



VI Congreso Iberoamericano de Acústica - FIA 2008  
Buenos Aires, 5, 6 y 7 de noviembre de 2008

FIA2008-A092

## **Análisis de Patrones de las Funciones de Transferencia de Cabeza**

Oscar Ramos<sup>(a,b)</sup>  
Fabían Tommasini<sup>(a,b)</sup>  
Mariano Araneda<sup>(a)</sup>

(a) Centro de Investigación y Transferencia en Acústica. UTN Regional Córdoba-UA del CONICET. Mtro López esq. Cruz Roja, Córdoba, Argentina. E-mail: oramos@scdt.frc.utn.edu.ar  
(b) Miembros del CONICET

### **Abstract**

The sound waves arriving at a subject's eardrum contain information which is used by the human being for locating accurately the sound source. These sound waves are transformed by the reflections produced by the torso and shoulders and the diffraction of the head and pinna. These transformations are different for each ear and each subject in function of the position of the sound source in the space. At the same time, these transformations are included in the impulse response (HRIR) measured between a sound source and the eardrums. The purpose of the present work is to study the outstanding aspects of the above mentioned transformations using a statistical method known as Principal Component Analysis (PCA). The aim is to advance in the physic-acoustic interpretation and its perceptual implications in order to design strategies for the development of a listener personal model. In addition, this method allows a significant reduction of the original data collection, making it more attractive for the development of efficient models of the human receptor. For such purpose, a group of HRIRs measured from 5 different subjects in the University of Wisconsin by Doctor Fred Wightman is used.

### **Resumen**

Las ondas sonoras que llegan a los tímpanos de un oyente contienen información que un ser humano utiliza para localizar con precisión una fuente sonora en el espacio. Las reflexiones originadas en el torso y los hombros y la difracción producida por la cabeza y la oreja, contribuyen a transformar las ondas producidas por la fuente. Estas transformaciones son diferentes para cada oído y sujeto y dependen de la posición en el espacio de la fuente sonora. Las Respuestas Impulsivas (HRIR) medidas entre una fuente sonora y los tímpanos engloban dichas transformaciones. Es propósito del presente trabajo, estudiar aspectos relevantes de las transformaciones mencionadas, utilizando un método estadístico conocido como Análisis de Componentes Principales (PCA). Se pretende avanzar en la interpretación físico-acústica y sus implicancias perceptuales de las Componentes Principales, con la finalidad de diseñar estrategias para desarrollar un modelo personalizado del oyente. Además, el método posibilita una reducción significativa del conjunto de datos originales, haciéndolo atractivo para el desarrollo de modelos eficientes del receptor humano. Se utiliza para tal propósito, un conjunto de HRIR medidas a 5 sujetos en la Universidad de Wisconsin por el Dr. Fred Wightman.

## 1 Introducción

Las HRIR caracterizan la transmisión en campo libre de las ondas sonora entre una fuente y los tímpanos de un oyente. Las HRTF (Head Related Transfers Functions) obtenidas por la transformada de Fourier, representan tales características en el dominio de la frecuencia.

Las alteraciones que producen la cabeza, los pabellones auriculares, el torso y los hombros en las ondas sonoras son claves auditivas que el individuo utiliza para localizar con precisión una fuente sonora en el espacio. Como es fácil advertir, estas claves son diferentes para cada sujeto ya que dependen de los rasgos anatómicos de cada uno.

Numerosos autores, desde diferentes abordajes, han intentado desarrollar modelos que se aproximen, en mayor o menor grado, a HRIR medidas en individuos o cabezas y torsos artificiales. Además, estos modelos deben ser *eficientes*<sup>1</sup> y ser capaces de resolver las diferencias individuales mencionadas.

Brown (1998) y Algazi et al. (2001) emplearon filtros paramétricos lineales diseñados a partir de modelos teóricos tanto de la cabeza como de las orejas que fueron validados con mediciones en cabezas y orejas artificiales. Otros autores, utilizan filtros de distintos tipos y orden (FIR y IIR) que son ajustados a mediciones de las HRTF (Kulkarni and Colburn, 2004; Ramos, et. al., 2006).

Modelar la dependencia espacial de las HRTF o HRIR mediante un conjunto de *funciones básicas* ha sido intentado por numerosos autores. Martens (1987) aplicó el método de Análisis de Componentes Principales (PCA) a HRTF preprocesadas con filtros de banda crítica, pero sólo en el plano horizontal. Kistler and Wightman (1992) aplicaron el mismo procedimiento a los logaritmos de las magnitudes de las HRTF. Demostraron que con la combinación lineal de cinco funciones básicas se podía reconstruir el logaritmo de las magnitudes y la fase con un retardo lineal e independiente de la frecuencia. Chen et al. (1995) utilizaron la expansión de Karhunen–Loève para obtener funciones básicas ortogonales a partir de los valores complejos (magnitud y fase) de HRTF medidas en una cabeza artificial y un gato. Estudios similares se realizaron en el dominio del tiempo. Wu et al. (1997) emplearon la expansión de Karhunen–Loève a HRIR obtenidas de un gato. Más recientemente, Park y Hwang (2008) aplican PCA a HRIR de un extenso grupo de sujetos solamente en el plano medio. Determinaron que 12 componentes son suficientes para reconstruir las HRIR, basándose en el error porcentual, fijado arbitrariamente en 5 %, entre las HRIR medidas y las reconstruidas.

