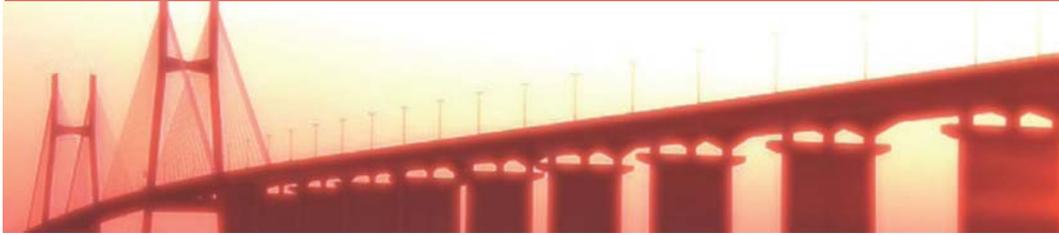


PRIMERAS JORNADAS REGIONALES DE ACÚSTICA AdAA 2009

19 y 20 de noviembre de 2009, Rosario, Argentina



AdAA2009-A004R

Empleo de técnicas de grabación para emulación de tridimensionalidad en una aplicación estereofónica

Mayra A. Cardozo Pérez^(a),
William F. Romo Martínez^(b).

(a) Ingeniería de Sonido, egresada Universidad San Buenaventura, AA 75010, Bogotá, Colombia. E-mail: mcardozo@academia.usbbog.edu.co

(b) Cátedra de Producción y Grabación, Facultad de Ingeniería, Universidad San Buenaventura, AA 75010, Bogotá, Colombia. E-mail: wromo@academia.usbbog.edu.co

Abstract

For a long time several investigations have been performed in order to explain how humans localize sound sources in a natural listening environment. Localization in the horizontal plane is mainly influenced by the time and intensity differences of the arriving sound at the ears, as the pinna's filtering information is important for vertical plane localization, when the location and distance of the sound source is changed. The stereophonic recording techniques provide considerable information about sound sources in the horizontal plane, supporting their localization. For that reason the emulation of the pinna's properties through artificial modeling is investigated, in order to obtain the synthetic pinna's impulse response, therefore a convolution with a stereophonically recorded signal can be done, resulting in a stereophonic signal with the pinna's characteristics. This provides a sound with a spatial hearing impression that approaches how is perceived in natural hearing and provides auditory sensation that depends on the position of the sound source.

Resumen

Por mucho tiempo varias investigaciones han sido dirigidas a la explicación de cómo los seres humanos ubican una fuente sonora en un ambiente de audición natural. La localización sonora en el plano horizontal es básicamente influenciada por las diferencias de tiempo y de intensidad con las cuales el sonido llega a los oídos; mientras que para ubicar sonidos en el plano vertical se utiliza la información dada por el filtrado de la pinna, cuando es variada la dirección y la distancia de la fuente sonora con respecto al receptor. Las técnicas de grabación estereofónicas proporcionan información de importancia para fuentes sonoras ubicadas en el plano horizontal, facilitando su localización. Por tal motivo se busca emular las propiedades del pabellón auditivo y mediante un modelado artificial de la misma, obtener su respuesta al impulso que permita por convolución, aplicar estas características a una señal estereofónica previamente capturada. Esto proporciona un sonido con una espacialidad muy cercana a aquellas de la escucha natural y causa ciertas sensaciones auditivas dependiendo de la posición de la fuente.

1 Introducción

Por mucho tiempo varias investigaciones han sido dirigidas a la explicación de cómo los seres humanos ubican una fuente sonora en un ambiente de audición natural. Nuestro sistema auditivo modifica el sonido mediante un proceso de filtrado que depende de su dirección (Proceso fisiológico y anatómico). Este sonido modificado, involucra una serie de propiedades espaciales las cuales son usadas por el cerebro para detectar su dirección, sin embargo, debido a la discrepancia en la anatomía del oído, la percepción de estas propiedades puede ser distinta para cada ser humano (Percepción subjetiva).

La forma en que el individuo percibe el sonido depende físicamente de sus oídos; esto es llamado audición binaural. La escucha binaural es la responsable de la detección del sonido, así como de su dirección. Desde 1900, los modelos binaurales han atraído la atención de muchos investigadores, particularmente cuando Lord Rayleigh postuló la teoría Dúplex, la cual se ha intensificado desde 1948 con Jeffress, quien propuso por primera vez el término binaural.

Los modelos binaurales han sido usados para crear indicios de localización de fuentes sonoras virtuales, que gracias a la ingeniería y a las ciencias que la soportan, se ha podido realizar aciertos y simulaciones muy reales lo que permite su aplicación en otras áreas. El sonido binaural produce la sensación de tridimensionalidad, el cual está directamente relacionado con la diferencia de intensidad y de tiempo recibida por cada oído de un sonido emitido por una fuente (Lord Rayleigh 1907); estas diferencias son de importancia para la ubicación del sonido en el plano horizontal. Otras características tales como el filtrado ocasionado por el pabellón auditivo “pinna”, contribuyen a la ubicación de fuentes sonoras en el plano vertical, donde los pliegues de la pinna producen reflexiones que generan picos y valles (Notch) en el rango de frecuencias de 4 khz-12 khz aproximadamente¹.

La revisión de estos antecedentes justifica el desarrollo de la presente investigación, cuyo propósito es crear sensaciones auditivas como aquellas de la escucha natural por medio de tres factores: el primero mediante la información en el plano vertical; el segundo, en el plano horizontal (Acimut) y el tercero que tiene que ver con los juicios de distancia. Para el primer factor, es de consideración el estudio y la emulación de las propiedades inherentes a la pinna. Para el segundo factor, plano horizontal, es de importancia el análisis e interpretación de las técnicas estereofónicas, ORTF y estéreo apantallado (Jecklin disk), técnicas escogidas premeditadamente para permitir aprovechar al máximo sus cualidades en el momento de la captura. La técnica ORTF (llamada así por ser ideada en la oficina de radio y televisión francesa) usa dos pequeños micrófonos cardioides, con una separación entre sus diafragmas de 0,17 m y un ángulo entre los ejes de sus cápsulas de 110°, es una de las técnicas de grabación estéreo más usada. La técnica Jecklin Disk o estéreo apantallado (Optimal Stereo Signal) fue creada por Josephson Electronics y se refiere a técnicas estereofónicas que utilizan una pantalla aislante para realzar la separación entre los canales de la señal estéreo, ésta técnica en particular consta de dos micrófonos omnidireccionales separados 0.165 m y tiene una pantalla de 0.30 m ubicada entre ellos; la pantalla es un disco recubierto con material absorbente cuyo fin es simular la cabeza humana².

Muchos investigadores han realizado estudios con objetivos similares a los presentados aquí, entre ellos están Frederic L. Wightman y Doris J. Kistler del Department of Psychology and Waisman Center de la Universidad de Wisconsin, Bloom en 1977³, Watkins en 1978⁴, Jens Blauert en 1983, Plenge en 1974, entre otros investigadores. Sin embargo, en Colombia no se han realizado estudios que aborden este tema, razón por la cual se hace necesaria esta investigación que procura aportar a la sociedad colombiana así como a la comunidad de ingenieros de sonido, pues el hecho de conocer las características de la pinna y su función en la audición humana permite mediante su emulación, aplicar

¹ Ramos, Oscar, et al. Modelo acústico de cabeza y torso mediante análisis de componentes principales. Córdoba, Argentina: Asociación Argentina de mecánica computacional, 2007, vol. XXVI, pp. 48.

² Bartlett, Bruce. Stereo microphone technique. Boston:Focal press, 1991. pp. 62-63

³ Bloom, P. Jeffrey Creating Source Elevation Illusions by Spectral Manipulation. J. Audio Eng. Soc. September 1977. Vol 25. pp. 560-565.

⁴ Watkins, Anthony J. Psychoacoustical aspects of synthesized vertical locale cues. J. Acoust. Soc. Am. Vol 63. pp. 1152-1165.

sus propiedades a sonidos previamente grabados, creando de esta forma sonidos más cercanos a los de la escucha natural y de esta manera lograr emplearlos no sólo en las áreas competentes a la ingeniería de sonido en sus distintas facetas (Producciones musicales, sonido para videos reproducidos a través de sistemas personales, “Ipods”, sonido para videojuegos), sino también en la medicina dirigida a personas con deficiencias visuales, entre otras aplicaciones.

