, según la posición de la fuente, queda oculto por la cabeza y son máximos para posiciones por debajo del eje inter-aural. Por el contrario, el mejor ajuste se consigue en el oído que está en el mismo hemisferio que la fuente y el error es mínimo cuando la fuente está justo arriba de la cabeza. A continuación analizaremos el modelo en el plano horizontal, medio y sagital.

3.2 Error de módulo y fase en el plano horizontal

En la Figura 3, se grafica el error medio (línea negra) de los 35 sujetos para el oído derecho. Es mayor para azimuts negativos alcanzando un máximo de 3.17 dB alrededor de -55°.

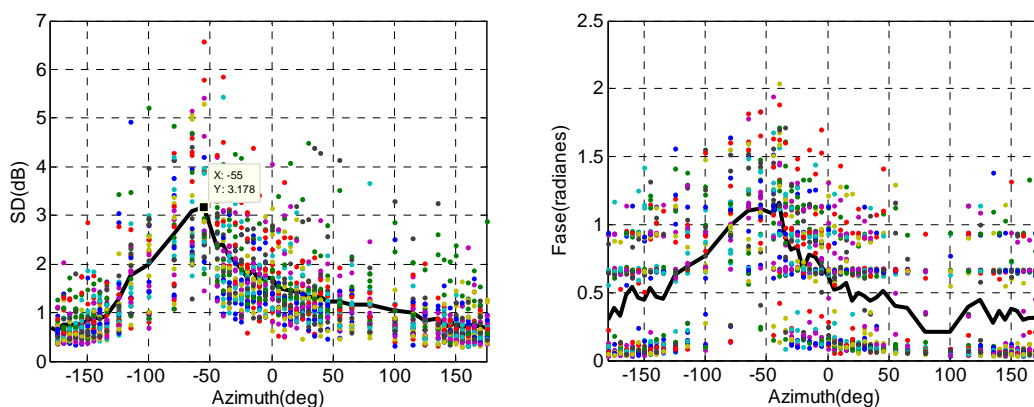


Figura 3: SD y PD del oído derecho en el plano horizontal

Para visualizar la dispersión de los errores también se grafica en puntos, los errores individuales de los 35 sujetos. En el caso de SD (columna de izquierda) se observa que el

error máximo también se produce -55° alcanzando un valor de 6.6 dB. Se comprobó que para un sujeto en particular, los errores no siguen un patrón definido a lo largo de todas las posiciones. Los errores máximos se producen sólo para algunas de ellas debido a HRIR defectuosas (quizás por movimientos involuntarios de la cabeza durante la medición). El patrón de los errores de fase es muy similar a los errores de magnitud

Morimoto, 2001, investigó la importancia del oído ipsilateral y contralateral para localizar una fuente sonora en el espacio. Concluyó que ambos oídos contribuyen de igual manera pero que la importancia de oído contralateral disminuye a medida que la fuente se desplaza lateralmente (en el azimut) del plano medio. Cuando este desplazamiento alcanza cierto valor (60°) la influencia de oído contralateral es irrelevante. En el panel izquierdo de la Figura 3 vemos que la SD máxima se produce en -55° muy próximo al valor determinado por Morimoto. Es de esperar que el error de ajuste sea irrelevante para el oído oculto por la cabeza entre -60° y -120° .

3.3 Error de módulo y fase en el plano vertical

El plano vertical, esto es, azimut igual a 0° , tiene la particularidad de que las diferencias inter-aurales de tiempo y de nivel (ITD e ILD) son casi nulas. Las pequeñas diferencias que pudieran existir se deben a las asimetrías antropométricas de los sujetos. La ausencia de estas dos claves, hace que los perfiles espectrales sean la claves dominantes en este plano (Blauert, 1997). Los perfiles espectrales de las HRTFs están determinados, especialmente, por la interacción de las ondas con las cavidades y protuberancias de las orejas y por las reflexiones en el torso para elevaciones entre -45° y $+50^\circ$ (Blauert, 1997; Shaw, 1997).