Por nuestra parte (Ramos et. al, 2007), aplicamos el método de PCA a un conjunto de HRTF medidas en 5 sujetos y realizamos pruebas psicoacústicas de discriminación entre estímulos sintetizados con las HRTF originales y las generadas por el modelo. El estudio reveló que la utilización de pocas funciones básicas –entre 20 y 25- son suficientes para preservar las características fundamentales de fase y magnitud de las HRTF originales. Se concluyó además, que el error medio cuadrático entre los logaritmos de las magnitudes de la originales y las reconstruidas, integrada a lo largo de todo el espectro, era menor a 2 dB. Decíamos además que, la cantidad de funciones básicas estaba determinada por el ajuste del oído oculto por la cabeza y que se debían realizar estudios complementarios para reducir aún más el número de funciones básicas.

Es propósito del presente trabajo, avanzar en la interpretación físico-acústica y sus implicancias perceptuales de las funciones básicas con la finalidad de diseñar estrategias que

---

<sup>1</sup> El tiempo de procesamiento y los requerimientos de memoria deben ser mínimos.

reduzcan el orden del modelo. Se utilizaron el mismo grupo de HRIR medidas a 5 sujetos en la Universidad de Wisconsin por el Dr. Fred Wightman.

El trabajo está organizado de la manera siguiente. En el punto 2, se introduce al lector en las bases del Análisis de Componentes Principales y su aplicación a nuestro caso; en el punto 3 se analiza objetivamente el error de aproximación; en el punto 4, se discute las características de las componentes principales y en el punto 5 se analiza la variabilidad intersujetos.

## 2 Análisis de Componentes Principales (PCA)

El análisis de Componentes Principales es un método estadístico que apunta, por un lado, a reducir la dimensión de un conjunto de observaciones o mediciones empíricas intercorrelacionadas, y por el otro, a obtener una serie de *funciones básicas* o Componentes Principales (CP) independientes entre sí, que evidencian características comunes del conjunto. El procedimiento consiste en proyectar la matriz de la covarianza del conjunto de datos originales sobre un grupo de vectores ortogonormales en el sentido en que dicha covarianza sea máxima (Shlens, 2005).

Un miembro en particular de los datos se puede reconstruir mediante la combinación lineal de las CP y un conjunto de *pesos W* que reflejan la contribución relativa que cada CP hace a ese miembro en particular.

Sea *H* la matriz de los datos empíricos de dimensión *nxm*; *CP* la matriz de las funciones básicas de dimensión *mxm* y *W* la matriz de los pesos de dimensión *nxm* tendremos:

$$H = CPW^t \tag{1}$$

**Tabla 1:** Estructura de los datos

Sujeto	Posición	Frecuencias			
AFW	1	f1	f2	...	fm
	...	...	..	...	...
	p	f1	f2	...	fm
SJX	1	f1	f2	...	fm
	...	...	..	...	...
	p	f1	f2	...	fm
SOS	1	f1	f2	...	fm
	...	...	..	...	...
	p	f1	f2	...	fm
SOU	1	f1	f2	...	fm
	...	...	..	...	...
	p	f1	f2	...	fm
SOW	1	f1	f2	...	fm
	...	...	..	...	...
	p	f1	f2	...	fm

En nuestro caso, los datos empíricos se componen de HRIR medidas a 5 sujetos 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 o acimut que pasa por ambos oídos) 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). Cada HRIR son secuencias de 256 puntos que muestreadas a 44100 Hz, dan una longitud de 5,8 milisegundos.

Antes de aplicar el procedimiento, las HRIR fueron convertidas al dominio de la frecuencia mediante la transformada de Fourier (HRTF). El ancho de banda fue limitado a 15000 Hz y la resolución en frecuencia fue de 172,4 Hz. Por lo tanto, se utilizaron las primeras 88 componentes en frecuencias. En la tabla 1 se muestra la estructura de los datos.

Cada sujeto contribuye con 505 (posiciones) x 2 (oídos) = 1010 espectros complejos de 88 componentes de frecuencias. Por lo tanto, la matriz global con los datos de los 5 sujetos será: 5x1010=5050 filas (*n*) por 88 columnas (*m*) igual a 444.400 números complejos.

