

**EUSKAL HERRIKO UNIBERTSITATEA
UNIVERSIDAD DEL PAÍS VASCO**



ESCUELA SUPERIOR DE INGENIEROS DE
BILBAO

BILBOKO INGENIARIEN
GOI ESKOLA

TRABAJO DE ELECTROACÚSTICA TEORICO-PRÁCTICO

SISTEMA DE SONIDO 3D

PRESENTA: Alejandro García
PROFESOR: Imanol Madariaga

Bilbao, Junio de 2005

ÍNDICE

1. MOTIVACIÓN DEL TRABAJO	3
2. INTRODUCCIÓN	4
3. MECANISMOS DE LOCALIZACIÓN DEL SONIDO	6
3.1. SISTEMA DE COORDENADAS	6
3.2. DIRECCIÓN	8
3.2.1. Localización en el plano horizontal	8
3.2.2. Localización en el plano medio	11
3.2.2.1. HRTF (Head Related Transfer Function)	13
3.3. DISTANCIA	16
3.3.1. Loudness	16
3.3.2. Motion Parallax	16
3.3.3. Reverberación	16
3.3.4. Proporcionalidad entre Reverberación y sonido seco	16
3.3.5. Balance Espectral	16
3.3.6. Difusión	16
4. SISTEMA DE SONIDO 3D	17
4.1. Circuito HRTF	17
4.2. Creación de sonido 3D mediante grabaciones binaurales	20
5. PRACTICAS	23
5.1. VSPACE	24
5.2. AMPHIOTIK ENHACER	40
6. PRÁCTICAS A DESARROLLAR	45
7. REFERENCIAS	46

1. MOTIVACIÓN DEL TRABAJO

La motivación de este trabajo surge de la necesidad de conocer como los sistemas convencionales de sonido, formados por un par de altavoces, son capaces de generar sonido envolvente.

Conocer mediante que métodos, los sistemas de sonido son capaces de generar mediante un par de altavoces sonidos virtuales procedentes de cualquier punto del espacio.

A día de hoy, los sistemas que mas nos pueden sonar en cuanto a sistemas de sonido envolvente, son los sistemas Dolby Surround, THX, etc.. Pero si bien es verdad estos sistemas utilizan varios altavoces, normalmente 5, además de un decodificador especial para generar sonido por cada uno de los altavoces.

El objetivo de este trabajo es llegar a comprender y analizar como es posible mediante un par de altavoces simular un sonido ambiental o 3D.

Como puesta final a este trabajo se realizarán una serie de prácticas relacionadas con el sonido 3D.

2. INTRODUCCION

Antes de la aparición del sonido Dolby, uno de los objetivos que se propusieron los ingenieros de sonido era conseguir mediante un par de altavoces generar sonido en tres dimensiones, es decir, poder experimentar como los sonidos que se transmitían por un par de altavoces, los podíamos percibir de todo el entorno.

La percepción que un oyente podría experimentar con un simple sistema de altavoces y un sistema capaz de generar sonido 3D, podría ser por ejemplo, escuchar los coches que se mueven, a su izquierda oír los pasos de los peatones, las obras de dos calles mas arriba, el sonido de los pájaros y de las palomas por encima del oyente.

En nuestro ambiente natural percibimos los sonidos en tres dimensiones y podemos localizarlos en cualquier instante. En realidad, el cerebro se ha entrenando toda la vida, y tiene memorizados una serie de parámetros basados en la experiencia y personalizados para cada uno.

Sin embargo, en la música grabada, esto no es así. El estéreo tradicional nos da algo de información espacial, pero no lo suficiente para recrear que estamos en una habitación escuchando un concierto en directo con la banda instalada en nuestra habitación. Es más, en vez de dar una sensación de que estás en un espacio tridimensional, la reproducción por altavoces da la sensación que estas en frente de la fuente de sonido.

En realidad el sonido estéreo, puede recrear una sensación de especialidad, ya que podemos tener la sensación de estar escuchando sonidos diferentes. En el caso del sonido 3D se crearán sensaciones espaciales que no se experimentan con el estéreo.

En el caso de los altavoces se espera la posibilidad de situar sonidos fuera de los límites de los altavoces y, de los auriculares, se espera que el sistema sitúe los sonidos fuera de la cabeza.

La clave del desarrollo en el sonido 3D, se debe al uso del procesado de señal mediante funciones de transferencia dependientes direccionalmente, o funciones HRTF.

El modelo científico del sonido 3D se puede dividir en tres disciplinas para llegar a un entendimiento del mismo.

- Acústica física
- Psicoacústica
- Neuropsicología auditoria.

La *acústica física* se centra en las ondas sonoras que alcanzan al receptor desde el pabellón aditivo y el fenómeno acústico que determina sus propiedades específicas.

La *psicoacústica*, estudia la relación entre las ondas acústicas en el pabellón auditivo y la percepción de la imagen espacial que experimentan los receptores.

La *neuropsicología auditoria*, trata la comprensión de las estructuras neuronales que dan sentido a la sensación del sonido