2 Contextualización

Como ya se ha anticipado la investigación tiene como finalidad la aplicación de técnicas estereofónicas (ORTF y Jecklin Disk) para la creación de percepción de espacialidad sonora. Para esto se realiza una emulación del plano vertical mediante la extracción de las características de filtrado obtenidas de una pinna artificial (diseñada especialmente en esta investigación); tales características de filtrado es aplicada mediante un proceso de convolución a las capturas estereofónicas que se realizaron en el estudio. Para lograrlo fue necesario establecer una serie de mediciones, donde inicialmente se buscó determinar las diferencias existentes entre las capsulas de los micrófonos, tanto de tiempo como de intensidad y a través de ello conocer cómo estas técnicas estereofónicas capturan el campo sonoro en el plano horizontal (Acimut), con respecto a las que el ser humano realiza en la escucha cotidiana (*Diferencias interaurales de tiempo y de intensidad*). Debido a que no se cuenta en Colombia con una cámara anecoica, la medición intertransductor se realiza en la sala de grabación de un estudio, pero teniendo en cuenta ciertos parámetros que serán detallados en la metodología. Al respecto es importante saber que las *diferencias interaurales de tiempo* (ITD) como de *intensidad* (IID) (Las del oído), son realizadas bajo condiciones de campo libre o en cámaras anecoicas, en donde el sonido viaja en línea recta desde la fuente hasta el oyente. De esta manera se puede considerar que en nuestra escucha cotidiana las cosas son un poco diferentes puesto que se anexan condiciones de reflexiones, difracciones y otros factores ocasionados por el contacto del sonido con paredes, pisos, techos, entre otros. Para efectos de este proyecto, se considera la diferencia entre dos términos: La *localización* del sonido, referido a la manera en que el ser humano percibe el sonido y la laterización, entendida como fuentes sonoras virtuales.

Ahora bien, las *diferencias intertransductores* son importantes porque dependiendo del funcionamiento del arreglo estereofónico, se puede conocer cómo capturan éstos el campo sonoro en el plano horizontal; ya que no funcionan para el plano vertical. Entre cuanto más parecidas sean estas técnicas con la percepción humana, más realismo se puede tener al momento de la reproducción y menores errores de localización en la misma.

Por otro lado es importante acotar que el plano vertical es primordial para tener una percepción de espacialidad. De acuerdo a estudios realizados por muchos investigadores, la localización en este plano se ve disminuida ampliamente cuando la pinna es removida y el estímulo no tiene contenido espectral en altas frecuencias. De igual manera es relevante destacar que cuando se presentan estímulos sonoros de banda ancha la localización en el plano vertical mejora.

El rango de frecuencias en los cuales la pinna produce retrasos y modificaciones se encuentra entre 4 kHz y 12 kHz, pero esto varía de persona a persona, pues los seres humanos desde el punto de vista fisiológico no tienen el pabellón auditivo estructuralmente igual, razón por la cual las funciones de transferencia relativas a la cabeza (HRTF) en muchas ocasiones no proporcionan informaciones coherentes, sino a aquellas personas que tengan sus aparatos auditivos parecidos a los utilizados en la grabación, bien sean artificiales o reales.

3 Metodología

Para obtener los resultados requeridos para la creación de espacialidad, se establece el siguiente orden:

1. Grabación con técnicas estereofónicas, ORTF y Jecklin Disk para determinar qué diferencias de tiempo y de nivel *intertransductores* son producidas en estas configuraciones, a partir de la emisión de distintas señales acústicas.

2. Medición de respuesta al impulso de una pinna artificial diseñada previamente, de manera que se obtengan datos relevantes para la localización en el plano vertical, sin descartar aquella información que pueda ser proporcionada en otras posiciones (Por ejemplo: acimut).

3. Grabación de instrumentos, voces y sonidos cotidianos para el ser humano mediante el uso de las técnicas estereofónicas ORTF y Jecklin Disk y posteriormente realizar el procesamiento de estas señales grabadas mediante convolución.

4. Grabación binaural en múltiples posiciones de fuente con pequeños micrófonos ubicados en el canal auditivo de un oyente real. Esta grabación será comparada con aquellas que han sido procesadas con el filtrado de la pinna artificial (Grabación con las técnicas estéreo ORTF y Jecklin Disk).

5. Comprobación de resultados mediante pruebas con personas reales.

A continuación se detalla el procedimiento de cada punto.

3.1 Diferencias intertransductores de intensidad y tiempo de las técnicas ORTF y Jecklin Disk, mediante la variación de ángulos acimut

Debido a que la Universidad San Buenaventura no cuenta con una cámara anecoica para realizar este tipo de medidas, se prosigue a realizar la medición en el laboratorio de sonido (Estudio híbrido) de la misma, pero para poderla llevar a cabo es necesario obtener la distancia crítica del recinto, con el propósito de realizar la captura en campo directo y así lograr datos con mayor precisión.

Ahora bien, para obtener la distancia crítica, es necesario conocer el tiempo de reverberación del recinto que se consigue por medio de la respuesta al impulso. A continuación se expone el procedimiento para adquirir estos datos.

3.1.1 Medición del tiempo de reverberación y distancia crítica, estudio híbrido Universidad San Buenaventura, Bogotá

La distancia crítica es obtenida a través de la medición del tiempo de reverberación, teniendo en cuenta la norma ISO 3382⁵, a través de la siguiente ecuación:

$$DC = 0.141\sqrt{QR} \quad (1)$$

donde:

Q = factor de directividad de la fuente. (1 para el dodecaedro de patrón polar omnidireccional)

R =constante del recinto

$$R = \frac{S\bar{\alpha}}{1 - \bar{\alpha}} \quad (2)$$

S =superficie de total del recinto.

$\bar{\alpha}$ =absorción promedio de la sala.

3.1.2 Cálculo de las Diferencias intertransductores de intensidad y tiempo

La medición se realiza en el estudio híbrido de la Universidad San Buenaventura sede Bogotá, en los siguientes ángulos acimutales: 0°, 30°, 45°, 60°, 90°, 120°, 135°, 180° (Sentido contrario a las manecillas del reloj y asumiendo simetría entre el lado izquierdo y derecho) y se emiten tres tonos puros: 500Hz, 1 kHz y 4 kHz; así se tiene una referencia de frecuencia media baja, media y alta. La distancia receptor-fuente se toma de acuerdo con el valor de DC calculado.

3.2 Medición de respuesta al impulso de una pinna artificial

En razón a que las técnicas no proporcionan información espacial en la reproducción por medio de audífonos, se realiza la emulación del plano vertical por medio de las características de filtrado de la pinna. Inicialmente se realiza el diseño de la pinna artificial y de un soporte para ésta, cuya finalidad

⁵ ISO 3382 Acoustics, measurement of the reverberation time of rooms with reference to other acoustical parameters. Segunda edición. Switzerland, 1997.

es permitir variaciones en ángulos acimut y de elevación de la fuente sonora. La pinna está hecha de silicona para moldeo y el soporte de metal y madera. La figura 1 muestra una fotografía de esto.

La silicona para moldeo es escogida por su estabilidad a la temperatura, lo cual no sucede con el látex, otro material comúnmente usado para estos fines. Además la silicona es usada para realizar moldes por su similitud con la textura de la piel. El molde para la elaboración de la pinna es tomado de una persona real y el pabellón auditivo diseñado es el izquierdo, por lo tanto la medición se efectúa sólo para ese lado asumiendo simetría para el lado derecho (En la realidad los seres humanos no son simétricos).



Figura 1. Fotografías de la pinna y su soporte; diseñadas para la investigación.

Una vez diseñada la pinna y su soporte, se prosigue a realizar la medición de la respuesta al impulso de la misma, medición que se realiza en dos lugares: al aire libre para “simular” las condiciones ideales de las cámaras anecoicas y en campo reverberante (Estudio híbrido de la Universidad San Buenaventura, sede Bogotá). El diagrama de conexión que contiene los instrumentos para la medición de la respuesta al impulso de la pinna artificial es el presentado a continuación (Figura 2):

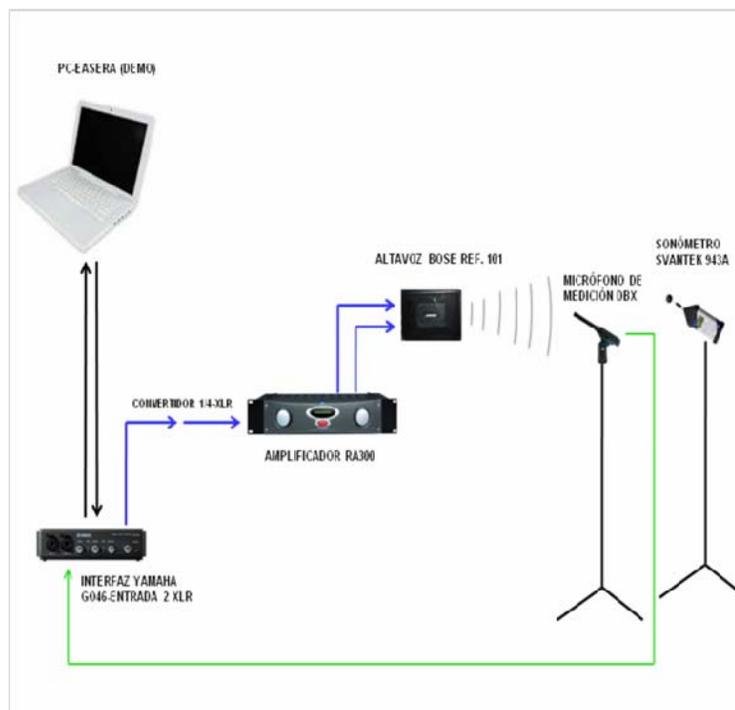


Figura 2. Diagrama de conexión de la medición de respuesta al impulso, pinna artificial.

