

DISTANCE THRESHOLD ON VIRTUAL SOUND LOCALIZATION ON HUMANS

PACS: 43.66.+y

Larisa Dunai, Ismael Lengua, Juan Manuel Sanchis Rico, Miguel Iglesias
Universitat Politècnica de València
Camino de Vera s/n,
46022 Valencia - España
Tel: +34 689215528
E-Mail: ladu@upv.es

ABSTRACT

The paper presents the distance threshold for complex virtual sounds localization in dependence of the Inter Click Interval. The method of measurement was a forced-choice, three-down one-up staircase. The sound is a complex record of a click sound of 47ms at 44,1kHz. The azimuth is at 0°, at the center of the human head where the ITD=0 varying the intensity and frequency. The investigation reported that the lowest distance thresholds of 0,06m occurred for high Inter Click Intervals at 200-300ms. At lower ICIs the distance threshold increases considerably.

RESUMEN

El artículo describe el umbral de distancia en dependencia del intervalo de tiempo entre dos sonidos en la localización de sonidos acústicos virtuales. El método de medida es elección forzada, escalera tres abajo y una arriba. El sonido acústico representa una grabación del sonido clic de 47ms con la intensidad de 44.1kHz. El sonido está representado para un azimut de 0° en el centro de la cabeza humana con un ITD=0 variando en intensidad y frecuencia. Los resultados demuestran que las distancias más cortas de 0,06m se han obtenido para intervalos de Tiempo entre Clics de 200-300ms. Para Intervalos de Tiempo entre Clics el umbral crece considerablemente.

INTRODUCCIÓN

Después de la segunda Guerra Mundial, con el desarrollo del sensor, se emplearon muchos esfuerzos en diseñar y desarrollar Sistemas Electrónicos de Movilidad capaces de percibir e interpretar el entorno. La idea del desarrollo de los dichos dispositivos surgió como objetivo de superar las limitaciones de los sentidos humanos así como la ceguera.

Actualmente, en el mundo hay más de 45 millones (Informe anual [1] de personas invidentes de donde solo en España [2] hay 14.089 personas invidentes, cuya mayor problema es la movilidad. Foulke en 1971 definió la movilidad como "la habilidad de viajar con seguridad,

confortablemente, con estilo y de forma independiente por el entorno” [3] Para las personas invidentes, la movilidad se basa en el mayor parte en la percepción de sonidos del entorno.

Desde antigüedad, el humano utilizó el sistema auditivo como medio de percepción de sonidos del entorno y en especial a los sonidos situados fuera del alcance visual, permitiéndole de esta forma detectar la dirección y posición de obstáculos evitando peligros [4] Como consecuencia, una de las características fundamentales y más importantes de la detección de obstáculos sonoros del entorno es la habilidad de estimar la locación en distancia y dirección de la fuente sonora.

Sin embargo, la habilidad de estimar las coordenadas de una fuente sonora por un lado, y por otro lado, aunque las personas invidentes tengan el sistema auditivo más desarrollado que las personas con visión normal, no es suficiente para localizar los sonidos espaciales con precisión de unos pocos centímetros en distancia y un grado en azimut. En este caso es estrictamente necesario un periodo de entrenamiento que ayude al cerebro humano a definir a un sonido espacial sus coordenadas tanto en distancia, azimut como elevación.

En este contexto, los investigadores han apostado por el estudio y el desarrollo de Dispositivos Electrónicos de Movilidad para personas invidentes en base a sonidos acústicos especializados, siendo el método más rápido y más eficaz en representar el entorno percibido mediante sensores ultrasonido, láser o estéreo-visión, así mismo en la mejora de los sistemas de visión artificial y los sistemas acústicos.