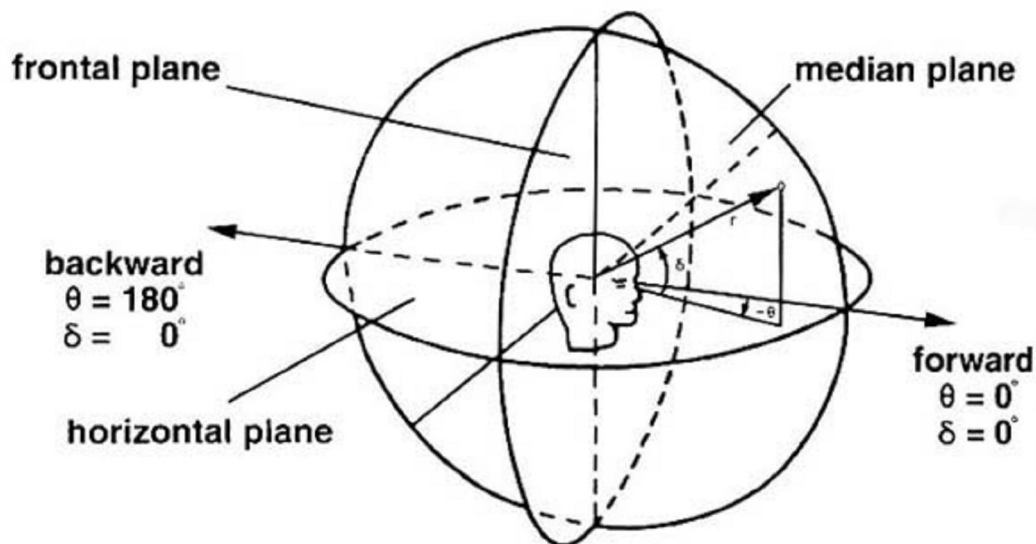
Antes de analizar como se emite sonido envolvente por un par de altavoces, se explicará como nosotros, los humanos, localizamos el sonido en un ambiente para comprender los trucos que se utilizan en el proceso de señal.

3. MECANISMOS DE LOCALIZACIÓN DEL SONIDO

Primero vamos a definir que es la localización. La localización define la capacidad del individuo de determinar la ubicación de una fuente sonora en el espacio.

La localización sólo es posible a partir de la audición binaural, es decir, teniendo dos oídos.

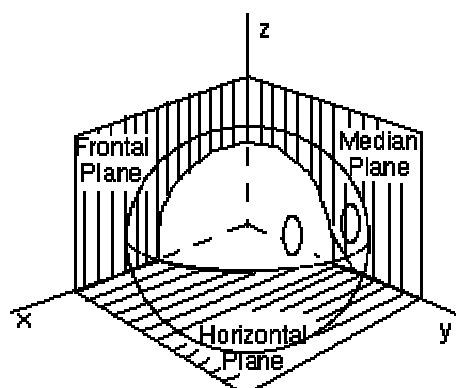
Por lo general se establecen tres planos característicos en los experimentos destinados a estudiar la localización por parte del ser humano.

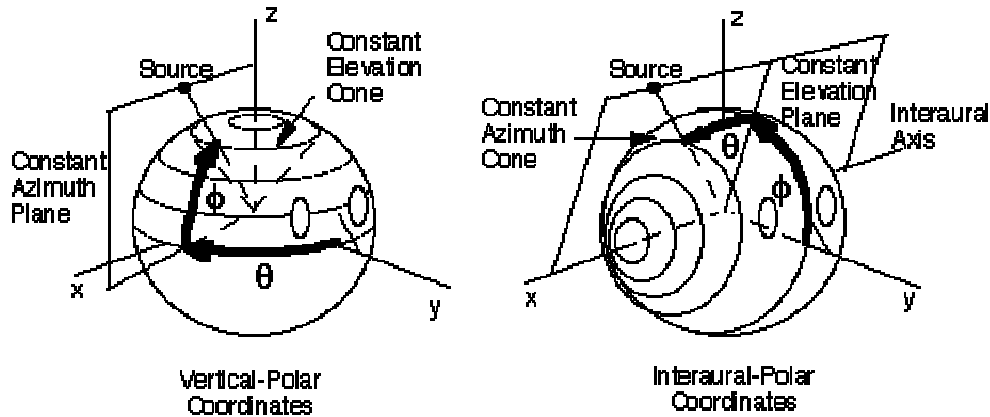


3.1. SISTEMA DE COORDENADAS

Para especificar la localización de un sonido en el entorno tridimensional, se necesita un sistema de coordenadas que nos ayude a localizar los sonidos en la cabeza del oyente, como el sistema que se muestra abajo.

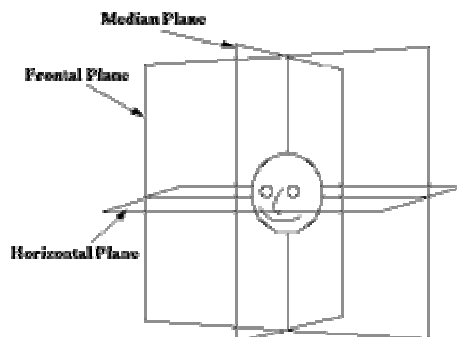
Los tres ejes coordenados definen tres planos estándar x-y el plano horizontal, x-z el plano frontal, y el y-z el plano mediano. Definiendo, arriba/abajo, delante/atrás y derecha/izquierda respectivamente. Pero como la cabeza es como una esfera se utilizan coordenadas esféricas, donde tenemos azimut, elevación y radio.





Los dos últimos dibujos de abajo, son los dos sistemas de coordenadas que se emplean:

1. vertical-polar coordinate system. Este se utiliza para describir las fuentes que están en el plano horizontal, cuando azimuth varía entre $(-180^\circ, +180^\circ)$
2. Interaural-polar coordinate system. Con este sistema el azimuth esta siempre entre -90° y $+90^\circ$. La distinción entre adelante y atrás, es 0 para superficies en el plano horizontal y 180 para las superficies en la parte de atrás del plano horizontal.



Por tanto, la localización se realiza a partir de la determinación de una *dirección* y una *distancia*.

3.2. DIRECCIÓN

La dirección de una fuente sonora, a su vez, se establece a partir de un *ángulo lateral* de un *ángulo de elevación*.

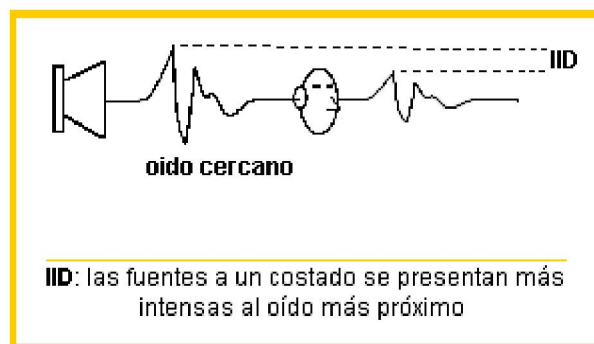
3.2.1. LOCALIZACIÓN EN EL PLANO HORIZONTAL – Ángulo Lateral

Para la ubicación lateral de una fuente sonora el sistema auditivo utiliza pistas provenientes principalmente de las diferencias de intensidad y tiempo con que las ondas sonoras llegan a cada uno de nuestros oídos.

