



## Diseño e implantación de un sistema de sonido tridimensional basado en la HRTF

Ivana Rossell Turull. Robert Barti. David Sansegundo.  
Departamento de Acústica Ingeniería La Salle.  
Universidad Ramón Llull. Pg. Bonanova,8. 08026 Barcelona.

### Introducción

Virtual Reality is one of the hottest buzzwords today in electronics industry. Until recently, the focus of these efforts was in providing stereoscopic three-dimensional graphics to stimulate our sense of vision. However, to create believable worlds, visual cues are not sufficient, the field of auralization, -three dimensional sound- is necessary. There are a number of motivations for developing auralization systems in addition to the advent of virtual reality. The three dimensional processing adds an auxiliary creative element to be manipulated by the commercial musician or record producer, providing a new realm of entertainment to explore.

In this paper, a tridimensional sound system for performing static sound source placement with headphone playback using the filtering by HRTF is developed.

### Localización de fuentes sonoras

Hasta hace relativamente poco la localización de fuentes sonoras se veía en función exclusiva de dos parámetros: ITD (Interaural Time Differences) y ILD (Interaural Level Differences). Ambos parámetros se basan en las diferencias interaurales de la persona humana; es decir, las diferencias de tiempo y amplitud de las señales recibidas en cada oído. Actualmente, por lo que respecta a la sensación del "oyente", a la experiencia subjetiva delante del suceso sonoro, se valoran otros factores que contribuyen de forma importantísima:

- La naturaleza de la fuente sonora:
- El ambiente acústico del oyente.
- La influencia de los hombros del oyente (torso).
- La influencia de la cabeza del oyente en la onda incidente.
- La influencia de las orejas del oyente para ultimar la sensación auditiva. Este factor es importantísimo debido a la modelación espectral que introduce el pabellón de la oreja dependiendo de la procedencia de la señal.
- La experiencia del oyente en este tipo de experimentos.

Todos estos factores que interaccionan en cada persona en el momento de recibir un estímulo acústico pueden englobarse en una función de transferencia que llamamos HRTF y caracterizar de esta manera la procedencia de un evento sonoro.

### La función de transferencia relacionada con la cabeza

Las teorías clásicas de localización del sonido ignoraban el filtrado que introduce el oído y mantenían que la posición aparente de una fuente sonora estaba determinada íntegramente por las diferencias interaurales. Estudios posteriores han demostrado que las variaciones del espectro causados por este filtrado son muy importantes, especialmente para localizar señales en el plano medio, donde los ITDs y ILDs son

minimizados. Intentando dar una definición clara de la función, podríamos entender la HRTF (Head Related Transfer Function) como el filtro que introduce una persona cuando percibe un sonido. Por lo tanto en su viaje hasta el canal auditivo el sonido pasará por este filtrado humano y, según la procedencia del estímulo, la onda incidente sufrirá diferentes reflexiones sobre todos los diferentes obstáculos provocando su filtro exclusivo; el de aquella posición espacial y el de aquella persona en concreto.

Las componentes de fase de las HRTFs incorporan las diferencias interaurales de tiempo y son importantes para determinar el *azimut* aparente de la fuente. Las componentes de magnitud codifican las diferencias interaurales de intensidad y las variaciones en la forma del espectro, que es de donde se deriva la elevación aparente de la fuente.

#### Método de cálculo

Si representamos  $x_1(t)$  como la señal eléctrica que enviamos a un altavoz en campo libre y  $y_1(t)$  la señal resultante captada por un micrófono en el canal auditivo (una sonda en nuestro caso). Si definimos  $x_2(t)$  como la señal eléctrica que enviamos a un auricular y  $y_2(t)$  la señal captada por la misma sonda. Pasando al dominio frecuencial y aplicando los principios clásicos de filtrado lineal podemos expresar las señales recibidas de la siguiente manera:

$$Y_1 = X_1 \cdot H_{\text{alt.}}$$

$$Y_2 = X_2 \cdot H_{\text{aur}}$$

El resultado de igualar  $X_1$  y  $X_2$  nos lleva a la expresión final de la función de transferencia:

$$\text{HRTF} = \frac{Y_1}{Y_2}$$

El método teórico descrito es el que se ha seguido en la realización práctica del filtro. Es importante recalcar la captación de la señal lo más próximo posible al canal auditivo gracias a una sonda, ya que es allí donde la información sobre la procedencia del estímulo es completa. Hemos despreciado el efecto de esta sonda sobre la señal gravada ya que se cancela en la expresión final. Sólo hemos asegurado (limitando las dimensiones de la sonda) una relación señal/ruido suficiente en toda la banda de frecuencias. Las respuestas en frecuencia del altavoz y de los auriculares utilizados así como el señal enviado para calcular la HRTF se han hecho con secuencias M.L.S.