Se enumeran los pasos seguidos:

1. Calibración eléctrica del micrófono de medición DBX® por medio del Pistófono RION Modelo NC-74 para asegurar que el nivel de presión sonora que reciba el transductor sea el mismo que indique el software de medición en su escala de SPL. El software utilizado es EASERA© (Versión demo).

2. Emitiendo un ruido rosa y teniendo en cuenta las sugerencias de la ISO 3382, se verifica con el sonómetro (Svantek® tipo2, referencia 943 A) que la señal emitida por el parlante esté 35 dB por encima del ruido de fondo. El sonómetro ha sido previamente calibrado con un filtro de ponderación A, detector *slow* y buffer *rms*.

3. Como señal de emisión es utilizado un barrido senoidal con una duración de 2.7 s. El barrido senoidal, “sweep”, es emitido por el parlante y con base en la calibración planteada se asegura que el ruido de fondo esté 35 dB por encima del ruido de fondo, norma ISO 3382.

Las posiciones de medición son las siguientes:

Acimut: 0°, 45°, 90°, 135°, 180°.

A cada posición acimutal le corresponde sus ángulos de elevación de la siguiente forma:

Acimut 0 grados, Elevaciones: 0°, 30°, 60°, 90°, 120°, 150°, 180°.

Acimut 45 grados, Elevaciones: 0°, 30°, 60°, 90°.

Acimut 90 grados, Elevaciones: 0°, 30°, 60°, 90°.

Acimut 135 grados, Elevaciones: 0°, 30°, 60°, 90°. La figura 3 lo muestra.

4. Con el micrófono DBX® ubicado a la entrada del meato auditivo de la pinna artificial, se captura el barrido emitido por el altavoz Bose®, el cual recibe todas las características de la pinna que implica el cambio de angulación de la fuente.

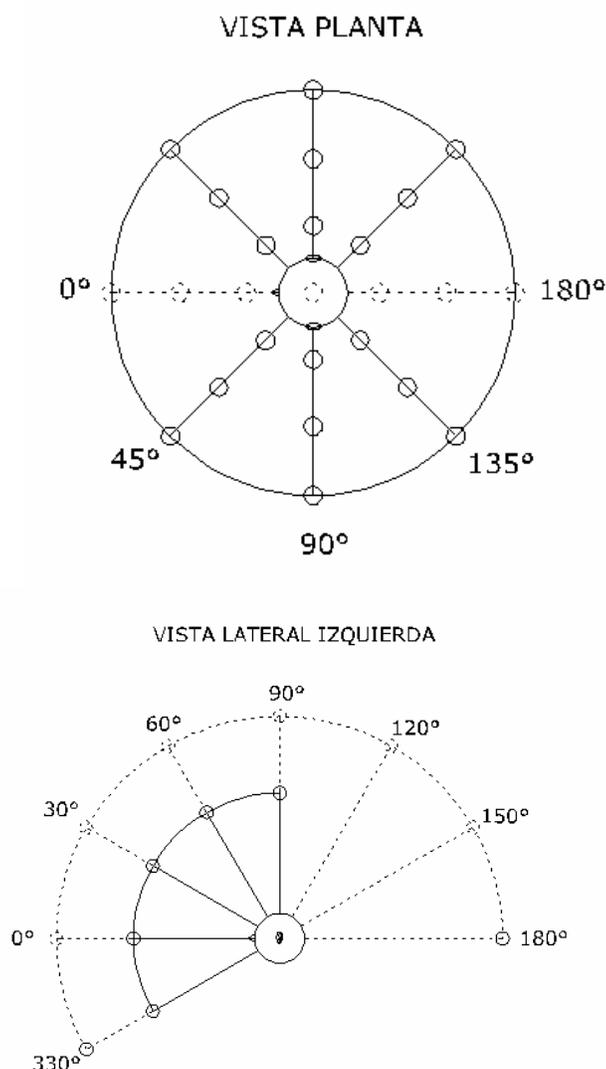


Figura 3. Posicionamiento de los ángulos de medición para RI (respuesta al impulso) de la pinna artificial. (a) vista planta, (b) vista lateral izquierda.

3.3 Grabación con técnicas estereofónicas ORTF, Jecklin Disk, Binaural y aplicación de algoritmo a estas capturas

Es importante establecer que para cumplir con los objetivos de la investigación se realiza un proceso de convolución, a continuación se explica:

Primero se realizan distintas capturas:

1. Grabación con las dos técnicas estereofónicas elegidas (ORTF y Jecklin Disk) en las mismas posiciones acimutales utilizadas para la medición de la pinna artificial. Los instrumentos fueron: shaker, voz y guitarra acústica. Estos instrumentos fueron escogidos premeditadamente ya que la percepción de elevación es mejorada cuando contiene frecuencias altas dentro del estímulo.

2. Grabación binaural para compararla con las estereofónicas. Los instrumentos fueron: shaker, voz y guitarra acústica

Como se ha mencionado las técnicas estereofónicas por si solas no proporcionan espacialidad por medio de audífonos, por tal razón se toma la respuesta al impulso de la pinna artificial y se realiza la convolución con la señal estereofónicamente grabada en las mismas posiciones, por ejemplo, para la guitarra grabada con la técnica ORTF a 90 grados, se aplica la respuesta al impulso de la pinna artificial en esa misma posición, pero con sus distintas elevaciones; por lo tanto, se obtendrá una guitarra filtrada, la cual produce ciertas sensaciones a diferentes alturas, tomando como referencia la misma grabada a 90° horizontal.

La convolución se realiza por medio de un algoritmo en un software, Matlab™ versión R2007a. Las sentencias aplicadas en este algoritmo para realizar tal proceso fueron las siguientes:

```
>> [y,fs]= wavread ('Arreglo ORTF guitarra 90 grados az 1m.wav'); % archivo de audio
>> left=y(:,1); % Canal izquierdo
>> right=y(:,2); % Canal derecho
>> [x,fs]= wavread ('AZ90EL30.wav');% Respuesta al impulso azimuth 90 elevación 30
>> left=y(:,1); % Canal izquierdo
>> right=y(:,2); % Canal derecho
>> convo(:,1)=conv(left,x); % Convolución
>> convo(:,2)=conv(right,x); % Convolución
>> wavwrite(convo,fs,24,'convolución guitarra AZ90EL30 1m');
```

3.4 Metodología para evaluación (Comprobación de resultados)

Finalmente para verificar qué tan adecuadas son las técnicas estereofónicas en proporcionar información para la correcta localización en el plano horizontal (Entendida como lateralización porque es un sonido virtual), así como también determinar si se percibe un cambio de elevación se realiza una comprobación de resultados (Resultados mostrados en el apartado 4.4). Un grupo pequeño de personas es requerido para certificar la coherencia correspondiente a los resultados obtenidos. La evaluación de 15 individuos se considera adecuada y suficiente para esta comprobación, ya que cada prueba requiere de un periodo de 50 minutos para un total de 12.5 horas, la población es básicamente personas con cierto grado de entrenamiento auditivo, por ejemplo: estudiantes de ingeniería de sonido, profesores y auxiliares de laboratorio de sonido. Ésta comprobación tiene dos divisiones:

1. Comprobación de resultados mediante la emisión de sonidos aleatoriamente presentados a los oyentes para que ellos determinen la posición del instrumento (Shaker grabado con arreglos estéreo y su respectiva convolución). Las posiciones utilizadas para esta comprobación son: 0° acimut con elevaciones: 0°, 30°, 60°, 90° y 90° acimut con elevaciones: 0°, 30°, 60°, 90°.

2. Comparación de las grabaciones estereofónicamente capturadas y su respectivo proceso de filtrado (Convolución) con una grabación binaural realizada mediante micrófonos miniatura (SP TFB-2) ubicados en los oídos de una persona real.

Obviamente para establecer una comparación se utilizan las mismas posiciones tanto horizontales como verticales, así como dos tipos de audífonos: In-Ear y Sony MDR V-6, por cuanto el tipo de transductor también influye en la percepción y por lo tanto en la ubicación de fuentes sonoras.