Unas y otras son más efectivas para distintos rangos de frecuencia.

DII - Diferencias Interaurales de Intensidad

Es el tiempo que tarda de más un sonido proveniente del lado izquierdo de una persona en llegar al oído derecho. Esto se debe a que el sonido se difracta alrededor de la cabeza hasta llegar al oído derecho, y por lo tanto tienen que viajar más.

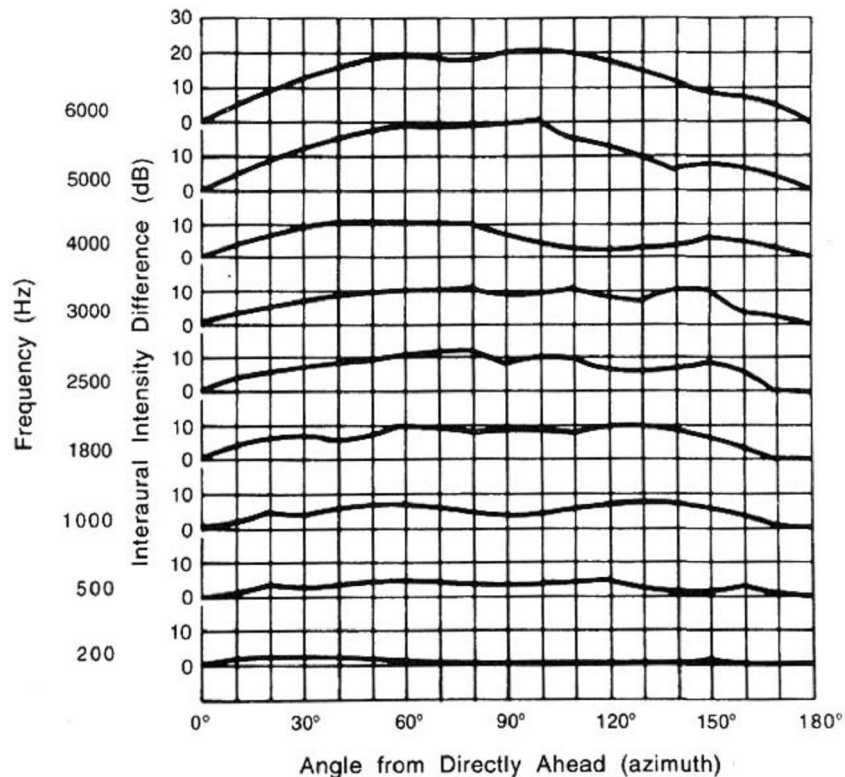


Las DII se dan principalmente a partir de las diferentes distancias que deben recorrer las ondas para llegar a uno y otro oído, pero también por la sombra acústica producida por la cabeza del individuo (difracción de la onda).

Los sonidos de bajas frecuencias tienen longitudes de onda relativamente grandes con respecto a las dimensiones de la cabeza. El estudio de la difracción determina que cuando la longitud de la onda es suficientemente grande con respecto al obstáculo que encuentra la onda, ésta se difracta fácilmente y no se genera una "sombra acústica" (o, al menos, se produce una sombra acústica pequeña). Por el contrario, cuando las longitudes de onda son pequeñas, se produce poca difracción y por lo tanto existe una sombra acústica mayor.

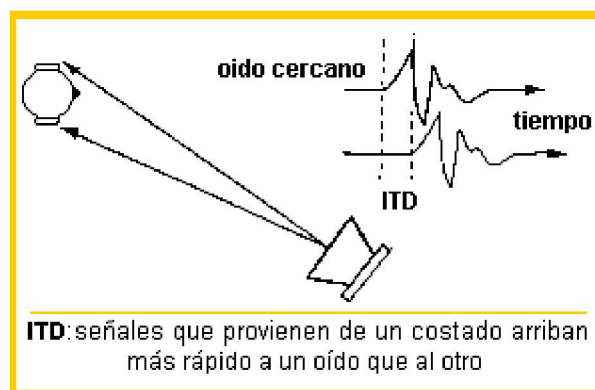
Para frecuencias de 500 Hz la longitud de onda del sonido es de unos 69cm, unas tres veces el diámetro promedio de la cabeza humana. La difracción es poca. Para frecuencias de 4 KHz (la longitud de onda del orden de los 8.5 cm) la sombra acústica es importante.

Las DII son prácticamente despreciables para frecuencias inferiores a los 500 Hz, pero pueden ser de hasta 20 dB para frecuencias mayores de 5 KHz.