La localización de sonidos acústicos virtuales representa un factor importante en la vida humana que supone un porcentaje elevado dentro del conjunto de características psicoacústicas. En mayoría, la localización de sonidos acústicos virtuales se estudia utilizando situaciones cuando el sujeto escucha mediante auriculares los sonidos, donde cada sonido ha sido medido en una cámara anecoica para cada posición del espacio para el oído izquierdo y el oído derecho individualmente para cada sujeto, o se han medido las Funciones de Transferencia con Respeto a la Cabeza llamados Head Related Transfer Function (HRTF en inglés) para cada posición del espacio tanto para el oído izquierdo como para el oído derecho con la ayuda de un maniquí KEMAR y se han convolucionado con el sonido seleccionado [5] Cualquiera de las modalidades las de medida o de generar los sonidos acústicos virtuales, ayudan a localizar sonidos acústicos en un entorno de dos o tres-dimensiones siempre que el sujeto está presente en los procesos de medida. La localización de sonidos acústicos virtuales en cámaras anecoicas mediante altavoces queda un problema debido a que la localización depende de los movimientos de la cabeza, siempre que la cabeza se mueve, los sonidos llegan a los dos oídos con distinta longitud de onda modificando la posición de la fuente sonora, afectando a la localización del dicho sonido. En estos casos se han aplicado varios métodos de fijación de la cabeza del sujeto como por ejemplo sujetando la cabeza a una barra fija en la parte superior de la cabeza o una barra que atraviesa la boca del sujeto, que no permita al sujeto mover siquiera un centímetro, que crea una sensación de tortura, incomodidad y agobio hacia el sujeto ya que cada experimento dura de 10 a 30 minutos en una sala pequeña. Por otro lado, los experimentos de localización de sonidos acústicos virtuales en cámaras anecoicas no utilizan ninguna interfaz visual donde el sujeto pueda percibir de forma visual el origen del sonido. Asimismo, los diferentes métodos aplicados en los experimentos de localización de sonidos acústicos virtuales se han basado en los métodos tradicionales.

Debido a estos y otros inconvenientes, durante los últimos años se ha propuesto el estudio de localización de sonidos acústicos virtuales mediante auriculares, como método para los Sistemas Acústicos. La localización mediante auriculares para los Sistemas Acústicos necesita de modo obligatorio tener medidos previamente las funciones de transferencia con respecto a la cabeza y convolucionados con los sonidos mediante la Transformada Fourier Discreta TFD dada por (1).

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N}kn} \quad k = 0, \dots, N - 1 \quad (1)$$

En este sentido, la Transformada Fourier Discreta se constituye como una herramienta que permite realizar la convolución y correlación cruzada entre la señal acústica y los HRTFs tanto para los sonidos monaurales como los sonidos binaurales o estéreo creando la sensación del espacio, frente a la convolución simple de dos señales dadas en la (2) y Figura 1.

$$y(n) = x(n) * h(n) \quad (2)$$

Donde, $x(n)$ representa la señal con el número n de muestras y el $h(n)$ representa la Función de Transferencia HRTF con el mismo número de muestras.

En cualquier caso, tanto para la localización de sonidos acústicos virtuales en cámaras anecoicas o mediante auriculares, la localización se ha estudiado en mayoría para dos situaciones: para la localización direccional que implica la localización en azimut, es decir a la derecha o izquierda y para la localización en elevación que representa el estudio de localización en un plano vertical.

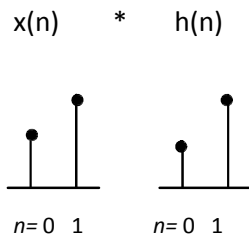


Figura 1. Esquema matemático de la convolución de dos señales acústicas $x(n)$ y $h(n)$ con el mismo número de muestras.

Menos estudios se han hecho en localización de sonidos para los tres planos: azimut, en la eje x , elevación en las ejes z , y en distancia el eje y y Figura 2.

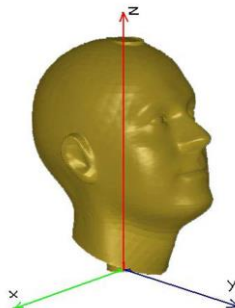


Figura 2. Representación gráfica del espacio tres dimensional.

Con el desarrollo de los dispositivos electrónicos y con la vasta aplicación del ordenador en los ensayos experimentales en acústica y psicoacústica, resulta en un mejor manejo y control tanto del experimento como de la interpretación y análisis de los resultados obtenidos. Durante las

últimas décadas los investigadores han intentado diseñar unas Interfaces Gráficas del Usuario con el objetivo de obtener una mejor evaluación de los resultados experimentales y reproducción de sonidos, así como permite diseñar unos experimentos mucho más complejos, controlar diferentes factores y variar los parámetros necesarios para el control humano. Como ventaja de las interfaces gráficas del usuario en los experimentos acústicos es la posibilidad de reducir los errores experimentales. La mayoría de las interfaces desarrolladas se aplican en los experimentos psicoacústicos de dos dimensiones.