#### Captación y generación de la señal

Como hemos comentado, para implantar la HRTF necesitamos la información en la entrada del canal auditivo con el fin captar las reflexiones que se producen en el pabellón de la oreja. La sonda que se ha construido se basa en un tubo capilar de 2.5mm de diámetro y 30mm de longitud que está introducido a una estructura cónica de cobre que encaja por el otro extremo con el micrófono de medida. Éste, introducido con cuidado en los oídos del individuo, recogerá la información deseada.

Respecto a la generación de la señal, la correcta colocación del altavoz es muy importante para obtener resultados fidedignos. Necesitamos un sistema que sitúe el altavoz en diferentes ángulos de elevación y *azimut* conservando siempre la misma distancia respecto al receptor y capaz de reproducir la misma situación cuando se necesite sin introducir diferencias importantes. Se ha construido un aparato que consta de un brazo de 3m de longitud unido por el centro a un eje rotativo que permitirá elevar la fuente a cualquier ángulo. En el extremo del brazo se ha colocado el altavoz siempre encarado al oyente que estará sentado

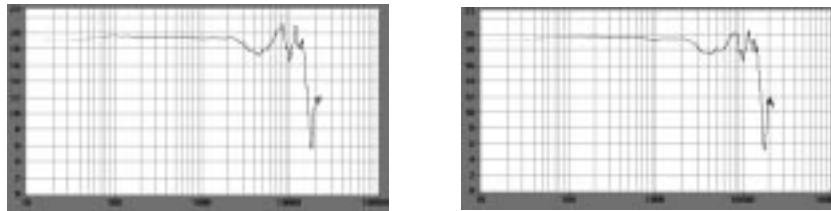


a 1.5m del estímulo. El barrido azimutal se realizará situando la persona en la posición que interese. A continuación mostramos unas fotografías de la sonda construida y del sistema de posicionamiento del altavoz.

#### Resultados obtenidos

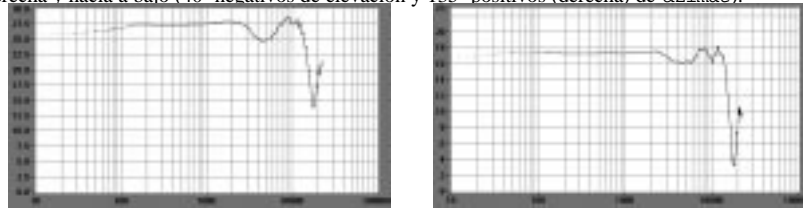
Cuando se habla de localización de fuentes en el espacio siempre se hace referencia a las fuentes procedentes del plano medio como a un grupo diferenciado susceptible a la conocida confusión back-front (fuentes dentro del cono de confusión).

Las HRTFs situadas en el plano medio (ángulo azimutal de cero grados) evidencian su propiedad de individualidad ya que el efecto tridimensional se agudiza cuando se realiza la escucha con la propia función de transferencia. A continuación presentamos unas gráficas que muestran las HRTFs de dos personas a 70° de elevación.



Las diferencias entre ambas ilustraciones se deben al cambio de individuo.

Las simulaciones obtenidas con otras posiciones han dado resultados más espectaculares sobre la mayoría de los voluntarios sujetos a las pruebas. Un ejemplo se muestra en las siguientes figuras donde las HRTFs (derecha e izquierda) mostradas dan a la señal que sea filtrada con ellas la sensación de proceder de detrás, hacia la derecha y hacia a bajo (40° negativos de elevación y 135° positivos (derecha) de azimut).



#### Simulación

Como se explica en el método de cálculo de la HRTF, la señal filtrada se enviará por auriculares (de esta manera evitamos el procesado obligado de eliminación de rayos cruzados cuando se usan altavoces). Este hecho incluye la imposibilidad de la externalización del sonido (la señal será percibida alrededor de la cabeza sin sensación de profundidad), y la movilidad de la escena auditiva al mover la cabeza (hay sistemas que incorporan un corrector de tracking).