En la Figura 4 se grafica la distorsión media (línea negra) tanto de la magnitud como de la fase correspondientes al oído derecho. Se observa que ambas son bastante menores a los conseguidos en el plano horizontal. La SD crece desde -45° hasta 0° (1.70 dB); se mantiene casi constante hasta los $+65^\circ$ y luego decrece casi monótonicamente hasta los 180° para volver a crecer hasta los 230° . En otras palabras, la distorsión es mayor por debajo del eje inter-aural y por encima hasta los $+50^\circ$. Estas evidencias están de acuerdo con que en esas posiciones las reflexiones y dispersión de las ondas sonoras provocadas por el torso y hombros son importantes.

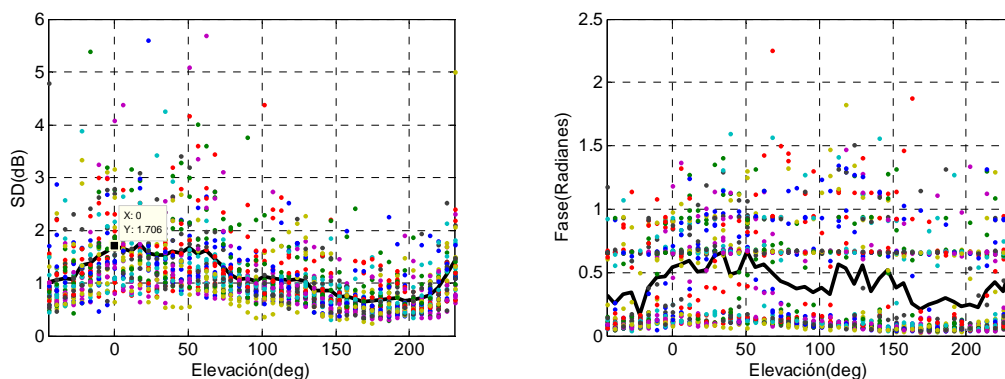


Figura 4: SD y PD del oído derecho en el plano vertical

La distorsión de fase tiene un comportamiento similar, salvo el aumento que se produce entre $+105^\circ$ y $+170^\circ$. Como en el plano horizontal, los errores máximos de cada sujeto son esporádicos y para algunas posiciones.

3.4 Error de módulo y fase en un plano sagital de azimut igual a -65°

Este plano, paralelo al plano vertical, también tienen sus particularidades. Las ITDs son iguales para todas las elevaciones. En consecuencia, esta clave es ambigua y es responsable de confusiones que un sujeto experimenta al juzgar la posición de una fuente como arriba o adelante cuando en la realidad se encuentra abajo o detrás respectivamente (Wenzel, 1993). La ILD y los perfiles espectrales son las claves dominantes en este tipo de planos.

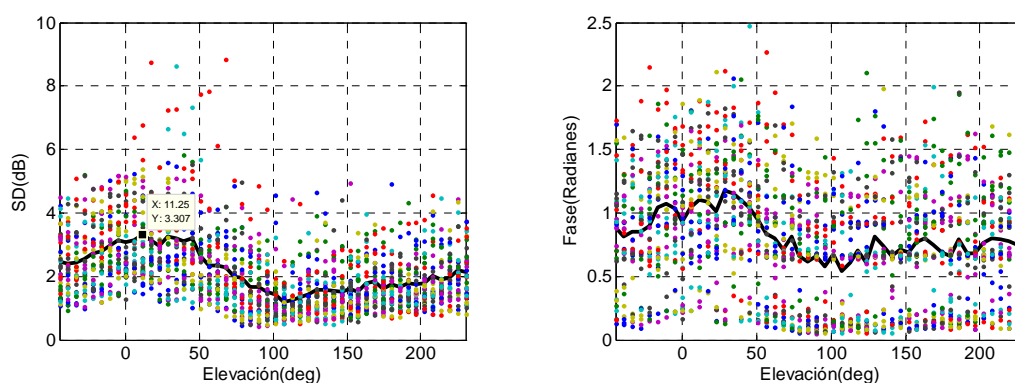


Figura 5: SD y PD del oído derecho en un plano sagital para un azimut $= -65^\circ$

