

CONSTRUCCIÓN DE HRTF INDIVIDUALES A PARTIR DE MEDICIONES ANTROPOMÉTRICAS

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

^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: HRTF individuales, PCA, percepción auditiva, realidad acústica virtual

Resumen: Las transformaciones que experimentan las ondas sonoras en la cabeza, torso, hombros y orejas, constituyen claves fundamentales que una persona utiliza para percibir y localizar una fuente sonora en el espacio. Estas claves son diferentes para cada individuo debido a las diferencias antropométricas tales como: tamaño de la cabeza y orejas; asimetrías de las orejas y entradas de los conductos auditivos; ancho de hombro y torso; etc. Una autentica y verídica percepción se consigue cuando una persona escucha por auriculares, sonidos filtrados por sus propias Funciones de Transferencia de Cabeza (HRTF). Por el contrario, si se utilizan HRTF medidas a otros individuos la percepción auditiva se degrada notablemente.

Desde el punto de vista práctico y económico es imposible medir a cada usuario potencial de un Sistema de Realidad Acústica Virtual (SRAV) sus propias HRTF. Por esta razón es esencial diseñar estrategias que permitan construir de manera sencilla y eficaz HRTF personalizadas.

En este trabajo se analiza la influencia del oído externo, la cabeza y el torso en las HRTF con el propósito de construir HRTF personalizadas mediante regresión lineal múltiple entre mediciones antropométricas y las HRTF medidas a 35 sujetos por el CIPIC (Center for Image Processing and Integrated Computing de la Universidad de California).

INTRODUCTION

Las ondas producidas por una fuente sonora, antes de alcanzar los tímpanos de un oyente, son alteradas por la cabeza, el torso y las orejas. Debido a la separación entre ambos oídos, las ondas recorren caminos diferentes dando lugar a una diferencia de tiempo conocida como diferencia de tiempo interaural (ITD). Según la posición de la fuente, uno de los oídos queda oculto por la cabeza (sombra acústica) originando una diferencia de nivel sonoro entre ambos (diferencia de nivel interaural: ILD). En las protuberancias y cavidades de las orejas se producen reflexiones y resonancias que modifican el espectro de las ondas incidentes. Los fenómenos mencionados, esto es, la diferencia de tiempo y de nivel interaural y los espectros son claves que el ser humano utiliza para detectar y localizar una fuente sonora.

Las Respuestas Impulsivas de Cabeza (HRIR) o las Funciones de Transferencia de Cabeza (HRTF), en el dominio de la frecuencia, caracterizan dichas transformaciones. Son diferentes para cada oído y varían sistemáticamente con la ubicación en el espacio de la fuente sonora. Las HRIR son medidas entre una fuente sonora ubicada en una superficie esférica y micrófonos miniaturas colocados en la proximidad de los tímpanos o a las estradas de los conductos auditivos.

Será posible crear una *imagen* auditiva virtual de una fuente sonora real si mediante auriculares, por ejemplo, se reproducen en los tímpanos de un oyente los mismos estímulos sonoros que produciría la fuente real. Para alcanzar este propósito, se debe filtrar un sonido con las HRTF. Las HRTF son diferentes entre individuos debido a las disimilitudes antropométricas tales como: tamaño de la cabeza y orejas; ancho de hombro y torso; entre otras. Si las HRTF utilizadas corresponde a las del oyente, la fuente se percibe compacta, externa y bien definida en una posición del espacio. Por el contrario, si las HRTF empleadas pertenecen a otro individuo, la fuente se percibe difusa y ubicada en el interior de la cabeza. Como no es práctico ni económico medir a cada usuario potencial de un SRAV sus propias HRTF, es decisivo desarrollar procedimientos que permitan construir las HRTF's de un individuo en particular sin tener que medirlas.

Desde la década del los 90 hasta nuestros días se han propuesto diferentes abordajes para personalizar las HRTF. Basándose en un estudio pionero de [Wightman y Kisler \(1989\)](#), [Wenzel y otros \(1993\)](#) propusieron utilizar las HRTF de un sujeto del estudio original, cuya habilidad para localizar fuentes sonoras fue excepcional. Se concluyó que si bien algunas claves, sobre todo las diferencias interaurales, fueron efectivas para localizar en el plano horizontal, los detalles espectrales, responsables para juzgar la elevación de una fuente, fueron deteriorados. [Middlebrooks \(1999\)](#) descubrió que los perfiles espectrales de las HRTFs, varían sistemáticamente de individuo a individuo a lo largo del eje de las frecuencias. Por lo tanto es posible construir las HRTFs de un individuo cualquiera desplazando a lo largo de la frecuencia las HRTFs medidas a otro sujeto. El valor del corrimiento lo determina por un factor de escala que calcula por una fórmula en la que intervienen el tamaño de la cabeza y del oído externo. [Seeber and Fastl \(2003\)](#) utilizaron un procedimiento psicoacústico mediante el cual un individuo podía seleccionar de un catálogo de HRTFs medidas a varios sujetos, aquellas HRTFs que mejoraban su rendimiento. [Brown y Duda \(1998\)](#) proponen un modelo estructural simplificado, en el dominio del tiempo, basado en modelos físicos-matemáticos de la propagación de ondas y difracción. Los parámetros del modelo pueden ser ajustados de manera tal de producir HRTFs individuales.