Sin embargo, la mayoría de las interfaces se desarrollan para analizar las propiedades acústicas de los sujetos [6], [7], [8] pero no hay interfaces que entren a los sujetos en localizar los sonidos espaciales en tres dimensiones tanto en dirección (azimut o sonidos situados en un plano horizontal), distancia y elevación (sonidos situados en posición vertical). No obstante, es de vital importancia diseñar una Interfaz Gráfica del Usuario que se utilice como medio de entrenamiento y aprendizaje de los sonidos virtuales con el fin de analizar el proceso de aprendizaje como el cerebro humano aprende a localizar los sonidos mediante los sistemas de visión y auditivos, y para utilizar el dispositivo acústico de detección de obstáculos y navegación para personas invidentes. El desarrollo de una Interfaz Gráfica del Usuario de libre acceso para tres dimensiones sería una tremenda utilidad en el estudio de localización de sonidos virtuales para no solo para las investigaciones de localización de sonidos en dirección y elevación sino para los experimentos psicoacústicos de localización para distancia, ya que la localización en distancia juega un rol muy importante en los experimentos con los dispositivos acústicos de navegación para personas invidentes.

El diseño y el desarrollo de las interfaces gráficas del Usuario para los experimentos acústicos requieren un análisis minucioso en los pasos de diseño, los modos de representación de la información acústica y visual y al final de representación de resultados.

METODOLOGÍA

Para llevar a cabo el experimento, se han desarrollado los siguientes pasos.

Determinados los requerimientos para la localización de sonidos virtuales donde se ha comprobado la dicha relación teórica se desarrolló un ensayo donde se cambiaron los parámetros acústicos del sonido tanto en tiempo como en intensidad para sonidos de frecuencia 44,1kHz. Dichos ensayos analizaron sonidos de materiales así como el cristal, la madera y sonidos generados de forma artificial que no represente ningún tipo de material. Se analizaron las propiedades de localización de cada sonido haciéndose una comparación entre las precisiones. Los resultados de este experimento ayudaron a obtener una información muy útil para una mejor comprensión del proceso de localización de sonidos acústicos y como va afectando los diferentes tipos de sonidos a la localización de sonidos acústicos virtuales [9].

Diseñado la interfaz acústica generando las coordenadas acústicas espaciales, se desarrollaron modelos de entornos para obtener señales HRTF utilizando el CATT ACOUSTICS. Con los modelos desarrollados mediante el software CATT ACOUSTICS se realizaron medidas de las coordenadas espaciales llamadas Head Related Transfer Function (HRTF) para las distancias de 1m a elevación de cero grados y azimut de cero grados. Se tomaron medidas tanto para entornos cerrados como para entornos abiertos utilizando señales puras. Se aplicó una ligera reverberación sobre las señales puras para obtener sonidos espaciales con reverberación.

Desarrollado el banco de sonidos virtual utilizando el método de convolución mediante la Transformada Fourier Discreta entre las Funciones de Transferencia a la Cabeza HRTFs y los sonidos medidos previamente. Se realizaron ensayos con sonidos para los entornos en dos-

dimensiones y tres-dimensiones, midiendo en cada caso la precisión de la localización de sonidos en azimut, distancia y elevación. Así mismo para desarrollar las distancias en las pruebas se utilizó la teoría de división por intensidad del sonido:

$$dB = 10 \log P2/P1 \quad (3)$$

En la localización de sonidos acústicos influyen una serie de parámetros acústicos. Los parámetros más importantes son la Diferencia Interaural de Tiempo ITD, la Diferencia Interaural de Intensidad ILD y el Tiempo de silencio entre dos fuentes sonoras [10].