La presentación de las muestras se realiza de acuerdo a lo establecido para este tipo de valoraciones subjetivas⁶. De acuerdo con esto, se tiene en cuenta una variable muy importante: presentar los datos aleatoriamente pero no tantas veces de forma que el individuo no sienta expectativa con respecto a la aparición de la señal, dado que esto puede causar errores en las respuestas. La cantidad de muestras presentadas y su orden son mostrados enseguida en la tabla 1

⁶ Schiffman, Harvey Richard. Psicofísica. En: La percepción sensorial. Segunda edición. Mexico:limusa,1997. pp. 35-44

Tabla 1. Orden de reproducción de las muestras

AZIMUTH		
Técnica de microfonería estéreo Jecklin disk	Técnica de microfonería estéreo ORTF	Técnica Binaural
Sonido 1 - 0 grados	Sonido 1 - 0 grados	Sonido 1 - 0 grados
Sonido 2 - 90 grados	Sonido 2 - 45 grados	Sonido 2 - 45 grados
Sonido 3 - 90grados	Sonido 3 - 45 grados	Sonido 3 - 135 grados
Sonido 4 - 45 grados	Sonido 4 - 135 grados	Sonido 4 - 180 grados
Sonido 5 - 0 grados	Sonido 5 - 0 grados	Sonido 5 - 135 grados
Sonido 6 - 45 grados	Sonido 6 - 180 grados	Sonido 6 - 180grados
Sonido 7 - 180 grados	Sonido 7 - 90 grados	Sonido 7 - 90grados
Sonido 8 - 135 grados	Sonido 8 - 90 grados	Sonido 8 - 45 grados
Sonido 9 - 135 grados	Sonido 9 - 135 grados	Sonido 9 - 90 grados
Sonido10 - 180 grados	Sonido10 - 180 grados	Sonido10 - 0 grados
ELEVACIÓN		
Técnica de microfonería estéreo Jecklin disk	Técnica de microfonería estéreo ORTF	Técnica Binaural
Sonido 1 AZ0EL0	Sonido 1 AZ0EL0	Sonido 1 AZ0EL0
Sonido 2 EZ0EL60	Sonido 2 AZ0EL60	Sonido 2 AZ0EL60
Sonido 3 AZ0EL90	Sonido 3 AZOEL90	Sonido 3 AZOEL90
Sonido 4 AZ0EL30	Sonido 4 AZ0EL30	Sonido 4 AZ0EL30
Sonido 1 AZ90EL0	Sonido 1 AZ90EL30	Sonido 1 AZ90EL0
Sonido 2 AZ90EL30	Sonido 2 AZ90EL0	Sonido 2 AZ90EL90
Sonido 3 AZ90EL90	Sonido 3 AZ90EL90	Sonido 3 AZ90EL30
Sonido 4 AZ90EL60	Sonido 4 AZ90EL60	Sonido 4 AZ90EL60

4 Resultados obtenidos

4.1 Resultados de distancia crítica y diferencias intertransductores

En la tabla 2 se muestra las características de la sala (Estudio híbrido) y su tiempo de reverberación, cuyos datos fueron usados para finalmente por ecuación obtener la distancia crítica de la sala

Tabla 2. Dimensiones de la sala y tiempo de reverberación medido.

Ancho	5 m
Largo	6,33 m
Altura	3,60 m
Superficie total	144,876 m ²
Volumen	113,94 m ³
Tiempo de reverberación, RTmid	0.52 s

Al tener en cuenta el procedimiento planteado para obtener el tiempo de reverberación y distancia crítica (Mostrados en 3.1.1) se obtienen el siguiente resultado:

Distancia crítica del estudio Híbrido = 0,977 m \approx 1 m. Por lo tanto las grabaciones y/o mediciones que requieran ser realizadas en campo directo deben ser efectuadas a una distancia menor de la distancia crítica.

De acuerdo con el valor de distancia crítica anteriormente mostrado, se prosigue a indicar los resultados de las diferencias entre capsulas para las técnicas de captura estereofónicas ORTF y Jecklin Disk. Los resultados de las diferencias intertransductores son mostrados en la figura 4.

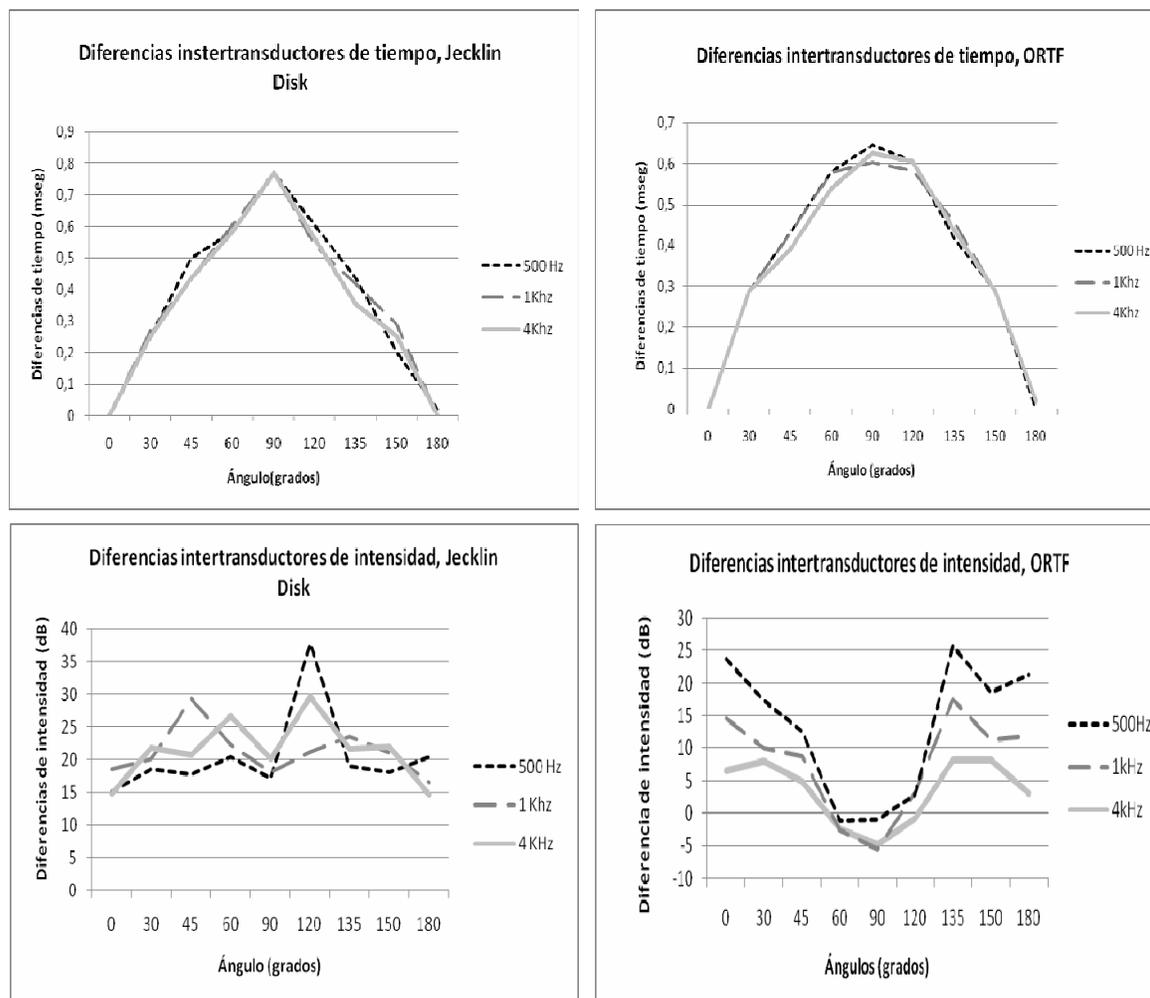


Figura 4. Diferencia intertransductores de tiempo y de intensidad ORTF y Jecklin Disk.

Interpretación de la ilustración:

Tomando como punto de comparación las *diferencias interaurales de tiempo y de nivel* medidas en la literatura⁷, se tiene que para las *diferencias intertransductores de intensidad* hay una variación aleatoria de datos, por lo tanto, no se puede establecer datos certeros. La teoría de las *diferencias interaurales de nivel* establece claramente que funciona para frecuencias altas por encima

⁷ Grantham, Wesley. Spatial hearing and related phenomena En: Handbook of Perception and Cognition, Second Edition. Second edition London, UK: edited by B. C. J. Moore. Academic, 1995, pp. 312.

de 2 kHz y a medida que se varía de angulación con respecto al axis (0 grados) se tiene un incremento, logrando su más alto nivel en 90 grados acimut. Ahora bien, de acuerdo a la teoría del arreglo de Jecklin Disk, éste se convierte en direccional cuando la frecuencia se incrementa, muy similar a la teoría *interaural*.

Sin embargo, en la banda de 4 khz de la figura 4 no se logra del todo ese valor, pues presenta una caída en 90 grados, lo cual no es coherente con la teoría acabada de citar.

Para ORTF no se tienen datos coherentes ni con la teoría *interaural* ni con la de micrófonos, esto es atribuido principalmente a las condiciones altamente reverberantes del recinto.

Para las diferencias de tiempo se tienen datos muy cercanos a los *interaurales*, pero no del todo correctos, ya que la respuesta también depende del patrón polar del arreglo. Si las mediciones se hubieran realizado en condiciones ideales (Cámara anecóica), los resultados posiblemente estuvieran más cercanos a los *interaurales*.

A pesar de no obtener datos totalmente exactos a los de la escucha natural, todas las grabaciones se realizan con éstas técnicas, asumiendo que proporcionan información relevante para la localización precisa en el plano horizontal (Acimut) teniendo en cuenta estudios realizados de éstas técnicas⁸.