[Martens \(1976\)](#) y [Kisler y Wightman \(1992\)](#) utilizan el Análisis de Componentes Principales (PCA) para modelar las HRTF's. Obtuvieron un conjunto de *funciones básicas* o Componentes Principales (CP) comunes a todos los sujetos y una serie de *pesos* (W) para cada posición de la fuente y para cada sujeto. Cualquier HRTF del conjunto puede ser

reconstruida por combinación lineal de las CP multiplicadas por los W correspondientes. Este método de descomposición de las HRTF en CP , permitió estudiar aspectos comunes de un grupo de HRTF e inferir del análisis de los pesos W las diferencias entre sujetos para diferentes posiciones de la fuente sonora (Kisler y Wightman, 1992; Hwand and Park, 2007 y 2008; Ramos y Tommasini, 2008a y b).

Hwand and Park, 2008 proponen un método en donde el sujeto, ayudado por un programa computacional, ajusta sistemáticamente los W de mayor desviación estándar para cada posición de la fuente.

En el año 2001, Algazi y otros publican una base de datos de HRTF medidas a 45 sujetos para 25 ángulos de acimut y 50 ángulos de elevación (1250 posiciones de la fuente). Además, la base de datos contiene mediciones antropométricas de cada sujeto. Utilizando este conjunto de HRTF de dominio público, numerosos autores se propusieron determinar los rasgos antropométricos de mayor correlación con las HRTF medidas. Apelando al método de PCA y usando variados procedimientos determinaron los W de un individuo en particular a partir de mediciones de alguno de los rasgos antropométricos (Hu y otros, 2007; Xu y otros, 2009). Tanto los resultados como la selección de las mediciones antropométricas preponderantes son disímiles.

En un trabajo reciente (Fels and Vorländer, 2009) se estudió la influencia de las dimensiones de la cabeza, torso y orejas en las HRTF de individuos que van desde la niñez (6 meses) a la adultez (23 años). Es propósito de los autores, dilucidar cuales son los rasgos antropométricos más importantes que determinan las claves que hacen a la audición espacial durante el crecimiento. La conclusión más importante es que las HRTFs de los niños no son una versión en escala de las HRTF de los adultos. Si embargo se pudo establecer cuales son las dimensiones antropométricas comunes a toda la población de sujetos que tienen influencia sobre sus HRTF. La utilización, por nuestra parte, de los resultados de este trabajo, nos permitirá generalizar la construcción de HRTF personalizadas para un rango importante de personas (sobre todo de niños, que por razones obvias, es imposible medir sus HRTF).

Es propósito de este trabajo, establecer las mediciones antropométricas más importantes que tienen influencia en las HRTF de un individuo y construir HRTF's personalizadas mediante regresión lineal múltiple entre dichas mediciones y los W que resultan de la aplicación del método PCA a un conjunto de HRTFs medidas a 35 sujetos por el CIPIC (Center for Image Processing and Integrated Computing de la Universidad de California).

1 LA BASE DE DATOS

1.1 Las HRIR

La base de datos contiene mediciones de HRIR 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 φ (25) y el ángulo de elevación θ (50) referenciados a un eje que pasa por ambos oídos denominado eje inter-aural (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.

1.2 Mediciones antropométricas

Diversas mediciones antropométricas se realizaron a cada sujeto. En la Figura 1 se muestran las 17 mediciones realizadas en la cabeza y torso y 10 en las orejas. De los 47 sujetos trabajamos con 35 ya que el resto no tienen realizadas todas las mediciones.

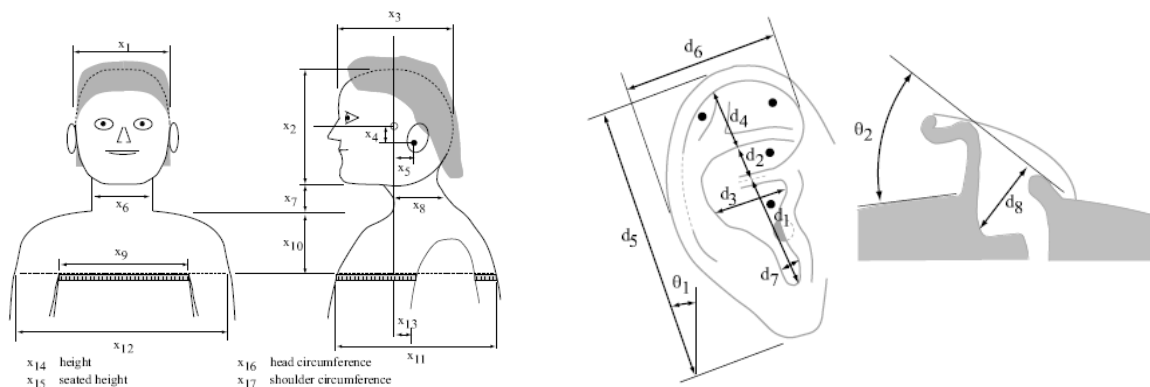


Figura 1: Mediciones de cabeza, torso y orejas (extraída de Algazi et al., 2001)

Algazi et al., (2001) realizaron un análisis estadístico entre pares de mediciones entre todos los sujetos. Manifiestan que, en general, las mediciones tienen entre sí una pobre correlación a pesar de que estadísticamente son significantes (95%). Sorprende a los autores, la baja correlación de las mediciones de las orejas entre sí y entre la altura de la cabeza (x_2) y la altura de las orejas (d_5). Relacionaron algunas características de las HRTF con algunas de las mediciones antropométricas. Encontraron que la máxima diferencia de tiempo interaural de cada sujeto está altamente correlacionado con el ancho (x_1) y profundidad (x_3) de la cabeza. Descubrieron, en el dominio de la frecuencia, que la prominente resonancia entre 3 KHz y 4 KHz y el “notch” característico de las HRTF en el plano vertical, se podía predecir mediante regresión lineal de la altura de la cavidad de la concha (d_1+d_2) y los ángulos Φ_1 y Φ_2 . Actualmente, no está claro cuáles son los rasgos antropométricos que mejor determinan las características relevantes de las HRTF. Por tal razón, es difícil establecer cuáles de estas mediciones pueden ser descartadas y cuáles son significativas.

