



## THREE DIMENSIONAL ENVIRONMENT LOCALIZATION THROUGH VIRTUAL SOUND EXPERIMENTAL

Larisa Dunai<sup>1</sup>, Andrés Prieto<sup>1</sup>, Ismael Lengua<sup>1</sup>, Beatriz Defez<sup>1</sup>

<sup>1</sup>Universitat Politècnica de València

([ladu@upv.es](mailto:ladu@upv.es), [anprlo@posgrado.upv.es](mailto:anprlo@posgrado.upv.es), [ilengua@dig.upv.es](mailto:ilengua@dig.upv.es), [bdefez@upv.es](mailto:bdefez@upv.es))

### Resumen

El artículo presenta el desarrollo de un experimental para la localización de 32 sonidos simultáneamente. El experimental se basa en el desarrollo de dos interfaces: la interfaz gráfica que representa un entorno virtual tridimensional que contiene distintas formas geométricas planas y tridimensionales, así como figuras del entorno real y la interfaz acústica virtual.

El experimental tiene como objetivo el estudio de localización de sonidos complejos múltiples reproducidos simultáneamente mediante los auriculares, así como estudiar las características psicoacústicas de localización de sonidos múltiples.

Para el desarrollo del experimental se ha analizado la Diferencia Interaural del Tiempo y la Diferencia Interaural de Intensidad así como dependencia de la frecuencia y el umbral del tiempo para la localización de sonidos variando el azimut y la intensidad.

Las pruebas experimentales previas han demostrado que los humanos son capaces de distinguir múltiples sonidos y localizarlos en el espacio.

**Palabras-clave:** sonido, virtual, localización, interaural

### Abstract

The paper deal with the development of an experimental for 32 virtual simultaneous sound localization. The experimental is based on the development of two interfaces: the graphic virtual environment that includes different images as plane geometric figures and three dimensional geometric figures and daily figures as trees, etc. and the virtual acoustic sounds. The experimental treats to study simultaneous sound sources localization through headphones, as the study of the psychoacoustic characteristics on simultaneous sound localization.

For the experimental was analysed the time dependence and intensity and frequency dependence as time thresholds. The virtual sounds are generated in real time in elevation, direction and depth.

Previous results demonstrate that the humans are able to perceive simultaneous sounds and to localize them in the environment.

**Keywords:** sound, virtual, location, interaural

**PACS no** 43.66.+y



## 1 Introducción

La supervivencia de los seres humanos depende no solamente de las necesidades de comer y beber. El ser humano como ser complejo, está dotado con cinco sentidos principales, sin los cuales no seríamos vivos: el sentido de la vista – que mediante el órgano de visión llamado ojo y mediante el nervio óptico se conecta al cerebro que interpreta las imágenes recibidas por los dos ojos que hace que podamos ver los objetos del entorno real y localizarlos; el sentido del oído – que mediante el oído nos permite percibir y localizar los sonidos del entorno; el sentido de gusto nos permite percibir el gusto o el sabor de los alimentos, el olfato o el sentido del tacto.

No es un secreto que los humanos tienen la habilidad de percibir y localizar una o varias fuentes sonoras y separarlas tanto en entornos silenciosos como entornos ruidosos y abarrotados, entornos controlados como entornos abiertos. Este fenómeno que nos permite localizar los sonidos del entorno se llama la audición binaural. En la localización de sonidos espaciales un gran rol juega los factores físicos así como la Diferencia Interaural de Tiempo (DIT), la Diferencia Interaural de Intensidad (DII), el torso y la forma del oído. Es bien sabido que la DIT es un factor que se crea especialmente debido a las distancias de la fuente sonora hasta los dos oídos, la cual es igual cuando la fuente sonora está situada a  $0^\circ$  en frente. La DII se genera debido a la sombra que crea la cabeza.

La mayoría de la investigación enfoca en la localización de los sonidos sinusoidales en dirección midiendo los umbrales DIT y los DII para frecuencias no más de 1.5 kHz. Así como analizando la localización dependiendo de la frecuencia e intensidad (Hartman, 2014, Kuhn 1977). En caso de que la frecuencia sobrepasa los 1.5 kHz, se analiza la Diferencia Interaural en Fase (DIF).

Batteau (1967) llegó a la conclusión de que la localización de sonidos no es solo trabajo de DIT o DII si no que la forma del oído o pinna también, que en teoría ayuda a atenuar las señales acústicas y sirve como un simple sistema de retardo. La pinna juega un gran rol en la localización de sonidos del entorno así como en la localización de sonidos en planos verticales.

## 2 Localización de múltiples fuentes sonoras

Durante muchos años se ha trabajado en la localización de una sola fuente sonora tanto en entornos abiertos como sonidos virtuales en planos horizontales y verticales. Brungart, 2005, demostró la posibilidad de localizar múltiples sonidos en el entorno, llegando a la conclusión de que la precisión disminuye sistemáticamente con el incremento del número de sonidos reproducidos. También menciono que la localización de sonidos es mejor cuando el comienzo de los sonidos es más rápido.