4.2 Resultados respuesta al impulso de la pinna artificial

En este ítem se obtienen dos resultados de importancia de la medición (Basado en el procedimiento 3.2):

1. Respuesta en frecuencia de la pinna dependiente de la ubicación de la fuente sonora (Dominio de la frecuencia).

2. Respuesta al impulso que tendrá como fin filtrar una señal estéreo grabada, mediante una convolución (Dominio del tiempo).

Con respecto al primer punto, se expone a continuación solo dos figuras (La figura 5 y 6) de todos los resultados obtenidos que tiene el propósito de mostrar la manera en que funciona la pinna de acuerdo a ángulos específicos de fuente sonora; la ilustración es tomada del programa EASERA (Versión demo).

Se toma como ángulo de referencia para su análisis 90 grados acimut en sus dos lugares de medición; campo libre y en recinto.

Interpretación de la ilustración:

La pinna para acimut 90° elevación 0°, presenta las características mostradas por la curva B, la que muestra una caída (notch) alrededor de 8 kHz, un incremento a partir de 10 kHz y una caída otra vez en 11,47 kHz y 16 kHz aproximadamente. Con respecto a las elevaciones se tiene la siguiente interpretación: decrece el nivel alrededor de 5 kHz con respecto la curva AZ90EL0 para todas las curvas de elevación y hay un notch común para todas las elevaciones en 11,47 kHz, pero con distinto nivel a medida que el ángulo de elevación aumenta. Para la elevación de 30° se encuentra una caída en 9 kHz.

De esta figura se puede deducir que a medida que aumenta el posicionamiento vertical, aumenta la frecuencia central del filtrado tipo notch (entre 7 kHz y 10 kHz). La curva A muestra el comportamiento si la pinna no existiera, lo que deja notar que sí existe una variación cuando la pinna es incluida. Se piensa en esta curva como ideal al no detenerse en los errores sistemáticos ocasionados por las reflexiones que se pudieron ocasionar con el soporte mismo, ruido de fondo, ruido de animales, las condiciones atmosféricas (factores de viento, humedad relativa), entre otros.

Se puede inducir que cada ángulo específico de elevación tiene una estructura de filtrado diferente, por lo tanto, en la frecuencia común de 11,47 khz para percibir un desplazamiento de 30 grados es necesario tener ese nivel, el cual es distinto de los demás ángulos.

⁸ Ceoen, Carl. Comparative Stereophonic Listening Test. J. Acoust. Soc. Am. January/February 1972. Vol 20. pp. 19-27.

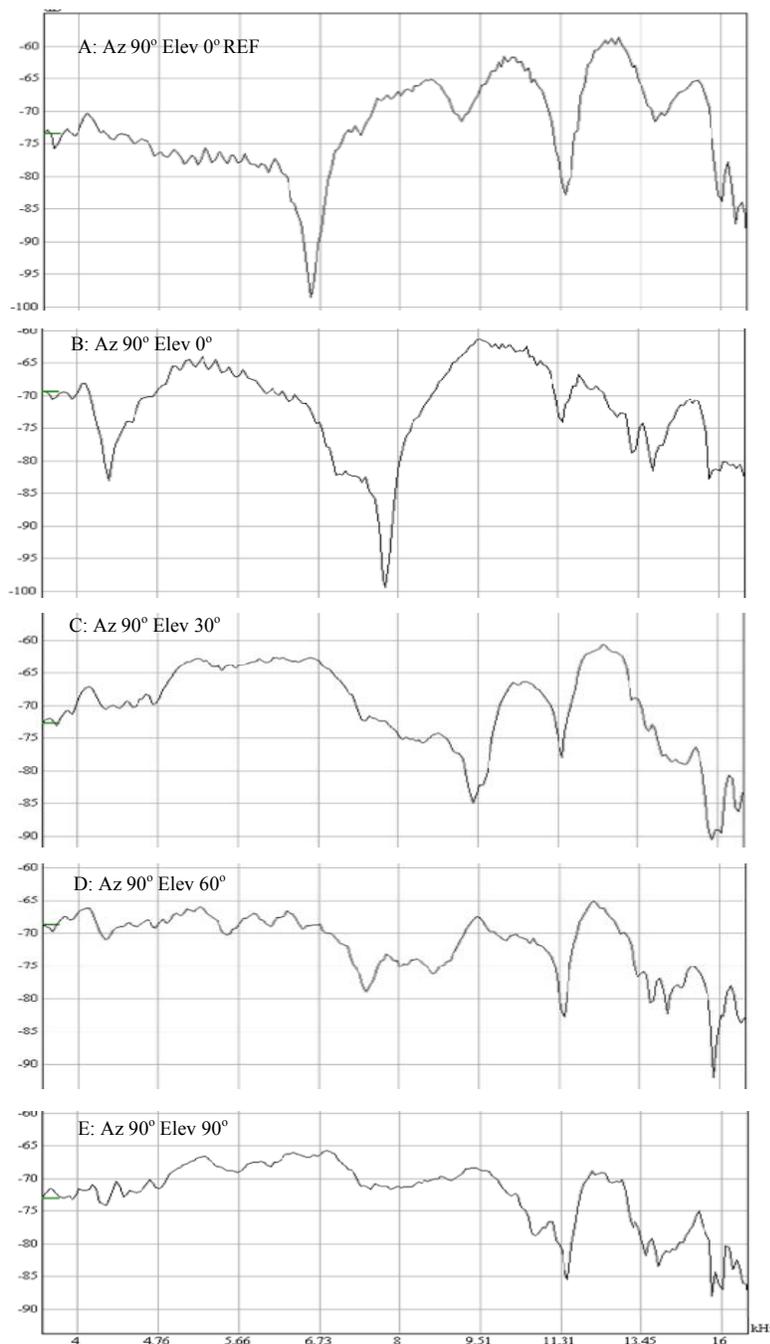


Figura 5. Respuesta en frecuencia posición acimutal 90° en elevaciones 0° (B), 30° (C), 60° (D), 90° (E) de la medición realizada en campo libre. AZ=acimut, EL=elevación. Figuras proporcionados por el programa EASERA© (versión demo).

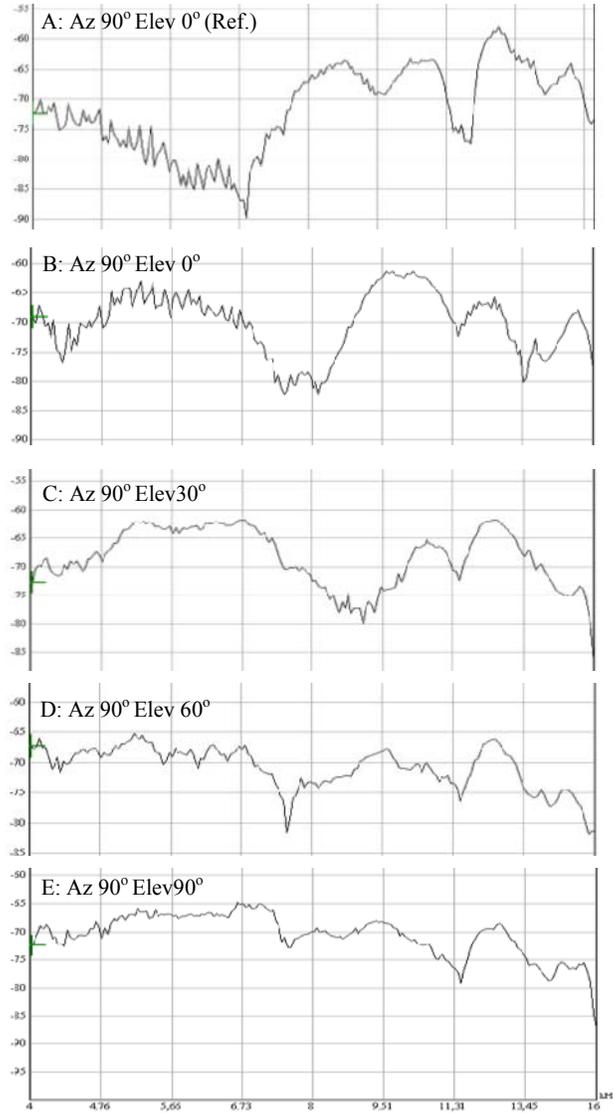


Figura 6. Respuesta en frecuencia de la posición acimutal 90° en elevaciones 0°(B), 30°(C), 60°(D), 90°(E) de la medición realizada en la sala de grabación. AZ=acimut, EL=elevación. Figuras proporcionados por el programa EASERA© (versión demo).

Interpretación de la ilustración:

De la figura 6 se pueden extraer algunos datos de importancia: en la elevación de 60 grados se presenta una caída en la frecuencia 7,49 kHz, pero no existe una comparación relevante con respecto a las otras curvas, por lo tanto, ésta puede ser ocasionada por reflexiones del lugar o del soporte. Al igual que en el gráfico presentado para aire libre surge una caída en la frecuencia de 11,47 kHz y 16 kHz para todas las elevaciones, obteniendo el mismo análisis planteado en esa figura. A pesar que la pinna esta soportada por una placa cubierta con el mismo material de las orejas artificiales (Mirar diseño en la figura 1), no está exenta de reflexiones ocasionadas por la base de madera y por el soporte metálico; sin excluir aquellas dadas por el recinto (las características de la sala fueron mostradas en 4.1). Otros errores sistemáticos como la manipulación del sistema, precisión en la angulación, ruido de fondo durante la mediciones (Ejemplo movimiento de sillas en el piso superior), también son tenidos en cuenta.