2 PROCESAMIENTO DE LAS HRIR

Mehrgardt and Mellert, (1977) demostraron 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 e independiente de la frecuencia. Estas evidencias permitieron desarrollar un modelo simplificado de las HRTFs conocido como: fase-mínima-más-retardo (minimum-phase-plus-delay). Las HRTFs pueden ser aproximadas por las magnitud de las secuencias de fase-mínima y la fase por: $\phi \approx -\omega\tau$, en donde τ es un retardo constante e independiente de la frecuencia. Variantes de este modelo tienen que ver con el método de obtención de los retardos (Kuhn 1977, Kulkarni, et al 1999). Pruebas psicoacústicas 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 Wightman, 1992). Adoptamos este modelo para evaluar el procedimiento de personalización o individualización de las HRTFs.

En primer lugar, se obtuvo las HRTFs mediante la transformada de Fourier. Luego, se aplicó el logaritmo a las magnitudes y se calculó la media para todas las posiciones; para cada oído y cada sujeto. Estas medias fueron restadas de cada HRTF's para obtener las Funciones de Transferencias Direccionales (DTF). Por último, aplicamos al conjunto de DTF el método

de PCA.

El análisis de Componentes Principales es un método estadístico que pretende reducir la dimensionalidad de un conjunto de datos intercorrelacionados, resaltando sus características comunes. Cuando la dimensión de los datos es grande, es posible además, develar patrones disimulados en el conjunto. Consiste en proyectar los datos empíricos u observaciones sobre un conjunto de vectores ortonormales o *funciones básicas*, en el sentido en que la covarianza común es máxima (Shlens, J. 2005).

De la aplicación del método se obtiene un conjunto de funciones básicas o componentes principales (*CP*) y un conjunto de *pesos* (*W*) que reflejan la contribución relativa de cada *CP* para reconstruir una HRTF en particular. Cada *CP* captura en mayor o menor grado algún aspecto común del conjunto. Cualquier DTF se reconstruye totalmente calculando:

$$\log \left[\hat{H}(f, \varphi, \theta, s) \right] = H_0(f) + \sum_{i=1}^N CP_i(f)W_i(\varphi, \theta, s) \quad (1)$$

en donde H_0 es la media global del conjunto; N es la totalidad de las *CP*; f la frecuencia y s un sujeto en particular. Para conseguir una efectiva reducción de la dimensión de los datos originales, se debe determinar un número L , bastante menor a N , que permita reconstruir una HRTF dada dentro de un error perceptualmente aceptable. (Ramos et. al, 2007; Ramos et. al, 2008a). Como se desprende de (1) las *CP* dependen de la frecuencia f y los W dependen del acimut φ , la elevación θ de la fuente y del sujeto s .

3 MÉTODO

La idea subyacente es determinar los W de un sujeto en particular, fuera de la muestra, a partir de sus mediciones antropométricas y poder reconstruir sus HRIRs o HRTFs sin tener que medirlas. Los métodos propuestos son disímiles. El más utilizado consiste en expresar los pesos W como una combinación lineal de dimensiones antropométricas seleccionadas por correlación (Hu y otros, 2007) o mediante redes neuronales (Xu y otros, 2009). Los criterios de selección de las mediciones están basados en procedimientos meramente matemáticos y los resultados presentados a menudo son parciales (sólo para algunas posiciones y sujetos) y muy difícil de evaluar globalmente.

W se puede estimar a partir de la matriz de las mediciones antropométricas X de los 35 sujetos.

$$\tilde{W} = \alpha X \quad (2)$$

La sumatoria de la ecuación (1) se puede expresar:

$$\hat{H} = CPW \quad (3)$$

Reemplazando en (3) por la (2)

$$\tilde{H} = CP\alpha X \quad (4)$$

en donde \tilde{H} es la matriz de HRTF calculadas por las mediciones antropométricas y α los coeficientes obtenidos por regresión lineal multidimensional. Estos pueden ser despejados para obtener:

$$\hat{\alpha} = (X^T X)^{-1} X^T (\tilde{H} W^T) \quad (5)$$

Una de las metas que perseguimos fue que la cantidad de las mediciones antropométrica sean las menos posibles y que sean fáciles de medir sin usar herramientas especiales. Para ellos nos basamos en un trabajo de [Feld y Vorlander \(2009\)](#) cuyo objetivo fue estudiar cuales son las dimensiones más importantes de la cabeza, torso y orejas durante el crecimiento de una persona, que determinan los detalles espectrales de las HRTFs y sus correspondientes claves binaurales. Este análisis fue realizado en 97 sujetos cuyas edades varían desde los 6 meses a 17,5 años. Concluyeron que las dimensiones más relevantes fueron: el alto y ancho de las orejas (d5 y d6); el alto, ancho y profundidad de la concha (d1+d2, d3 y d8); el ancho de la cabeza X1; la distancia entre el conducto auditivo y la nuca (X3÷2 -X5) y finalmente, la más significativa de todas, la distancia entre las orejas y los hombros (X2÷2-X4-X7). Lamentablemente no incluyeron en el estudio los efectos del torso. Afortunadamente las dimensiones mencionadas pueden ser extraídas o calculadas de nuestra base de datos.