Se utilizó la función *princomp* de Matlab y se obtuvo una matriz de 88 CP de 88 frecuencias y la matriz *W* de 88 pesos (uno para cada CP) por 5050 posiciones (*n*). Cualquier HRTF del conjunto, se reconstruye totalmente haciendo:

$$\hat{H}(f, \varphi, \theta) = H_0(f) + \sum_{i=1}^{88} CP_i(f)W_i(\varphi, \theta) \quad (2)$$

en donde  $H_0$  es la media global del conjunto que fue calculada y restada de cada componente antes de aplicar el procedimiento.  $H_0$  refleja los rasgos comunes e independientes de la posición de la fuente sonora. Cada CP representa una fracción de las características del conjunto y se expresa como un porcentaje de la varianza total. Las CP se ordenan de mayor a menor de manera que CP1 es la de mayor varianza, mientras que CP<sub>88</sub> es la de menor varianza. Los pesos  $W$  determinan la contribución que hace cada CP a una posición y un sujeto determinado. En la tabla 2 se muestra el porcentaje de la varianza con que contribuye

**Tabla 2:** Porcentajes de la varianza individual de cada CP, más la acumulada

CP	Varianza	Acumulada	CP	Varianza	Acumulada
1	48,44	48,44	9	0,77	97,48
2	19,50	67,94	10	0,54	98,02
3	12,36	80,31	11	0,44	98,46
4	7,00	87,31	12	0,33	98,79
5	4,1	91,41	13	0,26	99,04
6	2,67	94,08	14	0,20	99,24
7	1,58	95,66	15	0,15	99,40
8	1,05	96,71	16	0,12	99,52

cada CP, más la varianza acumulada hasta la de orden 16.

Para conseguir una efectiva reducción de la dimensión de los datos originales, se debe determinar un número  $L$  de CP, bastante menor a 88, que permita reconstruir una HRTF dada dentro de un error aceptable.

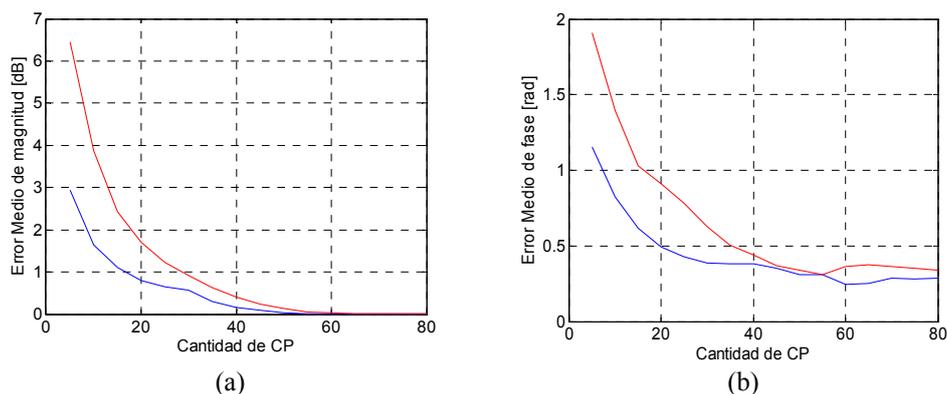
Algunos autores utilizan tantas CP de manera que la suma de sus varianzas resulte, por ejemplo, igual al 90 %. Por nuestra parte, sostenemos que un error aceptable es aquel que resulta de evaluaciones psicoacústicas, esto es, que estímulos sonoros sintetizados con las HRIR medidas y las reconstruidas por el modelo no puedan ser discriminados por el oyente (Ramos et. al, 2007).

### 3 Análisis Objetivo del error de reconstrucción

Las magnitudes de las HRTF contienen una serie de finos detalles espectrales (picos y ceros) que varían sistemáticamente con la posición de la fuente sonora. Resonancias y reflexiones ocasionadas en las cavidades de las orejas son las principales causantes de estas variaciones. La diferencia de magnitud entre el oído derecho y del izquierdo (conocida como diferencia de nivel interaural: ILD) es ocasionada por difracción de la cabeza y es una de las claves primarias que el ser humano utiliza para localizar una fuente en el espacio. De igual manera lo es la diferencia de fase que evidencia la diferencia de tiempo en que las ondas sonoras arriban a ambos oídos (diferencia de tiempos interaural: ITD). Estas características de las ondas sonoras mencionadas son claves fundamentales que un modelo del oyente debe preservar. En consecuencia, el error de aproximación debe ser evaluado tanto en magnitud como en fase.

Estudios previos que realizamos, revelaron que la distribución de los errores de magnitud y fase es independiente de la cantidad de CP utilizadas. Estos varían de manera sistemática con la posición de la fuente y el oído (Ramos y Tommasini, 2008). El error es mayor para el oído contralateral (el oído que está en el hemisferio opuesto a la fuente y oculto por la cabeza) para elevaciones negativas (debajo del plano horizontal) y disminuyen gradualmente para elevaciones positivas, siendo mínimo cuando la fuente esta sobre la cabeza.