Se realizaron ensayos experimentales de localización de sonidos de distintos materiales mediante auriculares, en los que se modificarán los parámetros acústicos. Inicialmente se realizaron pruebas experimentales con para cada tipo de sonido individual modificando los parámetros de tiempo e intensidad buscando el umbral mínimo de tiempo interaural y la precisión de la localización de sonidos virtuales. Se hará una comparación entre resultados de cada sonido. Se validaron los resultados obtenidos utilizando entornos con cantidad baja de ruido externo y en condiciones reales en los que se tiene la sospecha de la influencia de dichos ruidos en la precisión de percepción y localización de sonidos virtuales. Así mismo se seleccionará el sonido acústico con mejores resultados para la futura investigación mediante interfaces gráficas de dos-dimensiones y de tres-dimensiones.

Diseñado y programado la interfaz acústica. En esta fase se ha desarrollado mediante el paquete C++ la interfaz acústica, que proporcione una amplia variedad de métodos y posibilidades de representación de sonidos acústicos virtuales, sin modificar las propiedades de las señales. Se diseñó las estrategias de representación de sonidos y comunicación con las interfaces audio y la imagen. Finalmente, se desarrolló la interactividad de la interfaz acústica con las interfaces audio, la imagen y el usuario ver Figura 3. Para ello se utilizó el sistema de programación C++10 apoyándose en las librerías OpenSource QT y el IDE QtCreator, y algunas librerías MCR de Matlab.



Figura 3. La interfaz gráfica del usuario para introducción de datos.

Una vez introducidos los datos de ha diseñado una interfaz gráfica para la selección de áreas de trabajo así como en un área de 0m a 1m solo para distancias o solo para el azimut y la tercera solución es para las pruebas integradas con distancia y azimut distribuidas por columnas Figura 4.

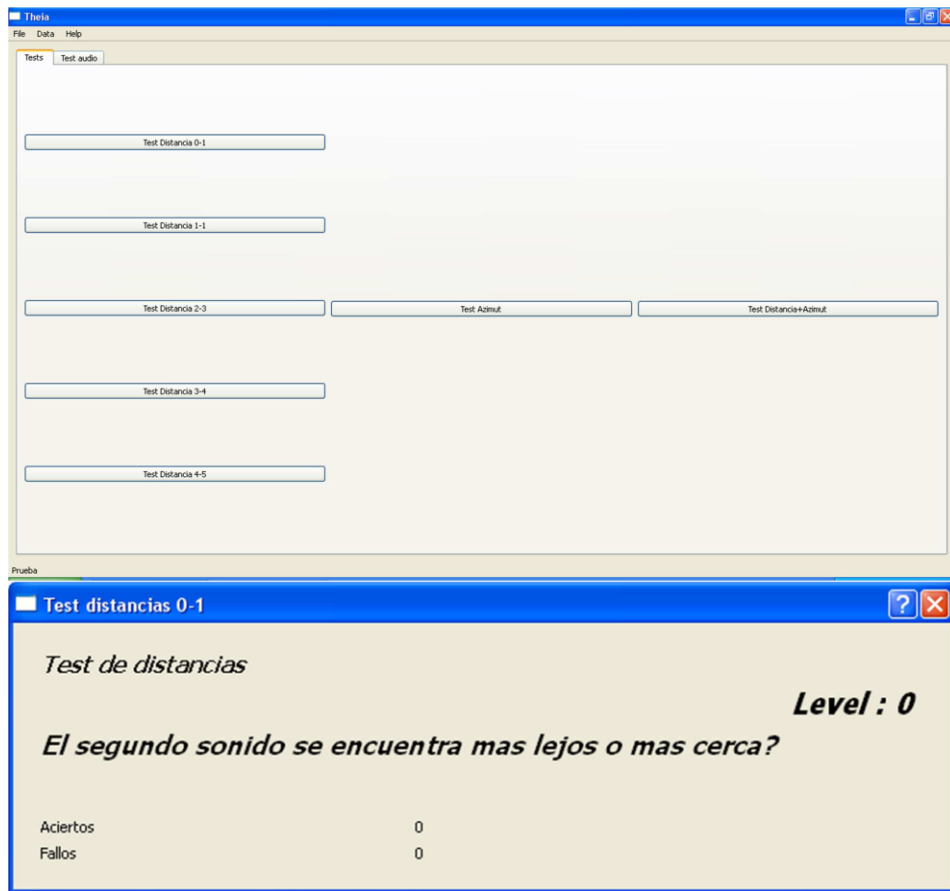


Figura 4. Interfaz gráfica de selección de áreas experimentales distribuida por columnas y filas.