Para evaluar los errores se utilizó la fórmula siguiente:

$$SD = \sqrt{\frac{1}{N} \sum_{i=1}^{i=n} \left(\hat{H}(\omega_i) - \tilde{H}(\omega_i) \right)^2} \quad (6)$$

En donde SD significa ‘distorsión espectral’; \tilde{H} son las HRTF calculadas por las mediciones antropométricas y \hat{H} las obtenidas por PCA y n es el número de frecuencias discretas que cubren el rango de interés (0-15000 Hz).

4 RESULTADOS

En un trabajo anterior ([Ramos et al, 2008a](#)) habíamos determinado que las 5 primeras CP reflejan la influencia de las orejas y la cabeza. Las reflexiones producidas en torso y hombros influyen a partir de la sexta CP.

En primer lugar reconstruimos las HRTFs con las 5 CP de mayor covarianza y calculamos las SD entre las HRTF medidas y \hat{H} (3) y las medidas y las \tilde{H} (4). Utilizamos las dimensiones de cabeza y orejas obtenidas por [Feld y Vorlande \(2009\)](#). El análisis del las SD se realizó para el plano horizontal (elevación=0) y vertical (acimut=0) y con respecto a la frecuencia (Figura 2). En el panel superior se grafican las SD medias para el oído derecho de todos los sujetos en el plano horizontal y en el panel inferior en el plano vertical. La SD (panel superior izquierdo) de \hat{H} (trazo negro) es mayor para acimut negativos, esto es, cuando el oído está oculto por la cabeza y disminuye significativamente, cuando la fuente está en el mismo hemisferio que el oído. Con \tilde{H} (trazo rojo) sucede algo similar, con la salvedad de que las SD son mayores. La diferencia entre ambas (trazo azul) es menor a 1 dB en todo el plano horizontal; es casi constante entre $\pm 100^\circ$ y aumenta levemente en los extremos, esto es, cuando la fuente está detrás del sujeto. En el panel superior derecho se muestra la SD respecto a la frecuencia, promediada entre todos los sujetos y las posiciones en el plano horizontal. Por debajo de los 5 Khz el error es casi cero; luego crece hasta 1 dB en 8 Khz y se mantiene, con pequeñas variaciones, en ese valor hasta el final del rango analizado.

En el panel inferior de la Figura 2 se ve los resultados obtenidos para el plano vertical. Las SD de \hat{H} y \tilde{H} son menores para 90° (encima del sujeto) y crece sistemáticamente hacia los extremos (debajo del eje inter-aural). La diferencia entre ambas SD es casi constante y menor

a 1 dB.

Hicimos lo propio con la *CP6* a la *CP10* para reconstruir las HRTF. Decíamos más arriba, que a partir de la sexta *CP* se evidencian, especialmente, las reflexiones producidas en torso y hombros. Estas reflexiones antes de ingresar a los conductos auditivos son afectadas también por la cabeza y las orejas. Agregamos a las mediciones antropométricas seleccionadas en el paso anterior, aquellas correspondientes al torso y hombros. Adoptamos, en primera instancia, las de mayor dimensión. Elegimos el ancho y profundidad del torso (x9 y x11) y el perímetro tomado alrededor de los hombros (x17). Pensamos también que el