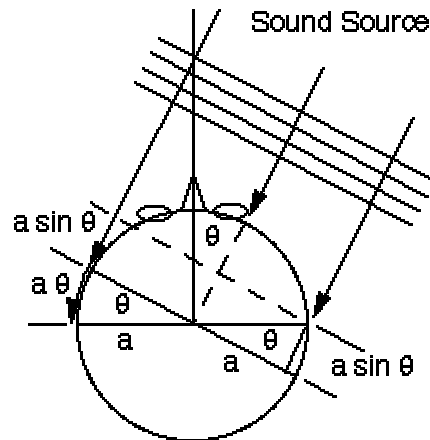


DIT - Diferencias Interaurales de Tiempo

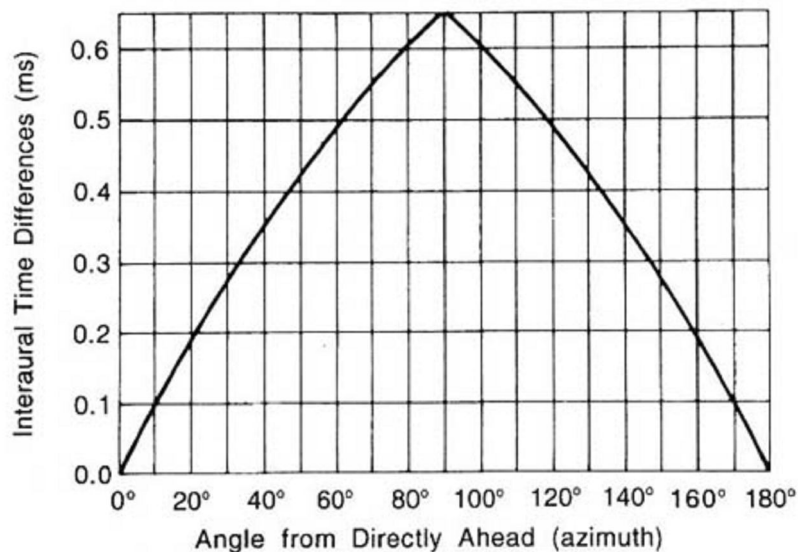
Este parámetro se refiere a la diferencia de tiempos de llegada entre los oídos. Mediante este parámetro se puede localizar una fuente de sonido. Este efecto es útil hasta una frecuencia en la que la longitud de onda del sonido se aproxima al doble de la distancia entre los dos oídos, a partir del cual, no se diferencia un sonido de otro.



Las DIT pueden calcularse a partir de las diferencias en las distancias que deben recorrer las ondas.



DIT van de 0 s para fuentes sonoras con un ángulo de 0° (exactamente delante del sujeto), hasta cerca de 0.69 ms para fuentes sonoras con un ángulo de 90° .



Según Raleigh: el sonido se propaga a una velocidad de “c” unos 343 m/s. Si una onda llega a una superficie esférica con un ángulo azimuthal, el sonido llega primero a la oreja derecha que a la izquierda, pues tiene que viajar una distancia superior para llegar a la oreja izquierda. Si lo dividimos por la velocidad del sonido, se obtiene la fórmula:

$$ITD = a/c (\theta + \sin \theta) ; -90^\circ \leq \theta \leq +90^\circ$$

Así, ITD es cero cuando la fuente está directamente en frente, y tiene un máximo cuando la fuente está en un lado. Esto representa una diferencia de 0.7ms para un tamaño típico de cabeza.

3.2.2. LOCALIZACIÓN EN EL PLANO MEDIO – Ángulo de elevación

Experimentos han mostrado la importancia de la cabeza, pero también del pabellón auditivo en la localización de sonidos que se encuentran en el plano medio. En tal caso las DII y las DIT son cero, dado que ambas ondas llegan al mismo tiempo y con la misma intensidad a los oídos.

El registro de los sonidos para estos experimentos se realizan por medio de micrófonos colocados dentro de los pabellones artificiales de un muñeco o dummy.



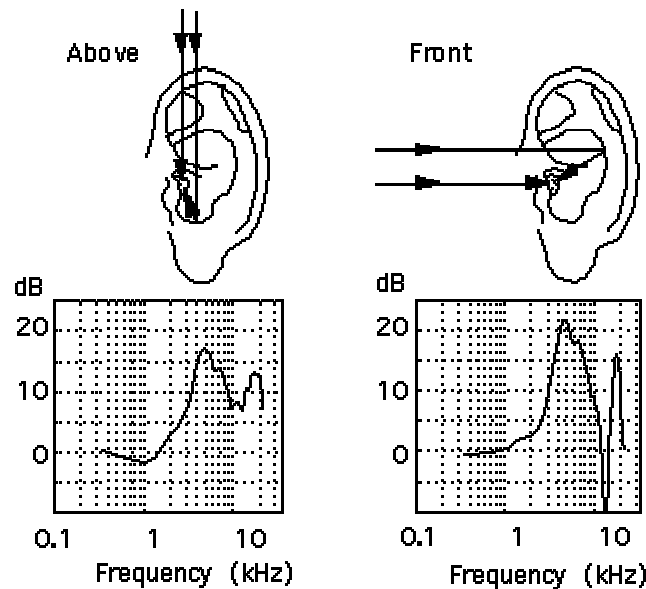
Tanto la cabeza, pero principalmente el pabellón auditivo, modifican el espectro de los sonidos en dependencia del ángulo de incidencia del sonido con respecto a la cabeza. Las diferencias espectrales entre el sonido original y el sonido medido junto al tímpano dieron lugar a las HRTF (Head-Related transfer Function) o funciones de transferencia relativas a la cabeza.

Las modificaciones espectrales producidas por el pabellón y la cabeza también son usadas por el sistema auditivo para determinar la localización de una fuente sonora. En este caso, es importante que el sonido tenga energía espectral a lo largo de un amplio rango de frecuencias. Las frecuencias superiores a los 6 KHz son particularmente importantes, dado que es en esa región en la que las longitudes de onda se hacen suficientemente pequeñas como para interactuar eficazmente con el pabellón.

Los distintos picos de resonancia en las HRTF corresponden a diferentes localizaciones de las fuentes sonoras en el plano medio.

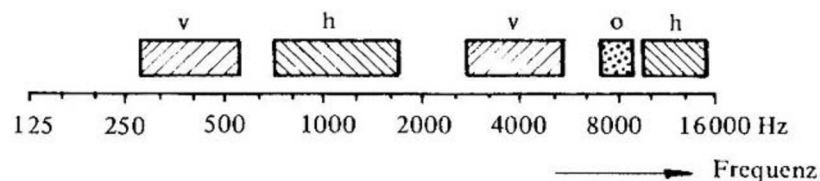


El pabellón auditivo actúa como una cavidad de resonancia, donde amplifica algunas frecuencias, y su geometría genera ciertos efectos de interferencia que atenúan otras frecuencias. Dicho de otra forma, su respuesta en frecuencia depende de la dirección. La figura de abajo mide las respuestas en frecuencia de dos direcciones distintas. En ambos casos vemos los caminos que siguen las ondas.