La señal usada para ser filtrada por la HRTF y, por lo tanto, el sonido que llevará la información tridimensional, muestra predilección por unas determinadas características. Después de realizar diferentes simulaciones con diferentes tipos de señales, los resultados más espectaculares se han obtenido usando ruido blanco que, además de proporcionar una ancha banda frecuencial, enmascara cualquier ruido de grabación o ruido de procesado introducido en el experimento. Señales de corta duración en el tiempo también han favorecido el efecto en las simulaciones.

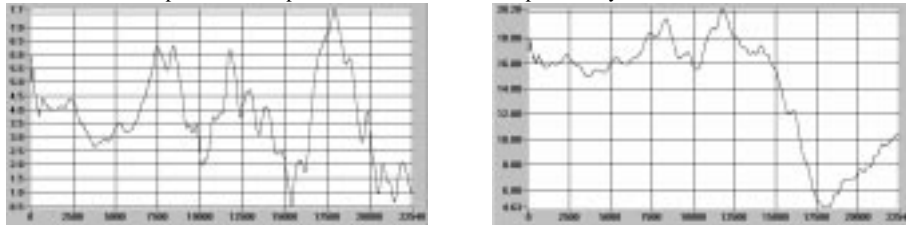
#### Valoración de la HRTF de un maniquí

Numerosas empresas del mundo de la acústica han invertido mucho tiempo y dinero en el desarrollo de

maniqués que simulen perfectamente la respuesta de la persona humana. Se los llama cabezas artificiales y gracias a estos modelos matemáticos se pueden hacer todo tipo de medidas binaurales en cualquier situación y condiciones reproduciendo la misma sensación que captaría una persona humana.

Después de implantar nuestro sistema basándonos en HRTFs humanas, hemos querido realizar lo mismo usando la función de transferencia estandarizada de un maniquí. Hemos intentado adecuar un maniquí vulgar y convertirlo en una cabeza artificial apta para cualquier tipo de grabaciones binaurales. Se ha moldeado el pabellón de la oreja haciéndolo más grande de lo que era originalmente, y dándole una forma más verosímil. Se ha recubierto toda la cabeza y cara con diversas capas de látex líquido para obtener una textura menos reflectante y más parecida a la humana, hemos recubierto el interior del busto y cabeza con absorbente para eliminar cualquier resonancia en el interior y, finalmente, se ha sellado cualquier apertura con el fin de evitar entrada de señal por sitios indeseados. Los micrófonos usados eran los mismos que en el primer experimento; totalmente aparejados en respuesta frecuencial y en fase.

A continuación mostramos las HRTFs obtenidas con el maniquí antes y después de la adecuación: Las HRTFs corresponden a una posición de 45° de azimut positivos y 40° de elevación.



En la ilustración de la izquierda se muestra la HRTF del maniquí antes de realizar las mejoras. Comparándola con la gráfica de la derecha (maniquí adecuado), se comprueban claramente los cambios obtenidos y la validez de la HRTF final, la cual nos ha dado los mismos resultados que las humanas. Por lo tanto, podemos afirmar que nuestro maniquí se acerca bastante a la realidad.

### Conclusiones

Las sensaciones provocadas en las personas sometidas a las diferentes simulaciones han sido muy satisfactorias. Queremos destacar la importancia de la experiencia del oyente en este tipo de escuchas y la importancia de la señal utilizada en la simulación. Las señales situadas en el plano medio eran localizadas con más dificultad que el resto, produciendo estas últimas efectos mucho más espectaculares. Presentar una simulación como resultado de diferentes filtrados a diferentes posiciones, dando así una sensación de movimiento brusco o continuado lleva a efectos muy reales y sorprendentes. Finalmente, sólo queremos remarcar la importancia del filtrado digital con la HRTF como base de la tecnología en tres dimensiones aplicada a las necesidades y aplicaciones de nuestro tiempo.

### Bibliografía

- Transaural 3-D audio. William G. Gardner. MIT Media Lab, Cambridge
- Transfer Characteristics of Headphones Measurements of Humans Ears. Henrik Möller, Michael Friis, Clemen Boje Dorte Hammershoi . JASA Vol 43 n°4
- A Computational Model of Spatial Hearing. Keith Dana Martin. Master of Science in Electrical Engineering at MIT
- Role of spectral cues in median plane localization. Futoshi Asano, Yoiti Suzuki Toshio Sone. JASA 88 July 90