Analizaremos a continuación como se comporta el error del oído ipsilateral y el contralateral en función de la cantidad de CP utilizadas.



**Figura 1.** Error medio de los cinco sujetos del oído contralateral (azul) y del oído ipsilateral (rojo) en función de la cantidad de CP utilizadas en la reconstrucción: a) magnitud y b) fase

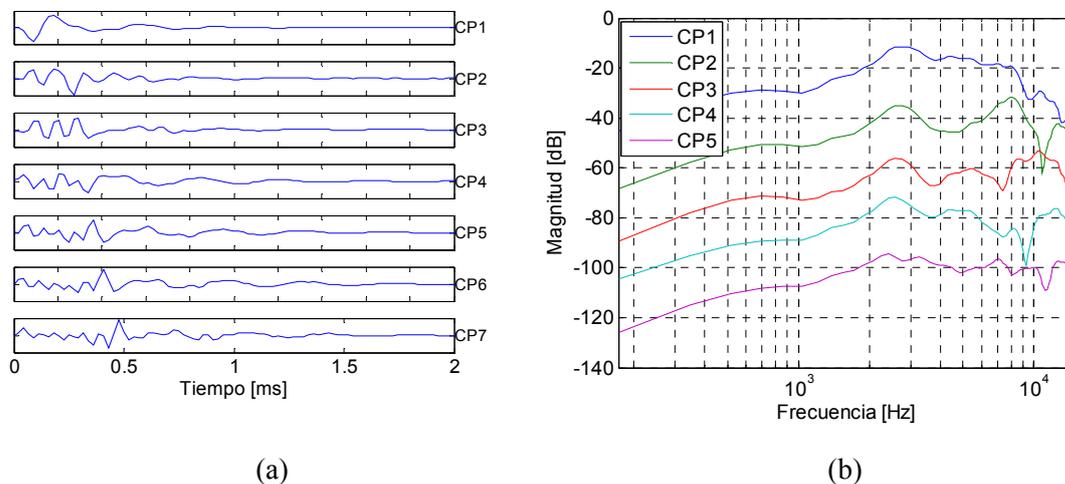
Para calcular el error del *oído contralateral* se agruparon los errores del oído derecho de todos los sujetos para ubicaciones de la fuente en el hemisferio izquierdo con los errores del oído izquierdo para la fuente ubicada en el hemisferio derecho. Para evaluar *el oído ipsilateral*, se agruparon los errores de cada oído para la fuente ubicada en el mismo hemisferio. En la figura 1 se muestran los resultados para la magnitud y la fase. En ambos gráficos, el error del oído contralateral (rojo) es siempre mayor que el del oído ipsilateral (azul). Después de 60 CP, los errores de magnitud (Fig. 1a) se estabilizan y se igualan; mientras que los errores de fase (Fig. 1b) parecen estabilizarse después de 70 CP. El error del oído contralateral, decrece más rápido entre la 5 y la 15 CP, mientras que el oído ipsilateral lo hace entre la 5 y la 10. El error del oído ipsilateral toma valores menores a 2 db para 10 CP, mientras que los errores contralaterales recién alcanzan estos valores para 20 CP. Esta diferencia en la convergencia del error, gobierna el orden definitivo del modelo. Estos resultados objetivos del error concuerdan con los obtenidos de las pruebas psicoacústicas realizadas en un estudio anterior (Ramos et. al, 2007) en que la utilización de 20 a 25 CP eran suficientes y que el error de ajuste era menor a 2 db.

#### 4 Características de las Componentes Principales

Las ondas sonoras que llegan a los tímpanos de un oyente son producto de una serie de eventos sonoros que recorren diferentes caminos. Hay un sonido directo; reflexiones y resonancias en las orejas; difracción en la cabeza y finalmente reflexiones en los hombros y el torso. Es imposible caracterizar estos eventos debido a que como están solapados son difíciles de aislar. Estudios realizados por Brown y Duda (1998) en una cabeza y torso artificial y Algazzi et. al (2001) en orejas artificiales, demostraron que la respuesta atribuida a las orejas se manifiestan en los primeros 500  $\mu$ s de las HRIR y que las reflexiones en el torso y los hombros se hacen presentes alrededor del milisegundo. Por otro lado, Shaw (1997) determinó a partir de mediciones realizadas en orejas de humanos, un conjunto de resonancias que se producen en las protuberancias de las orejas y que llamo modos normales.