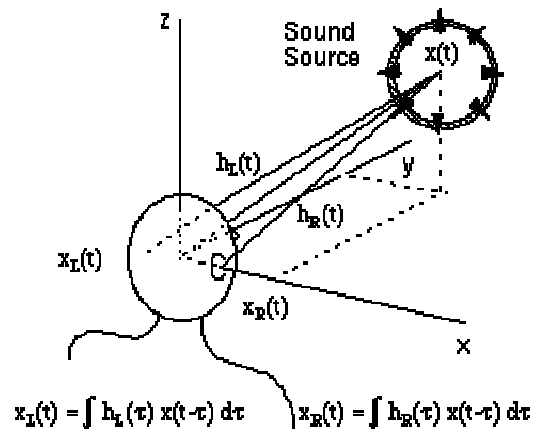
A frecuencias elevadas, la señal directa está en desfase con la señal que llega más tarde. La mayor interferencia ocurre, cuando la diferencia de recorrido d es la mitad de la longitud de onda. Por ejemplo si $F=c/2D$, como el pabellón auditivo es mucho más efectivo con sonidos que vienen de la parte delantera, este efecto es por lo tanto mucho más pronunciado.

Si se presenta un sonido de banda limitada con frecuencias centrales de 300 Hz o 3 KHz la imagen sonora siempre se formará delante del sujeto. Si la frecuencia central es de 8 KHz la imagen estará siempre arriba. Y si la frecuencia central es de 1 o 10 kHz la imagen se formará siempre detrás.



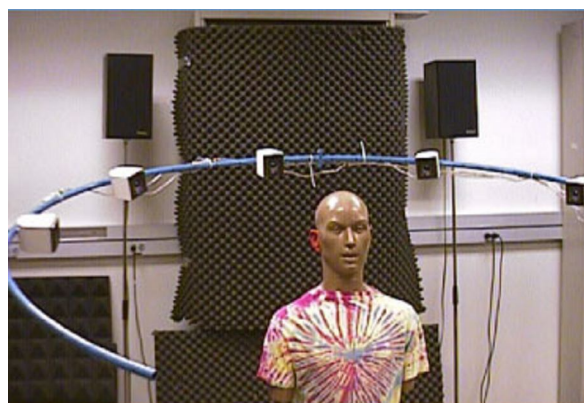
Donde:
V=adelante
O= arriba
H=atrás

3.2.2.1. HRTF - HEAD-RELATED TRANSFER FUNCTION



Este dibujo es el que se utiliza para encontrar la presión sonora en alguno de nuestros oídos. Para encontrar la presión sonora que una señal $x(t)$ produce en el pabellón auditivo, es necesario conocer la respuesta al impulso $h(t)$ de la fuente en el pabellón auditivo. Esto se llama Head Related Impulse Response (HRIR), y su transformada de Fourier se llama Head Related Transfer Function (HRTF). La HRTF captura muchas cualidades que utilizamos a la hora de localizar un sonido. Una vez conocido para ambos oídos, se puede sintetizar las señales binaurales a partir de una señal monoaural.

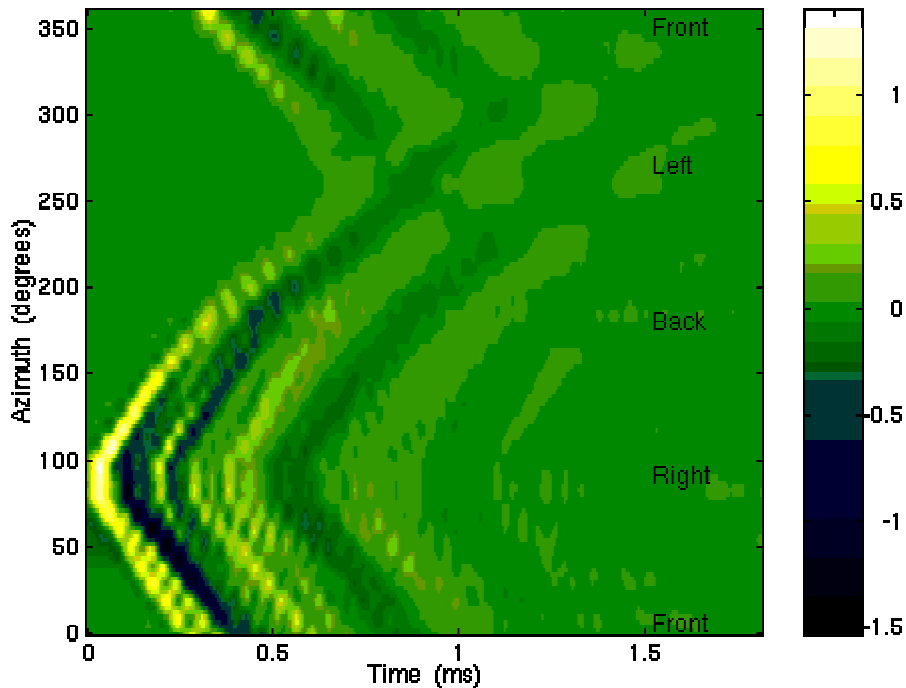
Esta función es bastante complicada y depende de 4 variables, tres espaciales y una frecuencial. Para distancias mayores de un metro, se dice que la fuente está en el campo lejano de audición y la HRTF cae con el inverso de la distancia. La mayoría de las mediciones del HRTF se miden en el campo lejano, ya que así dependerá fundamentalmente de azimut, elevación y la frecuencia.



Hay varias maneras de hacer estas mediciones, la más común, como se ha comentado antes, es usar un maniquí al que se le acoplan un micrófono en cada oreja y se realizan todo tipo de medidas binaurales en cualquier situación

y condiciones reproduciendo la misma sensación que captaría una persona humana.

Para poder hacerse una idea de cómo son estas funciones, se presenta un ejemplo de HRIR, en el plano horizontal.