En la Figura 5 se ve las distorsiones medias de magnitud y fase en un plano sagital para un azimut constante de -65° . El oído analizado es el derecho y el plano sagital considerado está en el hemisferio opuesto. La distribución de los errores es similar al encontrado en el plano vertical pero sus valores son casi el doble. La SD es mayor para elevaciones entre -45° y $+50^\circ$ (3.30 dB); disminuye hasta los $+110^\circ$ aproximadamente y luego aumenta levemente hasta los $+230^\circ$. Resumiendo: el error es mayor por debajo del eje inter-aural al frente del sujeto hasta los $+50^\circ$ (las reflexiones en el torso y sobre todo en hombros es importante). Para elevaciones mayores a 180° (por debajo del eje inter-aural, pero detrás de sujeto) el error aumenta pero menos que al frente del sujeto. La distribución de la PD es igual que en el plano horizontal, pero bastante mayor.

4 ANÁLISIS DE RESULTADOS

Veámos al comienzo que modelar las HRIR no era un hecho trivial ya que debían tenerse en cuenta aspectos psicoacústicos de la audición espacial. No se ha establecido aún cuales son los errores de magnitud y fase aceptables desde el punto de vista perceptual que hacen que una HRIR modelada no pueda distinguirse de una real. Está en discusión cuales de las claves mencionadas son binaurales (intervienen ambos oídos) y cuales son monoaurales (dependen de un solo oído). De todos modos, diversos son de factores que contribuyen a una verídica percepción de una fuente sonora simulada con auriculares.

Algunos estudios han sugerido que valores de SD de 5.6 dB son satisfactorios (Takanori et al, 1999). Por nuestra parte (Ramos et al, 2007) comprobamos, mediante pruebas psicoacústicas, que un SD de 2 dB era suficiente para que sonidos procesados con HRTF medidas y con HRTF aproximadas no puedan ser discriminados.

Con el objeto de estimar a priori el orden óptimo del modelo, realizaremos un estudio comparativo de las claves producidas por las HRIR medidas y las generadas por el modelo.

En primer lugar analizaremos el comportamiento del ITD. Para ello, aplicamos las formulas siguientes:

$$\tau_e(f) = \frac{\phi_e^r(\omega) - \phi_e^l(\omega)}{2\pi f} \quad (13)$$

$$\tau_a(f) = \frac{\phi_a^r(\omega) - \phi_a^l(\omega)}{2\pi f} \quad (14)$$

En donde ϕ es la fase de las HRTFs y el subíndice hace referencia a las HRTFs empíricas y aproximadas y el superíndice al oído. Los τ_e y τ_a son los retardos inter-aurales de tiempo (ITD) de las HRTFs empíricas y las aproximadas respectivamente y ω la frecuencia. Según Kulkarni et al., 1999 el ITD es una clave dominante hasta los 1500 Hz. Sin embargo hace la salvedad de que depende más de un valor global o medio que de sus variaciones en dicho rango de frecuencias. En la Figura 6 se muestra, en la columna de la izquierda y en trazo negro, el ITD medio en el plano horizontal para las HRTFs medidas entre 150 Hz y 1500 Hz promediado entre los 35 sujetos. En puntos rojos, se grafica lo mismo para las HRTF aproximadas. En la columna de la derecha se ve la diferencia entre ambos ITD (trazo negro) y en puntos las diferencias para los 35 sujetos.