Para facilitar el estudio de las CP se convirtieron al dominio del tiempo mediante transformada inversa de Fourier. En la figura 2a se grafican las respuestas temporales de las primeras siete CP. Se observa en todas, que la energía mayor está concentrada en los primeros 500  $\mu$ s y que a partir de la quinta son evidente reflexiones muy definidas. Se puede atribuir a CP1 al sonido directo, esto es, la onda que recorre el camino más corto entre la fuente sonora

y los tímpanos. No se notan retardos de la CP2 a la CP4, pero si que la energía se sostiene durante un tiempo, comportamiento típico de sistemas resonantes. En la quinta, sexta y séptima son notables reflexiones casi discretas a intervalos regulares.



**Figura 2.** En (a) se muestran las respuestas temporales de las CP más significativas y en (b) sus magnitudes en dB

En la figura 2b se grafican las magnitudes de la CP1 a la CP5 en el dominio de la frecuencia. Estas representan el 91,41 % de la varianza de las HRTF originales. Para mayor claridad, las ordenadas están desplazadas por valores arbitrarios. Se observa, en general, la existencia de picos y ceros marcados. En primer lugar y en las cinco, se ven picos que varían entre los 2400 Hz y los 2700 hz que corresponden a la resonancia principal del conducto auditivo. En la CP2 existe otro pico alrededor de los 8 khz; en la CP3 en 10 khz y en la CP4 en 4.4 Khz y 12 Khz. Estas resonancias coinciden aproximadamente con los modos de resonancias de la oreja definidos por Shaw como modos 1, 2, 3, 4 (4.1 Khz; 7.1 Khz; 9.6 Khz y 12.1 khz). Además define los modos 5 y 6 en 14.4 Khz y en 16.7 Khz respectivamente. El modo 5 no se observa mientras que el modo 6 está fuera del rango del análisis. Shaw determina estas resonancias a partir de promedios de mediciones realizadas en oídos de humanos y de modelos mecánicos del pabellón auricular (Shaw, 1997). Ceros muy notorios se notan en 10.1 khz en CP2; en 7.3 Khz en CP3 y en 9.3 khz en CP4 debido a cancelaciones producidas en las cavidades y protuberancias de las orejas.

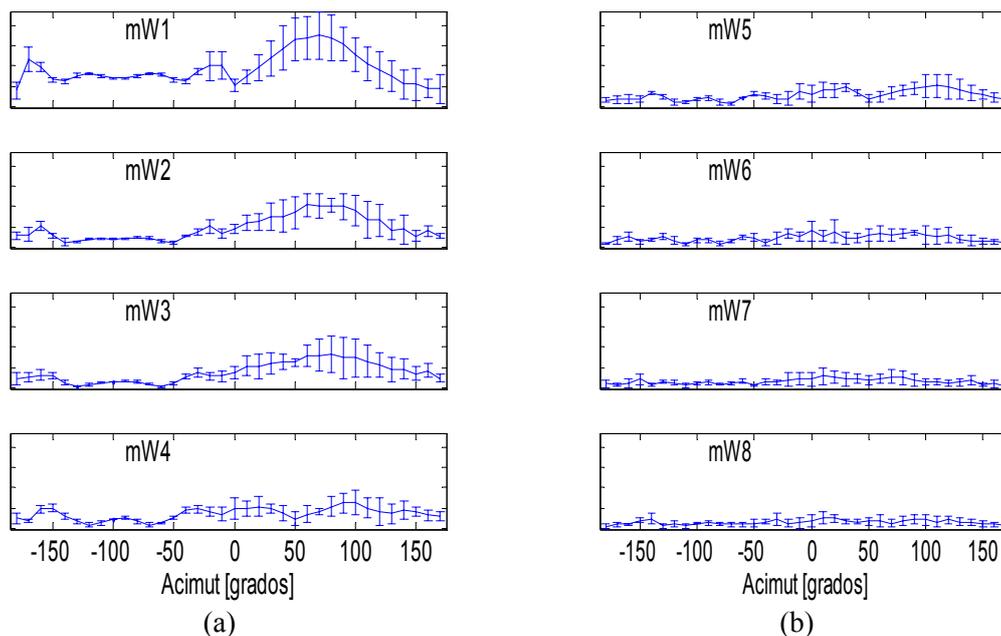
Por razones de espacio no se muestran gráficos de las CP subsiguientes. Estas tienen, en el dominio del tiempo, el aspecto de reflexiones bien definidas que se desplazan en el tiempo, compatibles con reflexiones en el torso y sus magnitudes muestran un perfil característico de los filtros “peine”.

## 5 Variabilidad Intersujeto

Las CP son comunes a todos los sujetos e independientes de la posición de la fuente sonora. Por el contrario, los pesos  $W$  de cada CP son doblemente dependientes: varían con la posición y el sujeto. Dada una posición y un sujeto en particular, existe un peso para cada CP. Si bien los  $W$  son valores complejos, para el análisis utilizaremos los módulos. Se han calculado los valores medios de los pesos ( $mW$ ) de los 5 sujetos para diferentes posiciones. Las posiciones analizadas fueron para el plano horizontal (elevación igual a cero) y para el plano vertical (acimut igual a cero).