El gráfico muestra la respuesta de la oreja derecha a un impulso en el plano horizontal. La fuerza de la respuesta se representa más brillante, así se puede apreciar como llega antes cuando viene de la derecha (azimut 90°). Y más débil cuando viene de la izquierda (azimut 270°). Nos damos cuenta que el tiempo de llegada varía más o menos como una forma sinusoidal. Este tiempo de llegada se aproxima muy bien al ITD. Se puede apreciar también, que la diferencia entre el tiempo más corto y largo de llegada es aproximadamente 0.7 ms, como se decía en la fórmula de ITD.

$$\text{ITD} = a/c (\theta + \sin \theta) ; -90^\circ \leq \theta \leq +90^\circ$$

Las interpretaciones que se pueden sacar además de las comentadas son:

Las secuencias iniciales de cambios rápidos (las bandas oscuras y brillantes) tiene que ver con las reflexiones. Así la primera reflexión es del pabellón auditivo, y la que está en 0.4 ms viene del hombro.

La respuesta cuando está enfrente o detrás es muy parecida. Las diferencias vienen de una falta de simetría en la línea de los 90°. Estas son por ejemplo la falta de la línea oscura que sigue el primer pulso, que existe en la parte

delantera, pero no en la trasera. Esto se debe a la reflexión del pabellón auditivo, que esta adecuada para los sonidos de frente.

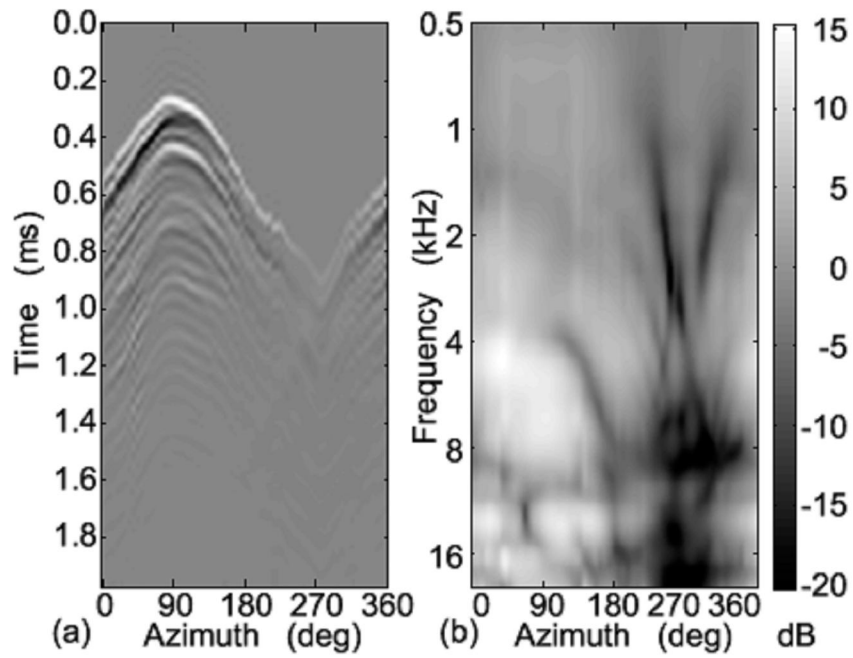
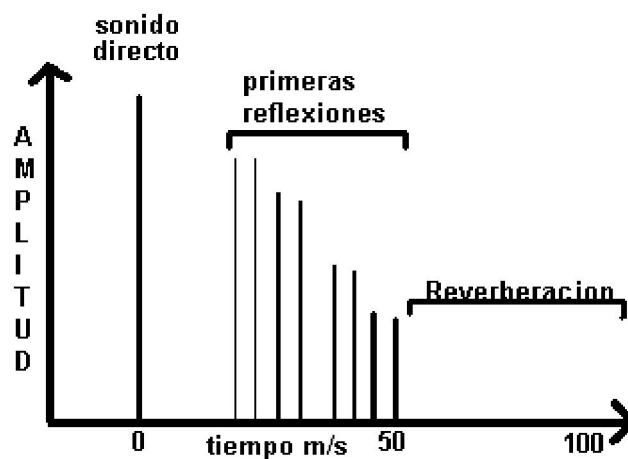


Figure 5: *Horizontal plane (a) HRIR (b) HRTF*

3.3. DISTANCIAS

Este parámetro es uno de los más difíciles de calcular. Hay ciertos parámetros que se pueden tener en cuenta. Entre los parámetros están:

- **Loudness:** Intensidad de sonido
- **Motion parallax:** esto se refiere a los movimientos de la cabeza. El mover la cabeza cambian los parámetros de azimut, y estos son dependientes de lo lejos que esté de la fuente. Así para fuentes sonoras que están cercanas al receptor, se puede observar un cambio grande en las variables azimut, mientras que para fuentes lejanas, es inapreciable. Así mismo, para fuentes cercanas, el ILD aumentará.
- **Reverberación:** otra característica que el oído utiliza para localizar los sonidos es la razón de reverberación con la fuente el sonido directo. Esto tiene que ver con la reflexiones del sonido en las paredes de una habitación. la intensidad del sonido cae con el cuadrado de la distancia, pero cuando existen reverberaciones las cosas cambian. A distancias cortas la reverberación es bastante grande, mientras que a distancias largas este ratio es más pequeño.