La mayoría de los experimentales trata de localizar un solo sonido de un entorno con múltiples fuentes sonoras. Los sonidos más utilizados son los tonos puros acompañados por un sonido enmascarado de ruido blanco. Good y Gilkey (Good and Gilkey, 48) utilizaron en su experimental un barrido de sonidos clic de banda ancha enmascarado demostraron que la localización en azimut no está influenciada por la interferencia de los ruidos enmascarados, sin embargo se ve una leve influencia en la localización de sonidos en elevación y una gran influencia en la localización en la parte trasera de la cabeza. Similares resultados han sido obtenidos por Lorenzi, (74). Todos los autores han llegado a la conclusión de que la precisión en la localización de sonidos en entornos con múltiples fuentes sonoras disminuye cuando el número de sonidos interferentes crece de 2 a 4. Brungart concluyó que al tener más de cinco sonidos interferentes, el movimiento de cabeza, que es un factor muy importante en la localización de sonidos, no influye en la mejora de la precisión de localización de sonidos.

Resultados muy similares han obtenido en la localización del habla o “cocktail party”. Kawashima y Sato simularon el cocktail party con el fin de obtener el umbral de localización e identificación de fuentes sonoras (en el caso se utilizó el habla). Llegaron a la conclusión de que el tiempo de razonamiento sobre la identificación y localización del habla disminuye considerablemente cuando el

número de personas que hablan crece. Los límites perceptivos se ha obtenido para un número de personas que hablan es entre 3 y 5, en este caso el valor de aciertos es de 0.5, teniendo en cuenta la posición del usuario. Yost et al 1996 han llegado a la conclusión de que para la percepción y localización de múltiples fuentes sonoras, en el caso se trata del palabras, letras y números, juega un rol importante el proceso binaural.

### 3 Desarrollo

Para el desarrollo del experimental se ha utilizado la multiplataforma abierta OpenAL desarrollada por Creative Labs y las bibliotecas de OpenGL para el desarrollo del entorno virtual. La plataforma OpenAL es una API de audio que permite renderizar y posicionar los sonidos multicanal en tres dimensiones.

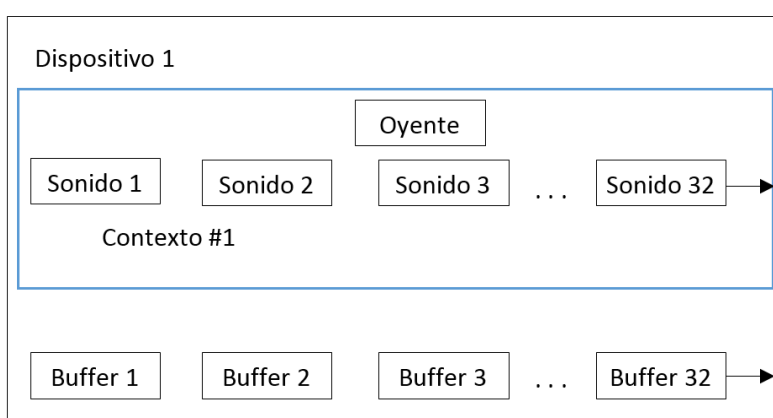


Figura 1 - Ejemplo de jerarquía entre objetos en OpenAL.

La segunda API es la de OpenGL que nos permite crear entornos en dos dimensiones y entornos en tres dimensiones a partir de primitivas geométricas simples tales como puntos, líneas y triángulos.

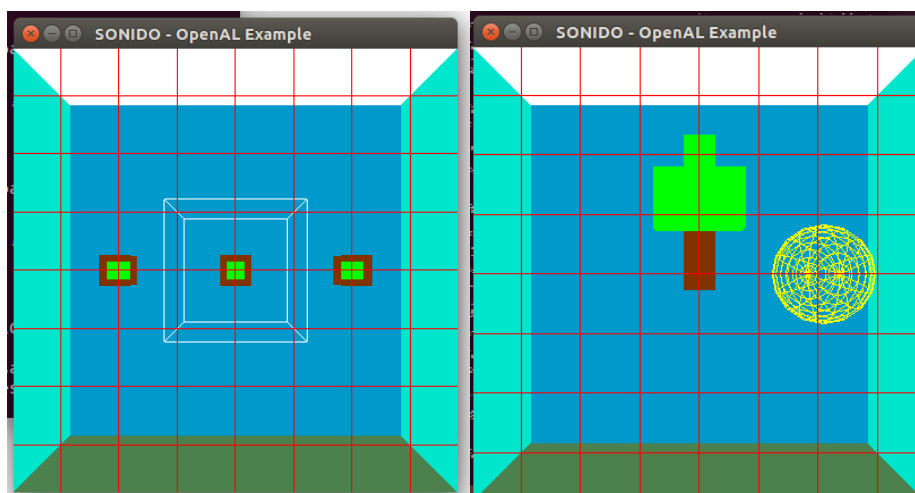


Figura 2 - Ejemplo de la interfaz gráfica del usuario en tres dimensiones nivel básico del experimental. El fondo de la habitación es de color azul y el oyente es de color blanco en la primera imagen y en la segunda es de color amarillo.