En la Fig. 3 se grafica desde mW1 a la mW8 ( $\pm \sigma$ ) para el plano horizontal y el oído derecho. Se observa que mW1 (Fig. 3a) crece monótonamente para acimuts positivos entre 0 grados y los 70 grados y que luego disminuye, de igual manera, hasta los 170 grados. Este aumento es congruente ya que el oído analizado está en el mismo hemisferio de la fuente. Las desviaciones estándares intersujetos son significativas para los acimuts positivos. Cuando la fuente está en el hemisferio opuesto y el oído analizado esta oculto por la cabeza, se advierte cierta simetría de la media respecto a -90 grados. Es casi constante entre -150 grados y -50 grados aumentado levemente entre -180 grados y -150 grados y -50 y 0 grados. Una particularidad notable es que las desviaciones estándares son significativamente menores para acimuts negativos que para los positivos. Debido a la asimetría de mW1 respecto a 0 grado, podemos afirmar que CP1 contribuye a la percepción de lateralidad de la fuente (derecha e izquierda) en el plano acimutal.

Comportamiento similar se observa para las restantes con la particularidad que el punto de simetría se desplaza desde 0 grados, hacia los azimuts negativos. Sin embargo se mantiene la tendencia observada para mWP1, esto es, la media y las desviaciones crecen a la derecha del punto de asimetría hacia los acimuts positivos. La variabilidad intersujeto se reduce significativamente entre mW5 y mW8 y sus medias tienden a ser constantes a lo largo de todas las posiciones analizadas.

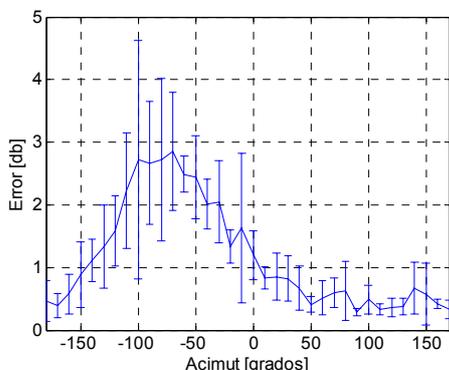


**Figura 3.** Medias y  $\pm \sigma$  de los pesos  $W$  para los cinco sujetos en el plano azimutal

Llama la atención que entre -150 grados y -50 grados la variación de las medias sean pequeñas al igual que sus desviaciones. Sólo se observa en mW2 a la mW6 un pico en -90 grados, debidos a la difracción que se produce en la cabeza. A partir de estas evidencias se puede afirmar que en la zona mencionada, los pesos  $W$  no reflejan ni las diferencias intersujetos ni las diferencias interposición. En la figura 4 se grafica el error de aproximación y  $\pm \sigma$  en el plano horizontal cuando se utilizan 25 CP. Contrariamente a lo esperado, se ve que los errores y las desviaciones son bastantes significativos para los acimutis negativos.

Esto corrobora que el modelo es deficiente cuando el oído está oculto por la cabeza mientras que lo hace con precisión cuando el oído está directamente enfrente a la fuente (acimuts positivos), en donde los errores y las diferencias intersujeto son menores.