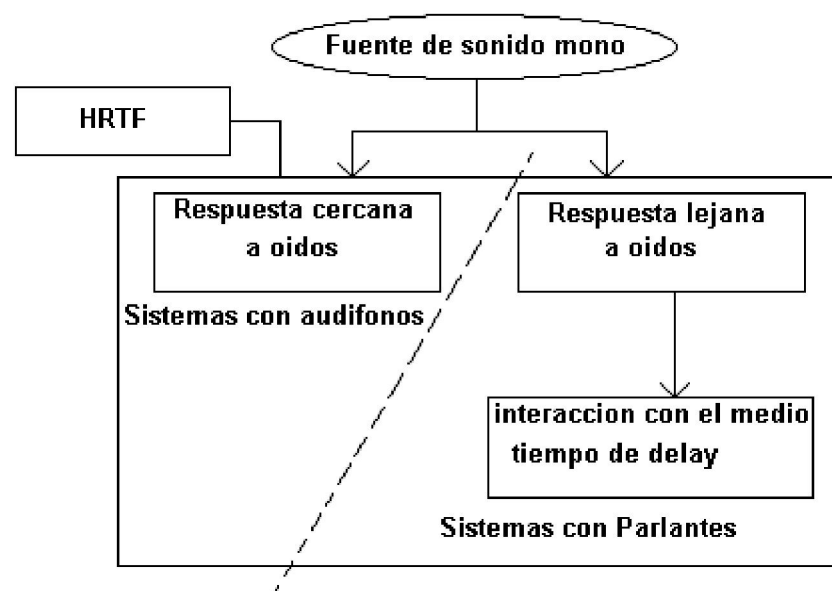


- **Proporcionalidad entre reverberación y sonido seco:** En habitaciones cerradas, la energía de una densa reverberación será más o menos constante, mientras que en distancias largas la energía que un sonido directo (seco, sin reverberación) caerá con la distancia.
- **Balance espectral:** las habitaciones actúan como filtros, modificando el espectro de las señales acústicas. El espectro de la respuesta al impulso de una habitación buena, debe ser lo suficientemente plana en todo el ancho de banda de la señal.
- **Difusión:** tiene que ver con la absorción de los materiales, normalmente las frecuencias altas caen más rápidamente

4. SISTEMA DE SONIDO 3D

Estos sistemas se basan en las funciones de transferencia HRTF, son capaces de simular elevación, distancia y azimut. En principio mediante este sistema se puede crear sonido 3D, pero en la práctica es algo distinto, debido a las diferencias entre las personas y las limitaciones computacionales. Aún así, los sistemas basados en HRTF son un estándar para los sistemas avanzados de 3D.

El esquema siguiente muestra el proceso para crear un sonido ambiental partiendo de un sonido mono.



4.1. CIRCUITO HRTF

Veamos ahora un ejemplo de cómo se ha creado una tarjeta de sonido con las funciones HRTF.

La parte principal de la tarjeta de sonido son dos convertidores A/D (DACs) utilizando un algoritmo específico para reproducir los sonidos en el plano horizontal alrededor del oyente.

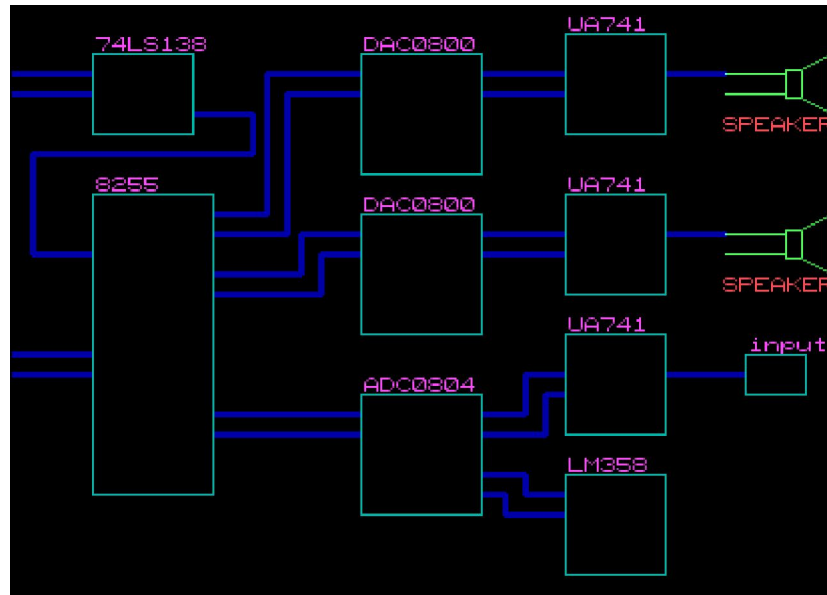
Los DACs que se utilizan en esta tarjeta, son un DAC 800 y otro DAC 803 los cuales tiene entradas y salidas de 8 bits. La salida de los DACs se lleva a los amplificadores operacionales, los cuales amplifican la señal para poder reproducir el sonido en los altavoces con la suficiente potencia.

Como dispositivo interfaz con el bus ISA del PC utilizamos el circuito integrado 8255, el cual es un PPI (Programable Peripheral Interface)

EL circuito 74LS138 es un decodificador/demultiplexor que se utiliza para activar el 8255.

Para convertir la señal de entrada utiliza un ADC 804.

Aquí se muestra un diagrama del circuito básico:



Veamos ahora el algoritmo que utiliza para posicionar las notas musicales en un entorno 3D:

```

Find the frequency representing the wanted note ;
Find the sound source angle deviation ;
If ( angle < 180 ) Then
Begin
    DAC1 := left_DAC;
    DAC2 := right_DAC
End
Else
Begin
    DAC1 := right_DAC;
    DAC2 := left_DAC
End;
If ( angle <> 180 ) Then
Begin
    For I := 1 To duration Do
    Begin
        turn DAC1 On;
        Delay ( angle_delay );
        turn DAC2 On;
        Delay ( note_delay - angle_delay );
        turn DAC1 Off;
        Delay ( angle_delay );
        turn DAC2 Off;
        Delay ( note_delay - angle_delay );
    End
End

```

La variable “angle” representa el ángulo requerido para posicionar la nota

La variable “duration” representa la longitud de la nota

La variable “note” representa la frecuencia de la nota

La variable “angle_delay” representa el retraso requerido antes de activar el segundo DAC.