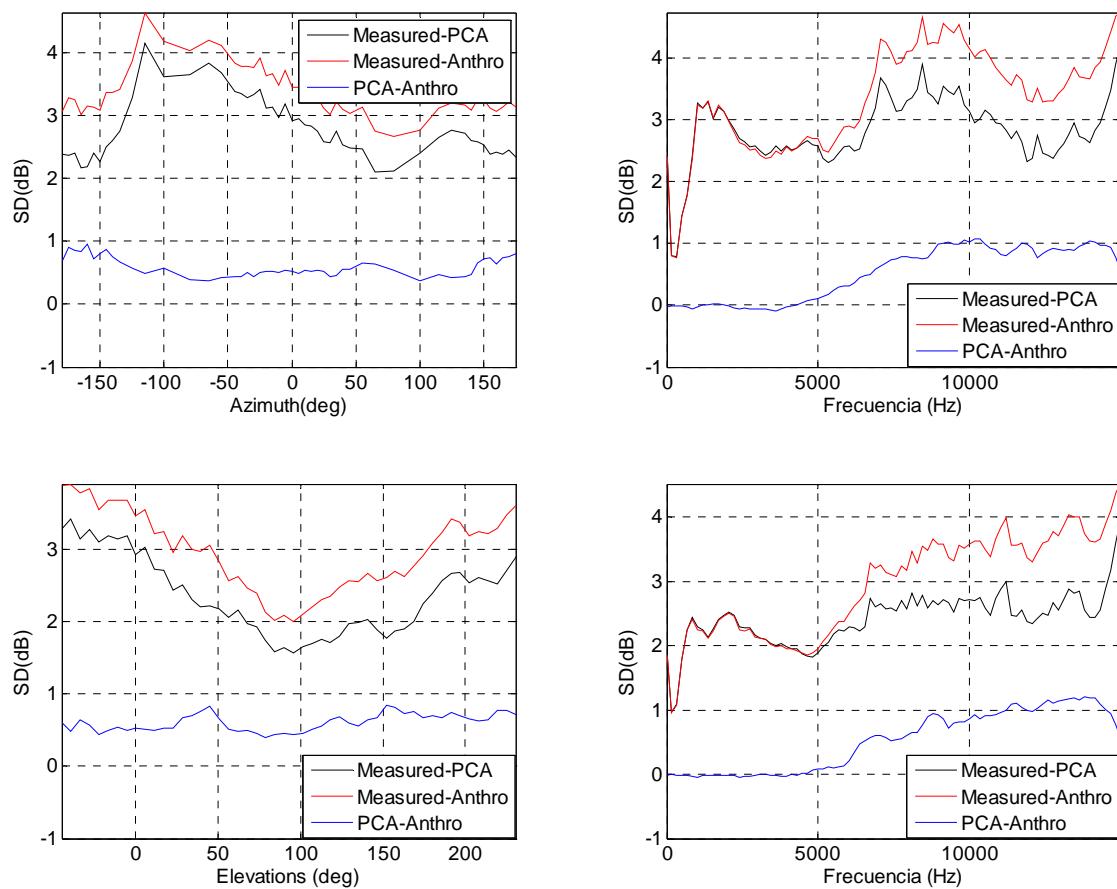


Figura 2: Error medio entre las HRTFs reconstruidas por PCA (trazo negro) y reconstruidas por mediciones antropométricas (trazo rojo) del oído derecho, utilizando las 5 *CP* de mayor orden. El panel superior corresponde al plano horizontal y el inferior, al plano vertical.

cuello tendría influencia ya que interviene, por un lado, en la distancia entre oídos y el torso (influye en la longitud recorrida por las reflexiones), y por el otro, produce difracción y obstrucción en el oído oculto. El alto del cuello (x7) ya está contemplado en la distancia entre los oídos y el hombro ya utilizada anteriormente. Consideramos al cuello como un cilindro. El ancho y profundidad del cuello (x6 y x8) son difíciles de medir, comparado con su perímetro. Por no contar con esta medición, lo estimamos tomando como valor del diámetro el promedio de $x6 + x8$.

En la Figura 3 se muestran los resultados. Sorprendentemente, las diferencias de las SD de \hat{H} y \tilde{H} son iguales a cero en ambos planos y a lo largo de la frecuencia. Se extendió el análisis de los errores de la sexta a la 20 *CP* y se comprobó que el comportamiento no difiere de lo

mostrado en la Figura 3.

5 ANÁLISIS DE LOS RESULTADOS

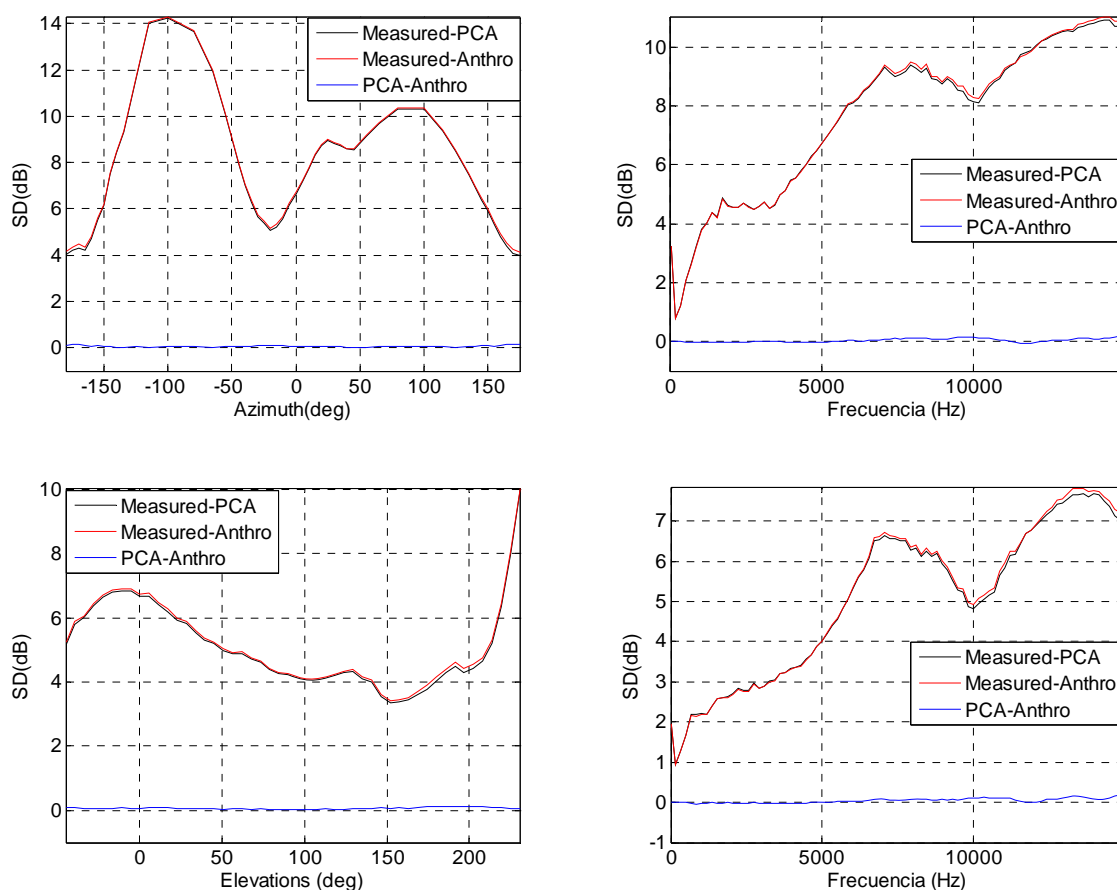


Figura 3: Error medio entre las HRTFs reconstruidas por PCA y reconstruidas por mediciones antropométricas utilizando de la 6 hasta la 10 CP. El panel superior corresponde al plano horizontal y el inferior al plano vertical.