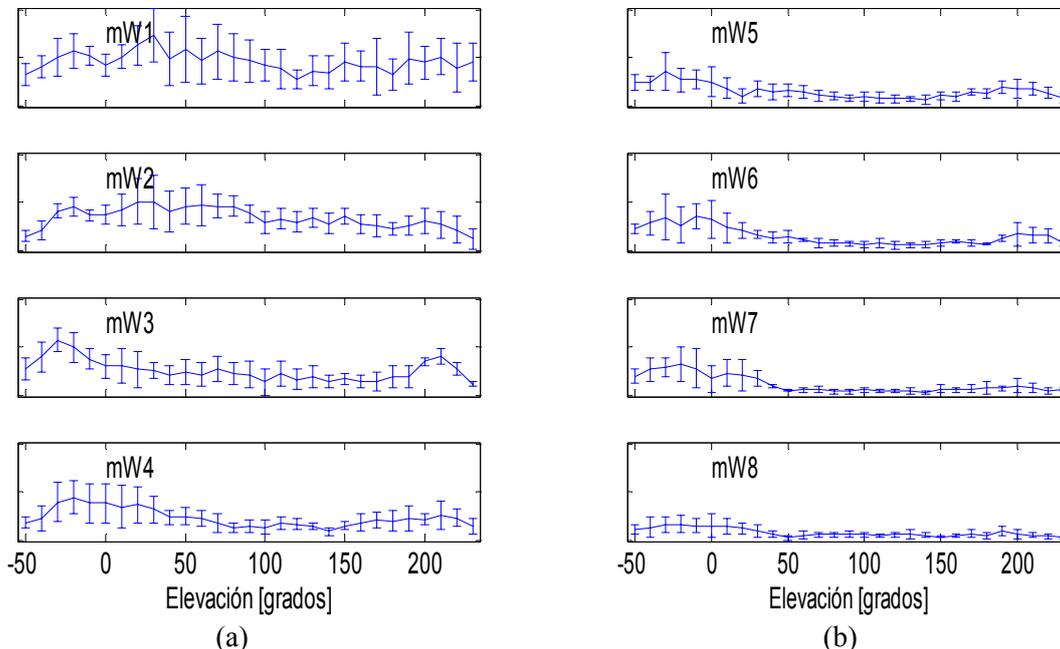
En la Figura 5 se grafican las medias de los pesos en el plano vertical (acimuts=0 grados). La media mW1 crece desde -50 grados (debajo del plano horizontal) hasta +30 grados y luego disminuye hasta un mínimo en +120 grados para crecer suavemente hasta los 230 grados. De igual manera, mW2 crece desde -50 grados alcanza un máximo en +30 y luego disminuye hasta los +230 grados. En ambas, las variaciones intersujetos son significativas en todo el rango de elevaciones analizadas. mW2 es asimétrica respecto a +100 grados (casi sobre la cabeza). Por lo tanto, CP2 contribuye a la percepción de adelante-atrás. mW3 es casi constante por arriba del plano azimutal (entre 0 y 180 grados) y es variable entre 0 y -50 grados y entre 190 y 230 grados (debajo del plano



**Figura 4.** Error medio de aproximación y  $\pm\sigma$  del oído derecho para el plano azimutal

azimutal). Podemos afirmar que CP3 contribuye con claves que determinan la percepción de arriba-abajo. Se aprecia una tendencia común a partir de mW4 y hasta mW8. Las medias de los pesos crecen desde -50 grados; alcanzan un máximo y luego disminuyen hacia un valor constante hasta aproximadamente 150 grados. Esto parece definir dos zonas: una adelante, de

grados y luego disminuye hasta un mínimo en +120 grados para crecer suavemente hasta los 230 grados. De igual manera, mW2 crece desde -50 grados alcanza un máximo en +30 y luego disminuye hasta los +230 grados. En ambas, las variaciones intersujetos son significativas en todo el rango de elevaciones analizadas. mW2 es asimétrica respecto a +100 grados (casi sobre la cabeza). Por lo tanto, CP2 contribuye a la percepción de adelante-atrás. mW3 es casi constante por arriba del plano azimutal (entre 0 y 180 grados) y es variable entre 0 y -50 grados y entre 190 y 230 grados (debajo del plano



**Figura 5.** Medias y  $\pm\sigma$  de los pesos  $W$  para los cinco sujetos en el plano medio.

medias más grandes, entre -50° y +50° y otra atrás, con medias más pequeñas entre 150° y 230°. En la zona del frente, la fuente sonora “ilumina” con mayor intensidad al torso y la

contribución de las reflexiones en él son importantes. En la zona de atrás, ocurre algo similar con la espalda pero sus reflexiones son apantalladas por la parte posterior de las orejas.

## 6 Análisis de resultados y conclusiones

Con la finalidad de desarrollar un modelo eficiente del oyente, se estudió un conjunto de HRIR obtenidas de un grupo de cinco sujetos. El propósito de este trabajo fue indagar sobre patrones subyacentes en las HRTF como así también las diferencias intersujetos. Se obtuvo, mediante el método de Análisis de Componentes Principales, un conjunto de funciones básicas (CP) independientes entre sí y de la posición de la fuente. Estas capturan aspectos comunes del conjunto y su relevancia se refleja por el porcentaje de la varianza que cada una representa. Por otro lado, un conjunto de *pesos* determinan la contribución que hace cada CP a una ubicación de la fuente y a un sujeto en particular.

Cinco CP representan el 91,4 % de la varianza en magnitud y fase del conjunto de HRTF mientras que 10 lo hacen con el 98,02 %. Sin embargo, a pesar de lo elevado de estos porcentajes demostramos en estudios anteriores que se necesitan entre 20 y 25 CP para que un sonido sintetizado con las HRTF medidas y reconstruidas no puedan ser discriminados (Ramos et. al, 2007). Indudablemente, CP de orden mayor a 10, contribuyen con pequeños y finos detalles que el oyente puede detectar. Se hace necesario estudiar el orden necesario mediante pruebas psicoacústicas de localización en donde el sujeto debe juzgar sobre la ubicación de una fuente sonora en el espacio y comparar sus rendimientos con las HRTF reales y las aproximadas.