Como se observa en la columna derecha de la Figura 6, la diferencia entre los ITD_e y los ITD_a es mayor para azimuts positivos y toma un valor máximo de 40 microsegundos para una posición de casi $+50^\circ$. Este error es menor a la “mínima diferencia notable” (jnd) según Durlach y Colburn, 1978. Kulkarni et al, 1999 establecieron que los sujetos no son sensibles a la fase monoaural, esto es, a la fase de los estímulos sonoros aplicados a cada oído por separado. Demostraron, por el contrario, que el ser humano es altamente sensible a la diferencia de fase entre ambos oídos. Observando las diferencias de los ITD para todos los sujetos, se comprobó que las grandes diferencias se corresponden para posiciones determinadas y no en todo el plano horizontal.

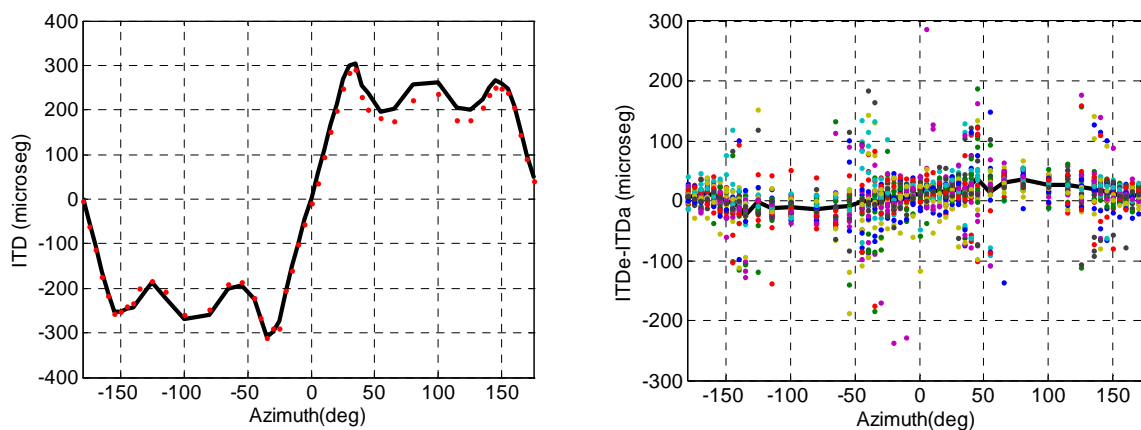


Figura 6: Panel izquierdo: ITD promedio en el plano horizontal de las HRTFs medidas (trazo negro) y de las HRTFs aproximadas entre 150 Hz y 1500 Hz (puntos rojos). Panel derecho: Error medio entre los ITD medidos y los ITD aproximados (trazo negro). En puntos, los errores individuales.

La diferencia de nivel interaural es importante y dominante para frecuencias en que la longitud de onda es menor al diámetro de la cabeza (aproximadamente 500 Hz). Sin embargo esta clave se hace dominante para frecuencias por encima de los 1500 Hz. La ILD es mayor para azimut de $+90^\circ$ y -90° (lateral izquierdo y derecho) y menor a 0° (justo al frente) y 180° (justo detrás). En la Figura 7 se muestra el ILD promediado para todas las posiciones del plano horizontal y todos los sujetos entre 1500 Hz y 15000 Hz. La diferencia entre los ILD_e y

los ILD_a es bastante menor a 1 dB y los errores individuales son también pequeños con la misma salvedad hecha en el análisis de los ITD.