Las primeras 5 CP contribuyen mayoritariamente al error ya que representan el 89.5% de la varianza total del conjunto de HRTFs, mientras que desde la 6 a la 10 CP representan sólo 4.7%. En la Figura 4 se muestran las SD cuando las HRTFs se reconstruyen con las 10 primeras 10 CP (94.3%). Se ve que la SD (línea azul) respecto a las posición en ambos planos (columna de la izquierda) son casi constantes, salvo en algunas regiones en que son levemente mayores. Las SD son menores a 1.5 dB, valor que está por debajo de la “mínima diferencia notable” que un ser humano puede detectar (Durlach and Colburn, 1978). Si se analiza el comportamiento de las SD respecto a la frecuencia (columna de la derecha) se comprueba que son casi cero hasta los 2500 Hz. En un estudio realizado con el mismo grupo de HRTFs por Xu y otros, (2007), comprobaron que las medidas del torso son preponderantes hasta los 4000 Hz mientras que la influencia de las de las orejas crece rápidamente a partir de los 1000 Hz. Además, Shaw (1997) demostró que la interacción de las ondas sonoras con las cavidades y protuberancias de los pabellones auditivos en el plano vertical, dan lugar a una serie de resonancias y cancelaciones entre los 4000 Hz y los 14000 Hz. Los valores de frecuencia en que estas ocurren, varían según el sujeto y la posición de la fuente y cumplen un rol preponderante en la localización de una fuente sonora en el plano vertical. De acuerdo a estas

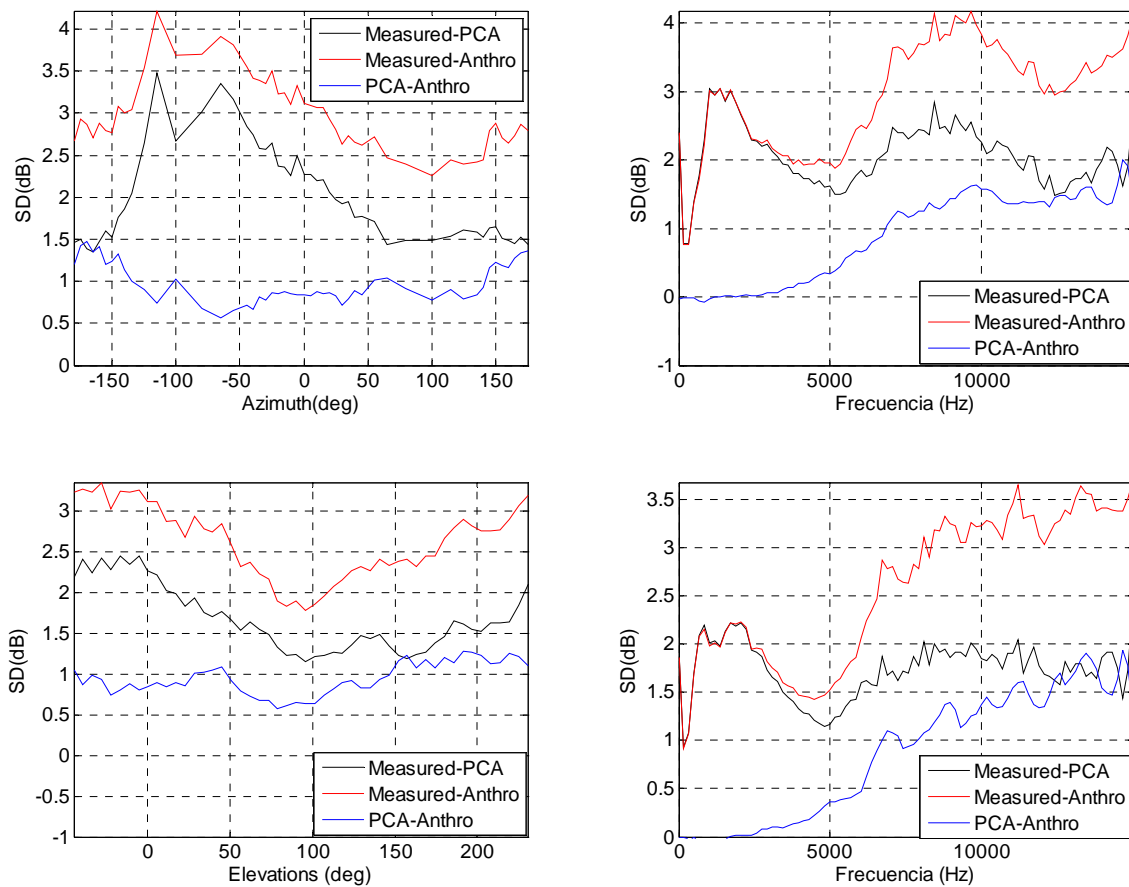


Figura 4: Error medio entre las HRTFs reconstruidas por PCA y reconstruidas por mediciones antropométricas utilizando las 10 CP. El panel superior corresponde al plano horizontal y el inferior al plano vertical.

evidencias, se puede pensar que las mediciones de torso adoptadas son adecuadas, mientras que las de las orejas no.