Pero en esta medición específicamente, se busca establecer el comportamiento en condiciones naturales de escucha humana, por tal razón estos errores sistemáticos ayudan a crear un ambiente más natural y pueden ser de cierta forma útiles.

4.3 Filtrado por convolución a las grabaciones estéreo

Así como en las dos figuras anteriores se toma como muestra la respuesta en frecuencia del instrumento grabado con la técnica ORTF a un ángulo de 90 grados y la captura realizada mediante la respuesta al impulso de la pinna artificial al mismo ángulo acimut de la técnica estéreo pero con una elevación de 30°, para así observar como es el funcionamiento del filtrado, de acuerdo con el procedimiento explicado en 3.3.

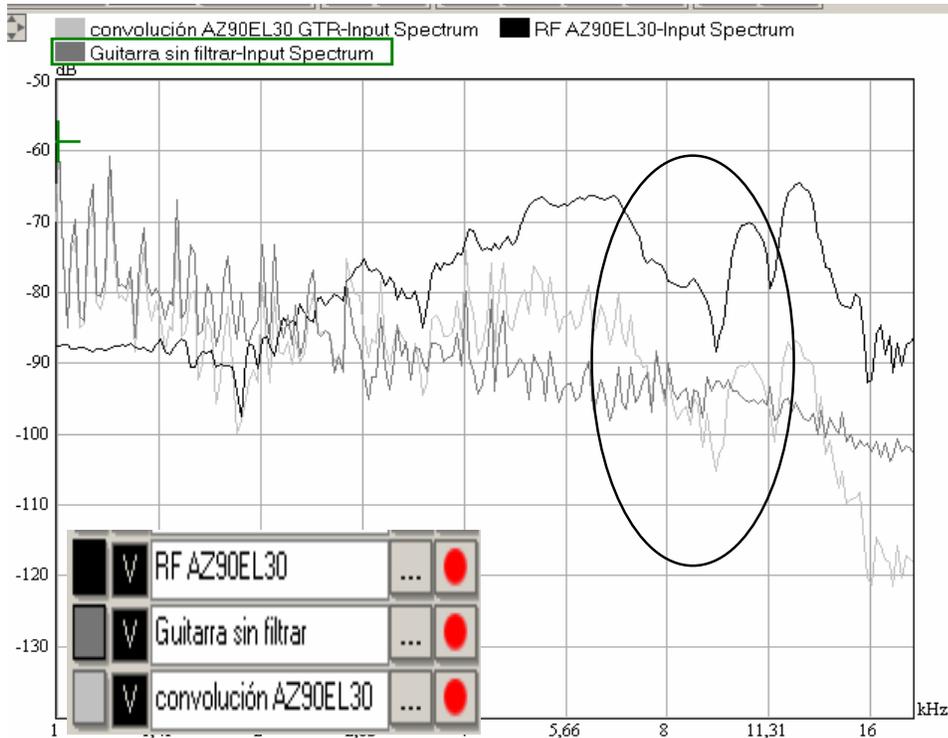


Figura 7. Respuesta en frecuencia de guitarra: sin filtrar (Gris oscuro), respuesta en frecuencia acimut 90° elevación 30° (Negro) y convolución (Gris claro). (Instrumento con Técnica de microfonería estéreo ORTF y respuesta al impulso de la pinna al aire libre).

En la figura 7 se distinguen las tres curvas, la guitarra tal y como se captura (Color gris oscuro), la respuesta en frecuencia de acimut 90 con 30 de elevación (Color negro), y el resultado de la convolución (Color gris claro), que como ya se ha mencionado fue hecha con la respuesta al impulso obtenida. Claramente se observa que la caída correspondiente a 9 kHz, 12 kHz y 16 kHz aproximadamente (Marcado en círculo), es aplicada a la señal (Guitarra) cuando es hecha la convolución; notando de esta forma que el filtrado se está realizando con éxito.

4.4 Resultados de la comprobación de resultados

Los resultados se fundamentan en el procedimiento explicado en el apartado 3.4. En términos porcentuales los resultados de la comprobación se muestran seguidamente (Figuras 8, 9, 10). Estos porcentajes indican los aciertos que las personas tuvieron a un mismo ángulo; Cada ángulo equivale a un 20%, el cual indica que de las 15 personas todas acertaron en ese ángulo específico.

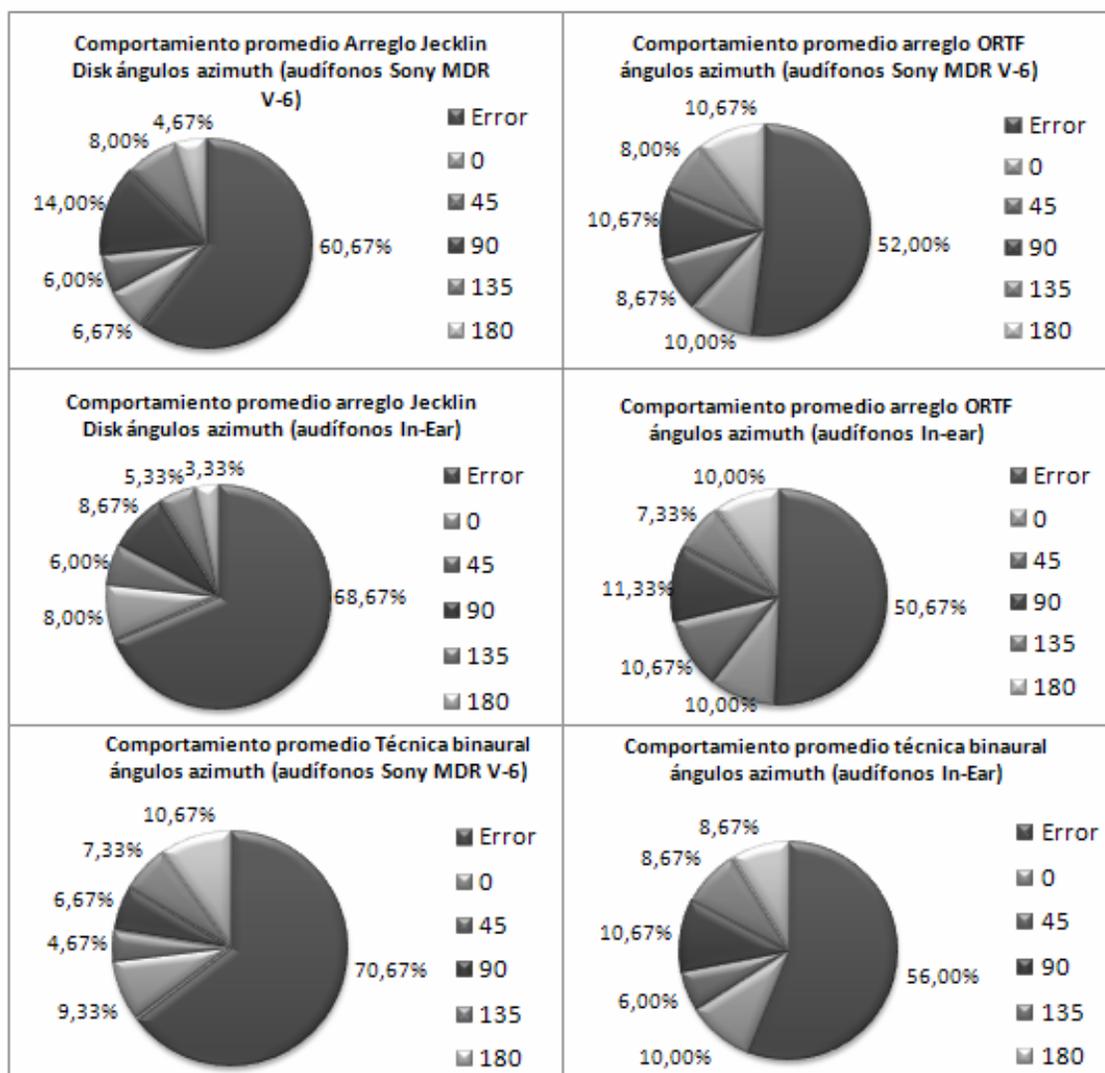


Figura 8. Indicación de porcentajes para posición horizontal. Audífonos Sony MDR V-6, In-Ear.