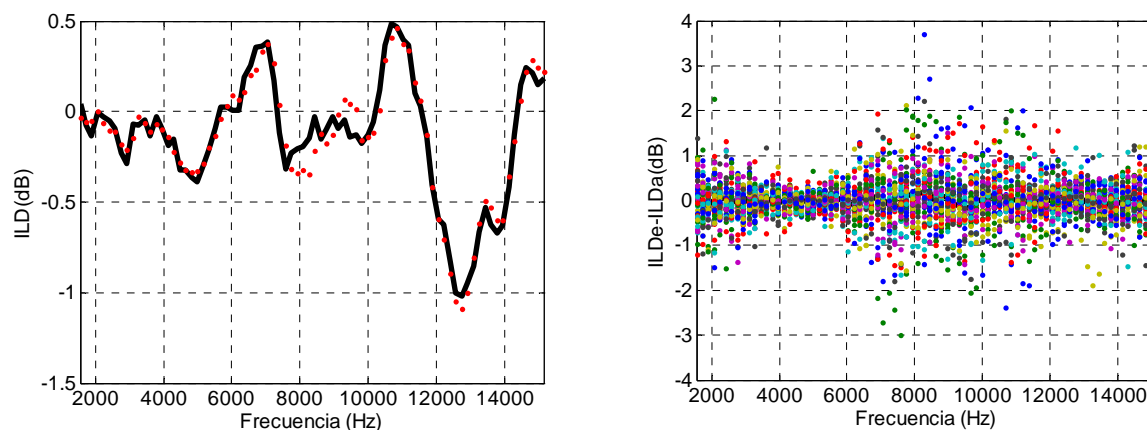


Figura 7: Panel izquierdo: ILD promedio en el plano horizontal de las HRTFs medidas (trazo negro) y de las HRTFs aproximadas entre 1500 Hz y 15000 Hz (puntos rojos). Panel derecho: Error medio entre los ILD medidos y los ILD aproximados (trazo negro). En puntos, los errores individuales.

Los contenidos espectrales de las HRIR son dominantes en aquellas regiones del espacio en que las diferencias interaurales eran nulas o ambiguas. Una de estas regiones es el plano vertical. La interacción de las ondas sonoras con las cavidades y protuberancias de las orejas dan origen a resonancias y cancelaciones que varían sistemáticamente con la posición de la fuente en este plano. Las HRIR del oído derecho y el izquierdo son muy similares (difieren sólo debido a la asimetría natural del sujeto) y por lo tanto las claves están codificadas en los contenidos espectrales de las ondas que llegan a cada oído individualmente (claves monoaurales).

Revisando la Figura 4 vemos que la SD media para todos los sujetos del oído derecho es pequeña y es menor a 1.70 dB en todo el plano vertical. Este valor está muy por debajo de la jnd establecido por algunos estudios (por ejemplo en [Durlach y Colburn, 1978](#)).

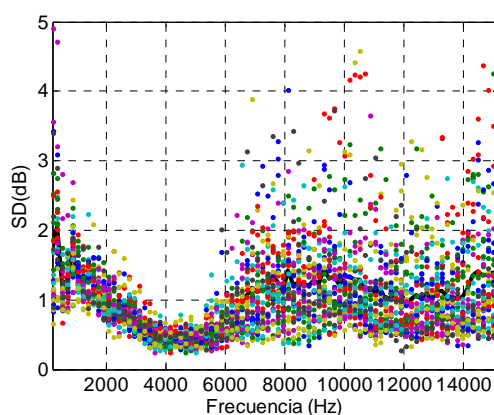


Figura 8: SD promediadas para todas las posiciones del plano vertical y todos los sujetos entre 1500 y 15000 Hz.

En la figura 8 se grafica la distribución de las SD promediadas para todas las posiciones del plano vertical y todos los sujetos entre 1500 Hz y 15000 Hz. Se ve que el valor medio está

por debajo de los 2 dB.

5 CONCLUSIONES

Se han diseñado filtros IIR utilizando el método iterativo propuesto por Steiglitz-McBrid, a partir de HRIR medidas a 35 sujetos. Se analizó objetivamente el ajuste de los filtros y se utilizó como criterio de aceptabilidad las “mínimas diferencia notables” establecidas para las claves fundamentales para localizar una fuente sonora en el espacio. El estudio se basó en filtros de orden 20, pero de los resultados obtenidos se puede inferir que este orden podría ser menor. En este momento se están realizando pruebas psicoacústicas que determinaran el orden adecuado.

Se analizó la distribución de las distorsiones de magnitud y de fase en el plano horizontal, vertical y sagital para un azimut de -65° . Se comprobó que los errores medios promediados para todos los sujetos, son máximos cuando la fuente sonora está por debajo del eje interaural y para el oído contralateral (el oído oculto por la cabeza). Por otro lado, los errores mínimos tanto de magnitud como de fase, se obtuvieron en el plano vertical.