Para comprobar si la SD mejoraba en las altas frecuencias en el plano vertical, empleamos todas las mediciones de las orejas disponibles. En la Figura 5 y en trazo negro, se muestran las SD obtenidas con las medidas seleccionadas y en trazo rojo, las conseguidas con todas las mediciones. La diferencia es máxima (0.5 dB) alrededor de los 12000 Hz y en el resto de las frecuencias la diferencia es muy pequeña. La mejora nos es significativa y decidimos adoptar definitivamente las medidas seleccionadas por su simplicidad para medirlas. Hay varias razones que podrían contribuir a este error. En primer lugar, las mediciones disponibles son unidimensionales. En segundo lugar, la ausencia de mediciones que podrían ser relevantes y en tercer lugar la influencia de la cabeza en el plano horizontal.

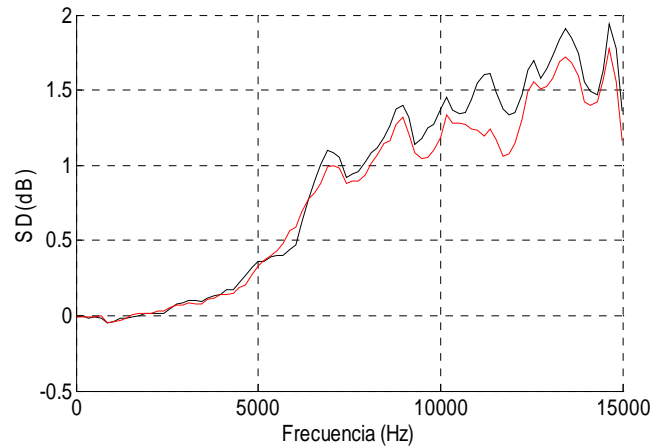


Figura 5: Error medio entre las HRTF medidas y las \tilde{H} utilizando las medidas seleccionadas de las orejas (negro) y todas las disponibles (rojo)

5.1 Comparación con otros estudios

Hu y otros (2007) y Xu y otros (2009) realizaron sendos trabajos sobre el mismo conjunto de HRTFs (CIPIC). Mediante el cálculo de correlación entre las mediciones antropométricas y las HRTF determinaron cuáles son relevantes. En un segundo paso, establecieron el grado de correlación de las mediciones obtenidas del paso anterior. Hu concluye que son suficientes 13 mediciones mientras que Xu sugiere sólo 10. De las mediciones utilizadas por ambos solo coinciden en dos: x3 (ancho de la cabeza) y d1 (alto de la cavidad de la concha). Lamentablemente Hu sólo reporta las SD de algunos sujetos y Xu lo hace para sólo para 10 posiciones que considera representativas. En la Figura 7, comparamos las SD obtenidas por Hu (trazo rojo), Xu (trazo azul) y los nuestros (trazo negro). Se observa en todos los gráficos que la SD son menores cuando se utiliza las mediciones antropométricas propuestas aquí.

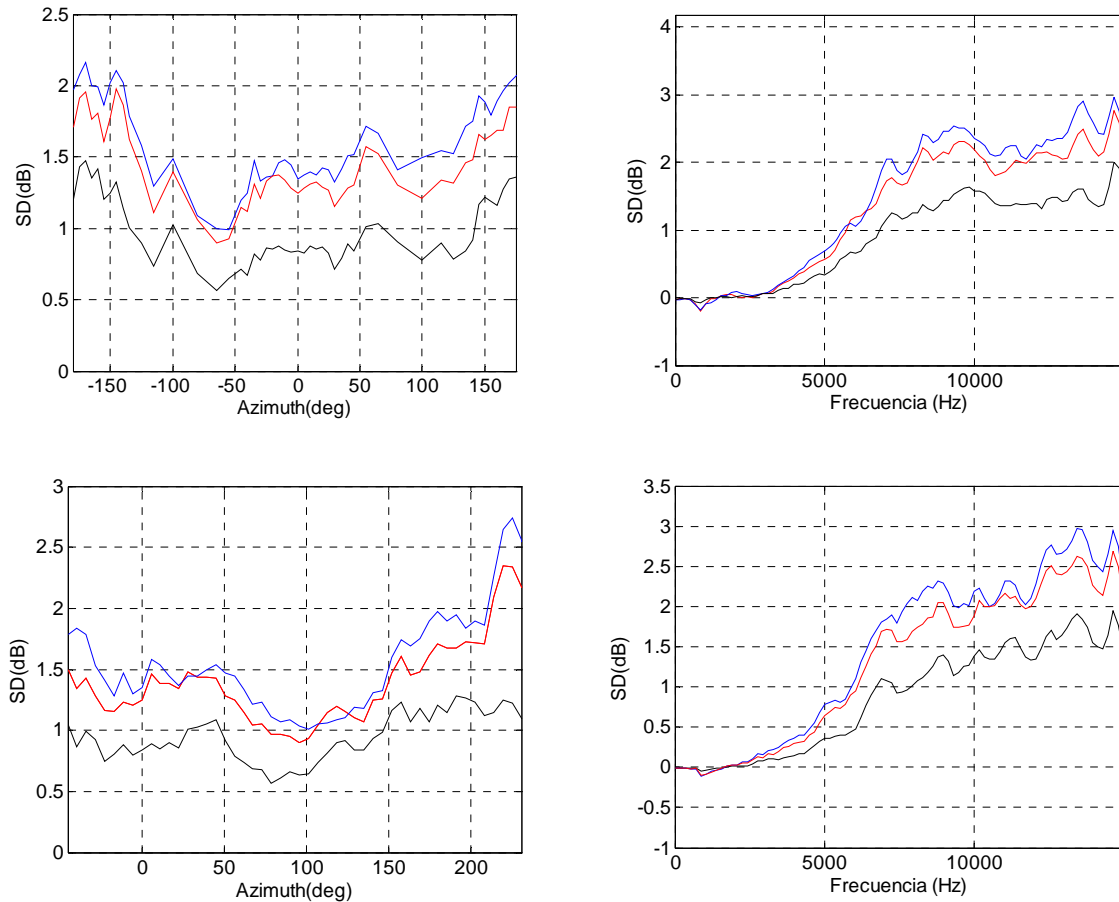


Figura 6: Comparación de los resultados obtenidos por [Hu y otros \(2007\)](#) (trazo rojo), [Xu y otros \(2009\)](#) (trazo azul) y los propios (trazo negro)