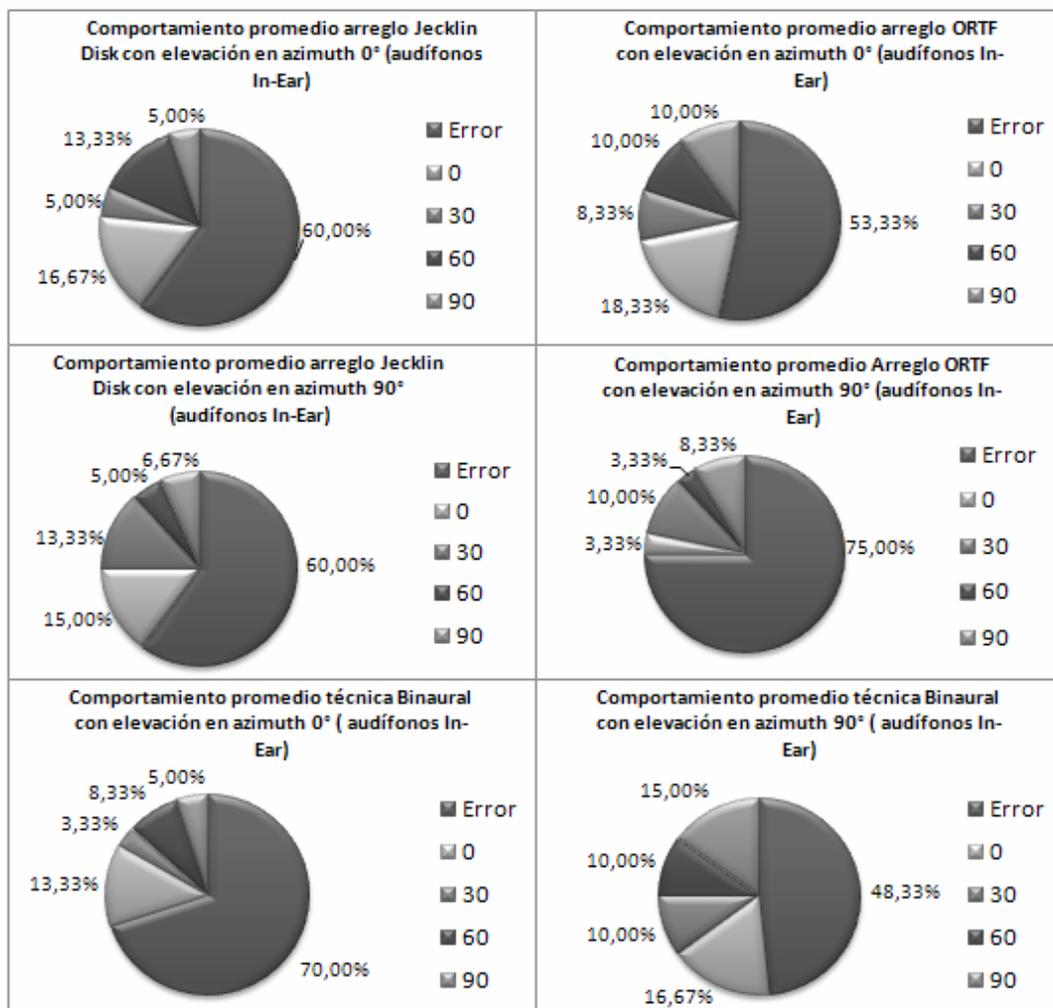


Figura 9. Indicación de porcentajes para elevación, audífonos In-Ear.

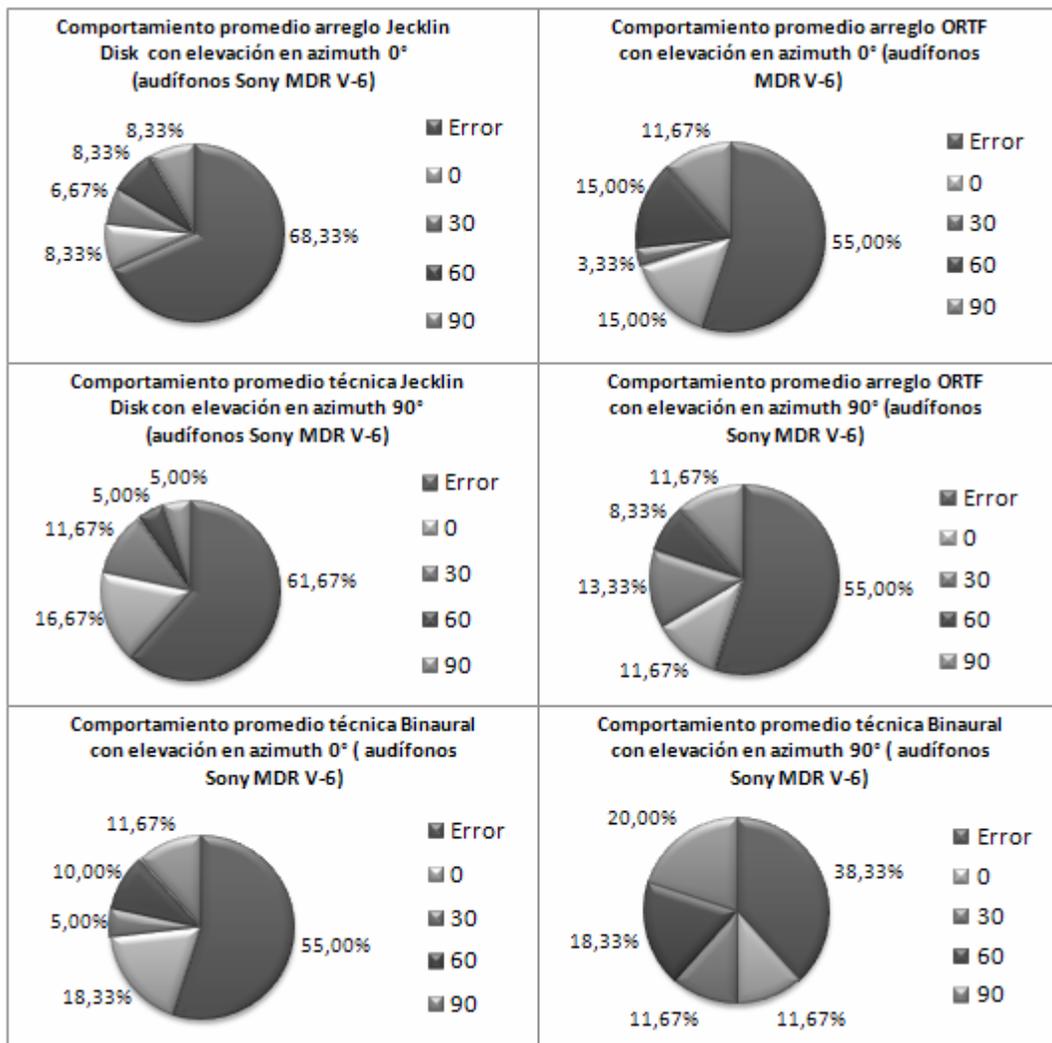


Figura 10. Indicación de porcentajes para elevación, audífonos Sony MDR V-6.

Interpretación de la ilustración:

Para los dos tipos de audífonos, se tuvo respuesta favorable con la técnica ORTF tanto horizontal como vertical. Sin embargo, tomando como límite 100 % los resultados no son del todo buenos. Esto es debido principalmente a las condiciones de medición, la respuesta al impulso de la pinna, las reflexiones con el soporte de lo misma que es de metal y a las condiciones del lugar de medición (Por no ser en una cámara anecoica). Debido a la cantidad de sonidos presentados (54 en total) se tuvo problemas de fatiga auditiva, por lo tanto los errores también están adjudicados a este problema.

4.5 Análisis y discusión de resultados

1. En la figura de respuesta en frecuencia de la pinna artificial al aire libre (Figura 5), se nota que hay caídas de importancia entre 12 kHz y 16 kHz. También se puede deducir que a medida que aumenta el posicionamiento vertical aumenta la frecuencia central del filtrado tipo notch (Entre 7 kHz y 10 kHz). Estas caídas también son notadas en la convolución de la figura 7.

2. En las técnicas estereofónicas cuando la convolución es aplicada, presenta resultados favorables tanto para la ubicación de fuentes sonoras en el plano horizontal como en vertical.

3. A pesar de la dificultad presentada para indicar la verdadera ubicación sonora, lo importante es que las personas lograron percibir un desplazamiento en el plano vertical. Las correcciones en las mediciones, como por ejemplo las realizadas en condiciones ideales (Cámara anecoica), lograrían mejorar esta percepción, así como también permitirían cuantificar sus datos.

4. Para la técnica de grabación binaural los resultados están alejados de lo esperado, debido a que se presenta un error del 70,66 % y 56 % para audífonos Sony MDR V-6 e In-Ear respectivamente. Sin embargo presenta buenas respuestas para acimut 90 en sus distintas elevaciones, con un porcentaje de 61,67 % para Sony y 51,67 % para In-Ear.

Al Comparar esta técnica, binaural, con aquellas aplicadas en la presente investigación (ORTF y Jecklin Disk), se puede concluir que estas últimas funcionan mejor para el plano horizontal, mientras que para elevación prevalece la técnica binaural.

Se tiene una explicación coherente para los resultados obtenidos, la cual tiene que ver con que la técnica binaural fue la última técnica presentada a los individuos, por lo tanto, ya tenían cierto grado de cansancio auditivo en la presentación de las señales binaurales. Por otro lado, en los laboratorios de sonido (Estudio híbrido) generalmente hay flujo constante de estudiantes debido a esto la concentración era dispersa en las personas a las cuales se les presentaron los sonidos.