La primera columna indica las pruebas solo en distancia a un azimut de 0°. La segunda fila representa las pruebas para la distancia de 1m en azimut y la tercera fila son las pruebas experimentales de localización de sonidos combinando las distancias y el azimut. En la segunda ventana es la interfaz gráfica para el observador o investigador que informa la aprobación de las respuestas del usuario.

Así mismo los resultados se guardan en un fichero .txt con las identificaciones del usuario y todos los datos de entrada y las respuestas del usuario.

La metodología experimental es “selección forzada” que obliga al usuario seleccionar si o si una de las dos respuestas. Para una mejora evaluación de los resultados se ha empleado el método experimental escalera tres abajo y una arriba. El método consta en que a cada tres respuestas correctas disminuye la distancia con un paso de 10 cm para las distancias de 1m hasta 10 cm y con 1cm desde 10 cm hasta 0m. Este método permite definir una precisión de 75% en las respuestas en localización de sonidos.

RESULTADOS

El objetivo principal del trabajo es definir el umbral en distancia para la localización de sonidos virtuales mediante auriculares para distancias cortas. En las pruebas los usuarios han escuchado dos sonidos situados a 0° en elevación a diferentes distancias. Todos los usuarios han demostrado tener capacidad de percibir los sonidos. Después de una serie de pruebas con seis usuarios se ha definido que la distancia mínima percibida entre dos sonidos situados a 0° en azimut, es decir que están situados en línea central con respecto a la cabeza, de 6cm. Este valor fue percibido para un Tiempo de Silencio entre Sonidos de 250ms, 350ms y 400ms.

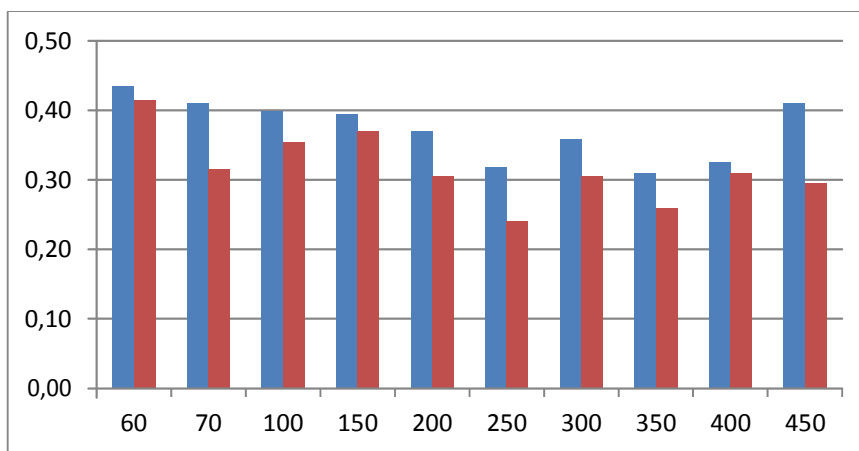


Figura 5. Resultados del umbral promedio para los cinco usuarios en m. El eje x

representa el Tiempo de Silencio entre sonidos y el eje vertical representa la distancia en metros. La barra de color azul representa la respuesta con auriculares de diadema cerrados modelo HD 201 de Sennheiser, los resultados de las barras de color rojo representan los datos obtenidos por los mismos usuarios con auricular óseo NOKIA.

Según los resultados los mejores datos se han obtenido para un Tiempo de Silencio entre Sonidos o el Intervalo de Tiempo entre Clics de 250ms. En cuanto más disminuye o incrementa el TSS incrementa la distancia de forma considerable.