En la primera imagen primitiva se colocaron los obstáculos sonoros en forma de cubos de color azul y marrón en una posición horizontal, en la segunda imagen los sonidos se colocaron en forma de árbol. Con la ayuda del teclado, se puede mover el oyente mediante el entorno arriba y abajo, a la derecha y a la izquierda y delante y hacia atrás. Con cada movimiento del oyente, al cambiar sus coordenadas cartesianas, los sonidos cambian de posición. De esta forma el oyente percibe un conjunto de barridos provenientes del objeto.

Una vez esta la interfaz gráfica lista, se combina con los algoritmos de reproducción de los sonidos en tiempo real desarrollados en OpenAL. El algoritmo permite la reproducción simultánea de los sonidos acústicos, así como a posibilidad de cambiar constantemente la posición de los sonidos en el entorno virtual generado.

En el experimental se han experimentado con uno, dos, tres y siete sonidos reproducidos simultáneamente. Dependiendo de la posición de los sonidos reproducidos, los participantes escuchaban como un sonido reproducido directamente en la cabeza. En cuando se reproducían los siete sonidos posicionados en el mismo plano horizontal y a la misma distancia en un ángulo de  $30^\circ$  izquierda a  $30^\circ$  a derecha con un azimut de aproximadamente  $8^\circ$ , se escuchaba una superficie casi continua que daba la sensación de una pared o un muro situado a una distancia de 1 metro delante del oyente.

Durante el experimental en 99% ha sido necesario bajar el volumen de los sonidos reproducidos mediante los botones de ajuste del volumen del ordenador para percibir mejor la posición de los sonidos. A alta intensidad del sonido, un 98% de los oyentes han percibido un sonido muy molesto y difícil de percibir y distinguir e número de sonidos reproducidos simultáneamente.

Debido a que el objetivo del experimental no es identificar cada sonido reproducido si no es percibir una imagen sonora en un entorno dos y tres dimensional, no se ha hecho enfoque en localizar cada sonido en parte si no percibir y detectar un conjunto de sonidos y comentar la forma geométrica que crean los grupos de sonidos en un conjunto. Para las imágenes en dos dimensiones, no se ha utilizado el tiempo de retardo entre cada sonido, hecho que nos permitió reproducir en tiempo real y simultáneamente los siete sonidos.

Hay que mencionar que el objetivo del experimental es crear un videojuego que sirva como plataforma de aprendizaje para la percepción y localización de imágenes acústicas asociadas a objetos del entorno real.

Para el futuro se pretende combinar imágenes del entorno real mediante una cámara web o la cámara del móvil y transformar las imágenes video en imágenes acústicas.

## 4 Conclusiones

El experimental para la localización de 32 sonidos simultáneamente ha sido desarrollado con el fin de estudiar la localización de sonidos en el entorno virtual y en el entorno real, así como estudiar la localización de múltiples sonidos simultáneos y su percepción por los usuarios. El experimental se desarrollo mediante dos interfaces: la interfaz gráfica y la interfaz acústica virtual.

Evaluando los resultados previos del experimental se ha observado que el ser humano es capaz de crear superficies sonoras o interpretar superficies de los objetos, estimular el cortex visual interpretando de esta forma las superficies de los objetos mediante sonidos.

El interval de tiempo entre sonidos establecidos en azimut es de 5 ms como umbral del tiempo definido por la señal utilizada.



## Agradecimientos

Los autores agradecen a todas las personas que apoyaron el desarrollo del experimental.

## Referencias

- [1] Brughera A., Dunai L., Hartmann W., “Human interaural time difference thresholds for sine tones: The high-frequency limit,” *J. Acoust. Soc. Am.* Vol 133 (3), pp.2839-2855, May 2013
- [2] Kuhn G.F., “Model for the interaural time differences in the azimuthal plane,” *J. Acoust. Soc. Am.*, Vol.34, pp.157-167, 1977
- [3] Batteau D., “The role of the pinna in human localization,” *Proc. R. Soc London B Biol. Sci.*, Vol. 168, pp.158-180, 1967
- [4] Brungart, D.S.; Simpson, B. D.; Kordik, A. J., “Localization in the Presence of Multiple Simultaneous Sounds,” *Acta Acustica united with Acustica*, Vol.91(3), pp. 471-479, May/June 2005
- [5] Good M. D., Gilkey R. H., “Sound localization in noise: the effect of signal-to-noise ratio,” *J. Acoust. Soc. Am.*, Vol.99, pp.1108-1117, 1996
- [6] Lorenzi C., Gatehouse S., Lever C., “Sound localization in noise in normal-hearing listeners,” *J. Acoust. Soc. Am.*, Vol.105, pp.1810-1820, 1999
- [7] Kawashima T. and Sato T., “Perceptual limits in a simulated “cocktail party,” *Atten. Percept. Psychophys*, Vol. 77(6), pp.2108-2120, August 2015
- [8] Yost W.A., Dye R. H., Sheft S., “A simulated “coctel party” with up to three sound sources,” *Percept Psychophys*, Vol.58(7), pp.1026-1036, Oct 1996