4.6 Conclusiones

1. Mediante la obtención de los atributos de la pinna artificial, se pudieron realizar filtrados a una señal estereofónicamente grabada, con distintas técnicas de captura. Se logra obtener en cierta medida la percepción de un desplazamiento en el eje vertical (altura) de la señal filtrada.

2. Efectuar mediciones en cámaras anecoicas permite obtener la información correcta de la fuente sonora, su ubicación y sus características espectrales. Es posible que al realizar mediciones en tal recinto, la sensación lograda hubiera podido mejorar en el plano vertical principalmente y los errores de localización de la fuente sonora serían mínimos.

3. La variación de datos es adjudicada principalmente a las condiciones de medición, factores climáticos, ruido de fondo, humedad relativa, reflexiones con el soporte de la pinna, errores mecánicos en la angulación de la fuente (altavoz) y errores mecánicos del soporte que permite la angulación de la pinna artificial.

4. Se consigue comprobar lo que múltiples investigadores han manifestado, con respecto al filtrado de la pinna, en relación a que cuando la fuente sonora varía de ubicación presenta caídas (notch), principalmente en el rango de frecuencias entre 4 kHz-16 khz.

5. A pesar de la proporción derivada de las pruebas con dos tipos de transductores, el modelo Sony MDR-V6 ofrece más confort y percepción de profundidad, según lo expresado por las personas, punto clave para la impresión de tridimensionalidad.

6. En los resultados de la prueba se presentan errores de localización frontal-trasero, que deja entrever que la pinna ayuda a mejorarlo; pero otros factores asociados con el movimiento de la cabeza, como reflexiones del torso y la cabeza, también son importantes, sin embargo no fueron objeto de investigación en este proyecto.

7. Las variaciones y/o errores en los resultados estuvieron sujetos al hecho de que el filtrado se realizó solo con el lado izquierdo de la oreja artificial, por lo tanto, la aplicación de los dos filtrados (izquierdo-derecho) podría mejorar las respuestas.

8. Está comprobado que la reverberación es importante para crear una sensación de ambiente parecida a la audición cotidiana; por tal motivo, los instrumentos musicales grabados en campo reverberante, permiten el incremento de la sensación de espacialidad, al contener ésta información del entorno real de captura; aunque la convolución se haya realizado con la respuesta al impulso de la pinna al aire libre. También el hecho de tener este tipo de información ayuda a sentir las fuentes externalizadas (fuera de la cabeza); este punto es reforzado con la distancia, pero dada las dificultades de medición presentadas a lo largo de la investigación, no se pudo determinar la importancia de la distancia para este tipo de representaciones.

9. Para la comprobación de resultados se aplica la grabación hecha en recinto con la respuesta al impulso de la pinna artificial al aire libre. Debido a que los instrumentos tienen información del entorno de captura, los errores fueron minimizados.

10. La aplicación del método utilizado en esta investigación (convolución de señal estéreo a través de una pinna artificial) para obtener sensaciones auditivas similares a las del sonido tridimensional fue importante en el desarrollo de herramientas análogas a las ya existentes, las cuales permiten ser usadas en distintas aplicaciones.

11. Delimitar un tamaño y forma del pabellón auditivo es realmente complejo, debido a que existe una cantidad tan amplia de pabellones auditivos como de personas. Los resultados obtenidos están sujetos a un tipo en particular de pabellón, razón por la cual los datos obtenidos de su función de transferencia al ser presentados a otras personas presentan cierto grado de variabilidad.

Aunque no todos los resultados son mostrados, aquí se concluye que los errores de localización son menores en la técnica estereofónica ORTF para fuentes ubicadas en el plano horizontal, mientras que la técnica Jecklin Disk entrega el mejor rendimiento en la impresión de elevación. A pesar de esto, el comportamiento en general de las técnicas para esta investigación, no fue el más óptimo.

Referencias

- Ángel, Francisco, et al. *Audiología básica* (2006). Universidad Nacional de Colombia, Olga Gómez Gómez, Bogotá D.C.
- Arias Claudia y RAMOS Oscar (2003). Audición espacial en espacios reverberantes: Aspectos teóricos relevantes. *Revista Interamericana de Psicología/Interamerican Journal of Psychology*. Cordoba. Vol.37, Núm. 2, pp. 373-382
- Bartlett, Bruce (1991). *Stereo microphone techniques*. Focal Press, Boston.
- Barlett, Bruce and Bartlett, Jenny (1999). *On-location recording techniques*. Focal Press Boston.
- Batteau, D. W. The role of the pinna in human localization (1966). *The Royal Society of London. Series B, Biological Sciences*. Vol. 168. Internet: <http://www.jstor.org/stable/75711>. pp. 158-180
- Blauert, Jens. *Spatial Hearing, The psychophysics of human sound localization* (1997). Revised edition. MIT, Cambridge, MA.
- Bloom, P. Jeffrey. *Creating Source Elevation Illusions by Spectral Manipulation* (1977). *J. Audio Eng. Soc.* Vol 25. pp. 560-565.
- Begault, Durant. *3-D sound for virtual reality and multimedia* (2000). Ames Research Center, California.
- Carrion Isbert, Antoni. *Diseño acústico de espacios arquitectónicos* (1998).: Editions UPC, Barcelona, España.
- Carrillo, Vicente. *Fisiología de la audición*, Página docente medicina Universidad del Mar. (2008). Internet: http://otorrino.homestead.com/files/Fisiolog_a_de_la_Audici_n.doc
- Cardozo, Mayra. *Empleo de técnicas de grabación para emulación de tridimensionalidad en una aplicación estereofónica* (2009). Trabajo de grado Ingeniera de sonido. Universidad de San Buenaventura, facultad de ingeniería, Bogotá D.C.
- Castro, Francisco. *Análisis comparativo de 4 configuraciones de microfonía estereofónicas* (2005). Trabajo de grado Ingeniero de sonido. Universidad de San Buenaventura, facultad de Ingeniería, Bogotá D.C
- Ceoen, Carl. *Comparative Stereophonic Listening Test* (1972). *J. Acoust. Soc. Am.* Vol 20, pp. 19-27
- Flynn, William E. *Role of the Pinna in Hearing* (1965). *J. Acoust. Soc. Am.* Vol 38, pp. 104.
- García, Luis. Universidad Autónoma de Madrid, página de docencia de fisiología. Internet: http://www.uam.es/personal_pdi/medicina/algvilla/audicion/audicion.htm.
- Gardner, Mark and Gardner, Robert (1973/1972). *Problem of localization in the median plane: effect of pinna cavity occlusion*. *J. Acoust. Soc. Am.* February. Vol 53, pp. 400-408 .
- Grantham, Wesley (1995). *Spatial hearing and related phenomena: Handbook of Perception and Cognition*, Second Edition. Second edition: edited by B. C. J. Moore. Academic, London, UK

- Howard, David and Angus, James (2001). *Acoustics and Psychoacoustics*. Second edition. Francis Rumsey, Oxford.
- International organization for Standardization. *Acoustics, measurement of the reverberation time of rooms with reference to other acoustical parameters*. ISO3382 (1997). Segunda edición. Switzerland.
- Middlebrooks, John and Green, David (1991).. *Sound Localization by Human Listeners*. Annual Review of Psychology Vol. 42, pp 135-159.
- Pérez, José y De la Cuesta, Mariana. *Sonido 3D, apartado 4.2, las coordenadas (Elevación)*. internet: http://www.lpi.tel.uva.es/~nacho/docencia/ing_ond_1/trabajos_03_04/Csound/42.htm
- Raykar, Vikas c, et al (2005). *Extracting the frequencies of the pinna spectral notches in measured head related impulse responses*. J. Acoust. Soc. Am. Vol 118, pp. 364–374.
- Ramos, Oscar, et al (2007). *Modelo acústico de cabeza y torso mediante análisis de componentes principales*. Asociación Argentina de mecánica computacional, vol. XXVI, Córdoba, Argentina
- Schiffman, Harvey Richard. (1997) *Psicofísica: La percepción sensorial*. Segunda edición. limusa, Mexico. pp. 35-44
- Vergara, Ramiro (1990). *Fisiología auditiva, visión actualizada*. Imprenta y Publicaciones de las Fuerzas Militares, Bogotá D.C
- Barti Robert; ORDOYO, Fabiola. *Comparación entre las técnicas binaural y estéreo para la localización de fuentes en el espacio 3D (2001)*. Departamento de acústica Ingeniería y Arquitectura la salle. Universidad Ramon Ilull. Barcelona, España. Internet: <http://www.sea-acustica.es/publicaciones/4366pn004.pdf>
- Watkins, Anthony J (1978). *Psychoacoustical aspects of synthesized vertical locale cues*. J. Acoust. Soc. Am. Vol 63. pp. 1152-1165.
- Wightman, Frederic and Kistler Doris (1989). *Headphone simulation of free-field listening I and II (a,b)*. J. Acoust. Soc. Am. Vol 85, pp. 858-867.