6 CONCLUSIONES

El método PCA permite descomponer las HRTFs de un grupo de sujetos en una serie de CP y pesos W de manera que por combinación lineal se puede reconstruir cualquier HRTF del conjunto. Trabajos anteriores nos habían permitido establecer que las 5 CP de mayor peso representaban la influencia de las orejas y que a partir de la sexta CP tenían en cuenta al torso y hombros. Basándonos en estas evidencias ([Ramos et. al, 2008a](#)) y en el trabajo de [Feld y Vorlande \(2009\)](#) se pudo establecer las medidas antropométricas más relevantes que permiten estimar los W para cada sujeto y posición. Mediante regresión lineal multidimensional se pudo estimar los pesos W utilizando 12 medidas antropométricas seleccionadas de las 37 disponibles. Estas mediciones son de fácil determinación y no requieren instrumentos especiales de medición. Los errores entre las HRTFs estimadas por PCA y las obtenidas por regresión son aceptables y están por debajo de la “mínima diferencia notable” en dB que un ser humano puede discriminar. Se demostró que las SD obtenidas son significativamente menores a las conseguidas en otros trabajos. Se comprobó que la media de las SD en el plano horizontal y vertical es independiente de la posición de la fuente mientras que la media de las SD respecto a la frecuencia aumentan por arriba de los 2000 Hz. Esto puede ser atribuido a

que: las mediciones son unidimensionales; la no disponibilidad de mediciones que podrían ser relevantes la influencia de la cabeza en el plano horizontal.

En resumen: es factible predecir los pesos W obtenidos por PCA de un conjunto de HRTF mediante regresión lineal multidimensional de medidas antropométricas de la cabeza, torso y las orejas. Los resultados obtenidos son promisorios restando probar, en el futuro, el comportamiento del método con sujetos externos al grupo de HRTFs procesadas. Finalmente, se deberán realizar pruebas psicoacústicas para validar perceptualmente el método propuesto.

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
- Brown, P. and Duda, R., A Structural Model for Binaural Sound Synthesis. *IEEE Transactions on Speech and Audio Processing*, (6), 5:476-488, 1998.
- 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.
- Feld, Janina and Vorlander, M., Anthropometric Parameters Influencing Head Related Functions. *Acta Acustica united with Acustica*. 95:331-342, 2009.
- Hwang, S. and Park, Y., HRTF customization in the median plane via Principal Components Analysis. *AES 31st International Conference, London, UK*, 2007.
- Hwang, S. and Park, Y., Interpretations on principal components analysis of head-related impulse responses in the median plane, *J. Acoust. Soc. Am. Express Letters*, 123, (4):E165-E171, 2008.
- Hu, H., Zhou, L., Ma, H. and Wu, Z., HRTF personalization based on artificial network in individual virtual auditory space. *Applied Acoustics*, 69:163-172, 2007.
- 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.
- Kuhn, G., Model for the interaural time differences in the azimuthal plane. *J. Acoust. Soc. Am.*, 62 (1):157-167, 1977.
- Martens, W. L., Principal components analysis and resynthesis of spectral cues to perceived direction. *Proceedings of International Computer Music Conference*, San Francisco, CA, 1987
- Mehrgardt, S. and Mellert, V., Transformation characteristics the external human ear, *J. Acoust. Soc. Am.* 61:1567–1576, 1977.
- Middlebrooks, J., Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency. *J. Acoust. Soc. Am.* 106(3):1493-1510, 1999.
- Ramos, O. and Tommasini, F., Simplificación de las Funciones de transferencia de cabeza mediante Análisis de Componentes Principales. *Mecánica Computacional*, Vol. XXVII, 431-442, 2008a.
- Ramos, O., Tommasini, F. and Araneda, M., Análisis de patrones de las funciones de transferencia de cabeza. *VI Congreso Iberoamericano de Acústica (FIA2008)*. Federación Iberoamericana de Acústica (FIA). Buenos Aires, Argentina, 2008b.
- 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).

- Seeber, B. and Fastl, H., Subjective selection of non-individual Head-Related Transfer Functions. *Proceedings of the 2003 International Conference on Auditory Display*. Boston, MA, USA:259-262, 2003.
- Shlens, J., A Tutorial on Principal Components Analysis. Systems Neurobiology Laboratory, Salk Institute for Biological Studies. La Jolla, CA 92037, 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 Andersson, T. Lawrence Erlbaum Associates, Inc., 1997.
- Xu S., Li D., Zeng L., and Salvendy G. (2007). A Study of Morphological Influence on Head-Related transfer Function. *Proceeding of the 2007 IEEM*. 472-476
- Xu, S., Li, Z. and Salvendy G., Identification of Anthropometric Measurements for Individualization of Head-Related Transfer Function. *Acta Acustica united with Acustica*. 95:168-177, 2009.
- Wenzel, E., Arruda, M. Kistler, 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 II: Psychophysical validation, *J. Acoust. Soc. Am.* 85:868–878, 1989b.