La variación del error de ajuste del oído ipsilateral toma valores menores a 2 dB para 10 CP, mientras que el oído contralateral (oculto por la cabeza) lo hace recién para 20 CP. Esto se debe a que el efecto de difracción que produce la cabeza provoca en el oído oculto perfiles espectrales y seguramente alinealidades muy difíciles de modelar y reproducir.

Del análisis de las CP de mayor orden en el dominio del tiempo se demostró que las cinco primeras engloban el efecto de los pabellones auriculares y que las restantes muestran reflexiones casi discretas espaciadas en el tiempo y que responden a reflexiones en los hombros y torso. Además se determinó, una serie de resonancias, los modos normales de Shaw (1997), en las primeras 5 CP.

Del estudio de las medias de los pesos  $W$  y sus desviaciones estándares, se desprende que las diferencias entre sujetos son importantes hasta la octava CP. En el plano azimutal las medias y las desviaciones son bastantes mayores para el oído que está en el mismo hemisferio de la fuente. En particular, la media de los pesos correspondientes a CP1 es simétrica alrededor de 0 grados y por ende, contribuye a la percepción de lateralidad, esto es, derecha-izquierda. Cuando el oído está oculto por la cabeza, en especial entre -50 y -150 grados, las medias son casi constantes y las desviaciones pequeñas. El análisis del error en ese plano reveló que son máximos para esas posiciones al igual que las desviaciones. Ninguna CP representa por si sola el comportamiento de ese oído siendo necesario el aporte de varias. Estas evidencias están respaldadas por pruebas psicoacústicas que realizamos previamente (Ramos et. al, 2007). En aquella oportunidad, las pruebas de discriminación realizadas con el oído ipsilateral revelaron que la cantidad de CP necesarias estaban comprendidas entre 10 y 15 mientras que para el oído contralateral fueron necesarias casi en doble: entre 20 y 25.

Del análisis de las medias y desviaciones en el plano vertical se concluyó que CP2 contribuye a la percepción de adelante-atrás mientras que CP3 contribuye a la percepción de arriba-abajo.

Finalmente, se piensa que se pueden construir HRTF personalizadas a partir de utilizar aquellas componentes principales que muestran una mayor diferencias intersujetos,

expresadas por las desviaciones significativas de sus pesos. Para ello, estamos trabajando con una población mayor de HRTF medidas a sujetos con diversas medidas antropométricas.

### Referencias

- Brown, Phillip. and Duda, Richard (1998). "A Structural Model for Binaural Sound Synthesis". IEEE Transactions on Speech and Audio Processing, (6), 5, 476-488.
- Algazi, V.; Duda, Richard; Thompson, Dennis and Morrison, Reed (2001). "Structural Composition and decomposition of HRTF". IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 103-106.
- Kulkarni, A. and Colburn, S (2004). "Infinite-impulse-response models of the head-related transfer function". *J. Acoust. Soc. Am.*, 115(4), 1714-1728.
- Ramos, Oscar., Jasá, Victor., Gorgas, Juan. and Céspedes Daza, Daniel. (2006). Simulación Binaural de Espacios Acústico. Memorias del *V Congreso Iberoamericano de Acústica*. Santiago, Chile.
- Martens, W. L (1987). Principal components analysis and resynthesis of spectral cues to perceived direction. *Proceedings of International Computer Music Conference*, 274-281.
- Kistler, D. and Wightman, F (1992). A model of head-related transfer functions based on principal components analysis minimum-phase reconstruction. *J. Acoust. Soc. Am.*, (91), 3, 1637-1647.
- Chen, J., Van Veen, B. D. and Hecox, K. E (1995). A spatial feature extraction and regularization model for the head-related transfer function. *J. Acoust. Soc. Am.* (97), 1, 439-452.
- Wu, Z., Chan, F. H.Y., Lam, F. K., and Chan, J. C. K (1997). A time domain binaural model based on spatial feature extraction for the head-related transfer function, *J. Acoust. Soc. Am.* 102, 2211-2218.
- Hwang, S. and Park, Y (2008). Interpretations on principal components analysis of head-related impulse responses in the median plane, *J. Acoust. Soc. Am. Express Letters*, 123, (4), E165-E171.
- Ramos, O., Calvo, G. and Tommasini, F (2007). Modelo Acústico de Cabeza y Torso Mediante Análisis de Componentes Principales. *Mecánica Computacional*, Vol. XXVI, pp 46-58.
- Shlens, J (2005). A Tutorial on Principal Components Analysis. Systems Neurobiology Laboratory, Salk Institute for Biological Studies. La Jolla, CA 92037.
- Ramos, O., and Tommasini, F (2008). Simplificación de las Funciones de transferencia de cabeza mediante Análisis de Componentes Principales. *Mecánica Computacional*, en prensa.
- Shaw, E. (1997). 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.