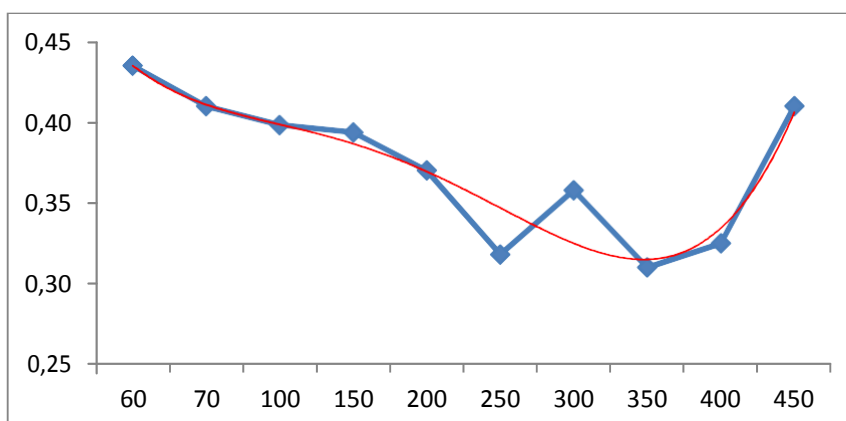


Figura 6. Resultados de obtenidos mediante los auriculares HD 201 de Sennheiser de diadema cerrados y su línea de tendencia.

De Figura 6 se observa que los resultados mejores obtenidos en distancias son para los TSS entre 200ms y 350ms. La línea de tendencia o "fit curve" es polinómica de grado 4. La línea de tendencia nos permite definir una fórmula matemática del comportamiento del oído humano. También el valor R de la línea de tendencia $R^2=0.8779$.

CONCLUSIONES

En el artículo se han presentado los datos obtenidos para definir el umbral de distancia dependiendo del intervalo de tiempo entre dos sonidos en la localización de sonidos acústicos virtuales. El método de medida utilizado es elección forzada, escalera tres abajo y una arriba. El sonido acústico representa una grabación del sonido clic de 47ms con la intensidad de 44.1kHz. El sonido está representado para un azimut de 0° en el centro de la cabeza humana con un ITD=0 variando en intensidad y frecuencia. Los resultados demuestran que las distancias más cortas de 0,06m se han obtenido para intervalos de Tiempo entre Clics de 200- 300ms. Para Intervalos de Tiempo entre Clics mayores de 400ms e inferiores a 100ms el umbral crece considerablemente. Así mismo se ha percibido una diferencia entre los resultados con auriculares cerrados y con auriculares óseos. Con los auriculares óseos los usuarios mejoraron la distancia con aproximadamente 8cm. Lo que indica que para sonidos cercanos 8cm es una cuestión de vida o muerte para una persona con discapacidad visual.

AGRADECIMIENTOS

Desarrollo realizado dentro del Proyecto financiado por la Generalitat Valenciana GV/2014/015 proyectos de I+D para grupos de investigación emergentes.

REFERENCIAS

- [1] WBU (World Blind Union) Available: <http://www.who.int/mediacentre/factsheets/fs282/en/> [Citado 28 de Abril 2013]
- [2] ONCE España, "Informe sobre ceguera," Ernser & Young S.L., 2012
- [3] Foulke E., "The perceptual basis for mobility," in AFB Res. Bull, vol. 23, pp. 1-8, 1971.
- [4] Fay R. R., Propper A. N, "Introduction to sound source localization," Book chapter 1 on sound source localization, Springer, pp. 1-5, 2002
- [5] Minnaar P., "Simulating an Acoustical Environment with Binaural Technology – investigations of binaural recording and synthesis", Ph.D. thesis obtained at Aalborg University, Denmark, September 2001.
- [6] Laneau J., Boets B., Moonen M., Van Wieringen A., Wouters J., „A flexible auditory research platform using acoustic or electric stimuli for adults and young children," Journal of Neuroscience Methods 142, pp.131-136, 2005
- [7] Usher J., Woszczyk W., "Design and testing of a graphical mapping tool for analyzing spatial audio scenes," AES 24th International Conference on Multichannel Audio, Banff, Canada, 26- 28.06.2003

- [8] Choisel S., Zimmer K., "A Pointing-Technique with Visual Feedback for Sound-Source Localization Experiments," AES 115th Convention, October 10-13, New York, printing number 5904, 2003
- [9] Dunai L., Peris-Fajarnés G., Ismael Lengua Lengua, Ignacio Tortajada (2012), Material sound source localization through headphones, Acoustical Physics, 58(5), 610-617
- [10] Hartmann W.M., Dunai L., Tianshu Qu (2013), Interaural Time Difference Thresholds as a Function of Frequency, Advances in Experimental Medicine Basic Aspects of Hearing, Physiology and Perception, 787, 239-246