Se investigó el comportamiento de las diferencias de tiempo y nivel (ITD e ILD) y los contenidos espectrales en el plano vertical. Se observó que las diferencias interaurales de tiempo y nivel en el rango de frecuencias en que son preponderantes, están por debajo de las “mínimas diferencias notables” (jnd).

Los errores de cada sujeto en todos los planos analizados, no siguen un patrón definido respecto a la posición de la fuente. Los errores máximos se producen sólo para algunas de las posiciones pudiéndose atribuir a HRIR defectuosas (quizás por movimientos involuntarios de la cabeza durante la medición).

REFERENCIAS

- Algazi, V., Duda, R., Thompson, D. and Avendano, C., The CIPIC HRTF database. *IEEE Workshop on applications of Signal Processing to Audio and Acoustics*, New Paltz, New York, USA, 99-102, 2001.
- Blauert, J. *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge, MA. The MIT Press, 1997.
- Durlach, N. I., and Colburn, H. S., “Binaural phenomenon,” in *Handbook of perception*, Vol. 4, edited by E. C. Carterette and M. P. Friedman (Academic, New York), 365–466, 1978.
- Hartman, W. and Wittenberg, A., On externalization of sound images. *J. Acoust. Soc. Am.*, (99), 6:3678-3688, 1996.
- Kistler, D. and Wightman, F.. A model of head-related transfer functions based on principal components analysis minimum-phase reconstruction. *J. Acoust. Soc. Am.*, 91, 3:1637-1647, 1992.
- Kulkarni, A., Isabelle, K. and Colburn S. Sensitivity of human subjects to head-related transfer-function phase spectra. *J. Acoust. Soc. Am.*, 105, 5:2821-2840, 1999.
- Mehrgardt, S. and Mellert, V. Transformation characteristics the external human ear, *J. Acoust. Soc. Am.* 61:1567–1576, 1977.
- Morimoto, M. The contribution of two ears to the perception of vertical angle in sagittal planes. *J. Acoust. Soc. Am.* 109 (4):1596-1603, 2001.
- Ramos, O., and Tommasini, F. Simplificación de las Funciones de transferencia de cabeza mediante Análisis de Componentes Principales. *Mecánica Computacional*, Vol. XXVII: 431-442, 2008.
- Ramos, O., Calvo, G. and Tommasini, F. Modelo Acústico de Cabeza y Torso Mediante Análisis de Componentes Principales. *Mecánica Computacional*, Vol. XXVI: 46-58, 2007.

- Scarpaci, J. and Colburn, S. 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.
- Shaw, E. Acoustical Features of the Human External Ear in *Binaural and Spatial Hearing in Real and Virtual Environments*. Editor's Gilkey, R. and Andersosn, T. Lawrence Erlbaum Associates, Inc. 1997.
- Steiglitz, K., and McBride, L. E. "A technique for the identification of linear systems," *IEEE Transactions on Automatic Control* AC-10:461–464, 1965.
- Takanori, M., Shoji, K., Kazuya, and Fumigada, I. Interpolation of the head-related transfers function on the horizontal plane. *J. Acoust. Soc. Jp* 55:91-99, 2009.
- Wenzel, E., Arruda, M. Kisler, D. and Wightman, F. Localization using nonindividualized head-related functions. *J. Acoust. Soc. Am.* 94, (1) :111–123, 1993.
- Wightman, F. and Kistler, D. Headphone simulation of free-field listening I: Stimulus synthesis, *J. Acoust. Soc. Am.* 85:858-867, 1989a.
- Wightman, F. and Kistler, D. Headphone simulation of free-field listening II: Psychophysical validation, *J. Acoust. Soc. Am.* 85:868–878, 1989b.
- Zahorik. P. y otros. Perceptual recalibration in human sound localization: Learning to remediate front-back reversals. *J. Acoust. Soc. Am* 120, 1:343-359, 2006.