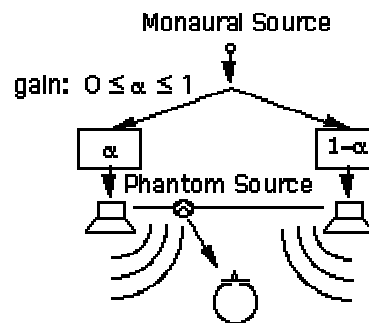
4.2. CREAR SONIDO 3D MEDIANTE GRABACIONES BINAURALES

Los sistemas de audio espaciales, se limitan a localizar el azimut. Hay tres clases de estos sistemas simples:

- Dos canales (estéreo)
- Multicanal
- Grabaciones binaurales

A. Sistemas de dos canales

En la industria de entretenimiento, el estéreo fue el primer producto comercial que daba la sensación de sonido 3D. La idea era muy simple, si se quería situar un sonido a la izquierda, se llevaba la señal al altavoz izquierdo, para situarlo al altavoz derecho. Aquí se muestra un gráfico explicativo.

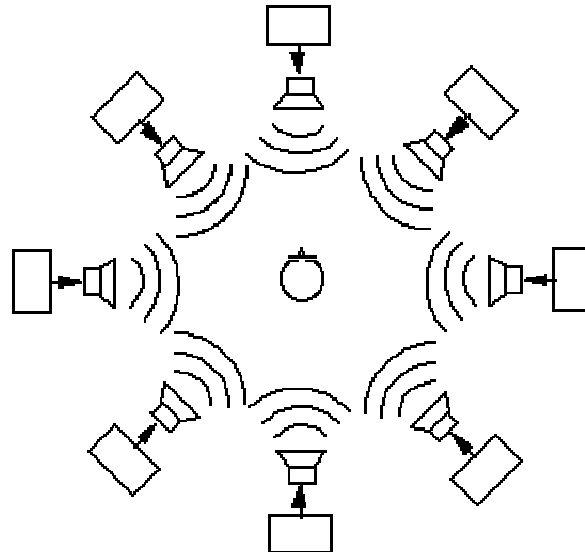


Así, la misma señal se envía tanto al altavoz izquierdo como al derecho, si los altavoces están en fase, y si el que escucha está entre los dos altavoces, entonces el oyente percibirá la fuente de sonido como si la fuente de sonido estuviera realmente delante suyo. Pero aún así no consigue nunca la impresión de que la fuente de sonido este fuera del segmento que une los dos altavoces.

B. Sistemas multicanal (Surround sound)

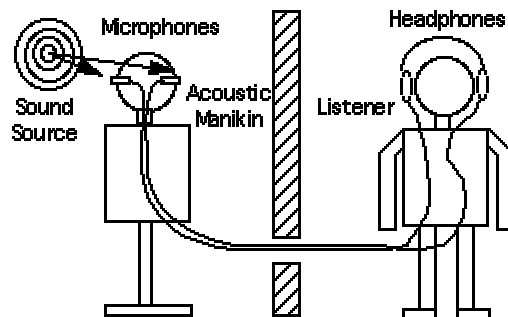
Otra forma de localizar los sonidos es separar los canales para cada una de las direcciones que deseamos, incluyendo arriba y abajo. Esto es lo que se utiliza en los sistemas de teatro, como Dolby Pro Logic Surround Sound.

Aunque producen efectos espaciales muy satisfactorios, los sistemas multicanal son caros e inconvenientes por el espacio que ocupan.



C. Grabaciones binaurales

El sistema anterior está bien si se quiere recrear sonido envolvente, pero existe otro sistema para crear sonido 3D mediante dos canales. El truco que utiliza es la recreación de las presiones de sonido en ambos oídos, como si el oyente se encontrara ahí mismo, en el lugar real.



Una aproximación que se podría hacer, consistiría en situar dos micrófonos en los canales auditivos de un maniquí, o incluso de dos maniqués situados cerca de tus orejas. Para grabar lo que nuestros oídos registran

De hecho, si el maniquí tiene la misma forma y tamaño de cabeza los mismos tiempos de ITD y ILD serán recibidos, así, si tuviera el mismo pabellón auditivo con la misma forma y tamaño, la información de elevación será recibida perfectamente.

A estas grabaciones se les llama grabaciones binaurales, y pueden producir casi un sonido 3D. De hecho se pueden usar este tipo de grabaciones en HCI (interacciones entre máquina hombre) para producir sonido en 3-D, con sistemas simples.

Prueba de grabación binaural:



Mediante este método se consigue grabar música para que posteriormente, mediante unos auriculares, se reproduzca la música en el oyente como si estuviese realmente en el lugar de grabación. Se consigue grabar tanto los efectos producidos en el plano horizontal, como del plano medio (DII, DIT, HRTF).

Para comprobar que efecto produce escuchar el sonido mediante este método de grabación, se puede bajar la música directamente de este link <http://www.sonicstudios.com/nobanjo.htm>

Otro ejemplo de grabación binaural:

En este caso la grabación se produce en una sala de orquesta



<http://www.sonicstudios.com/bigband.htm>

5. PRACTICAS

Para ver como se puede generar sonido envolvente mediante un par de altavoces, podemos recurrir a dos programas como, VSPACE y WDSPA. Tanto un programa como otro son capaces de convertir un sonido monoaural en un sonido envolvente. Para realizar la conversión se modificarán ciertos parámetros como, dimensión de la sala en la cual se va a reproducir el sonido, niveles de reverberación, retrasos o desfases entre los dos canales.

Analizaremos y comprobaremos como se generara un sonido 3D en el programa VSPACE.

5.1. VSPACE

VSpace realiza grabaciones de sonidos sobre espacios acústicos virtuales. El espacio acústico, la clase de micrófonos y los sonidos que se mueven dentro de él son generados a partir de un fichero de configuración.

VSpace es un programa autónomo que forma parte de MNLlib, un sistema de bibliotecas y de programas en C++ para el procesado de música de audio. MNLlib está diseñado para que sea portable tanto para Linux como para Windows.

VSpace se basa en un programa script. El cual se pueden configurar parámetros en un fichero para que el programa los procese. Este programa cuenta con una librería especial para la codificación de la señal de entrada (sonido monoaural) y convertirla en sonido ambiental.

Para obtener el ejecutable lo descargaremos desde este link:

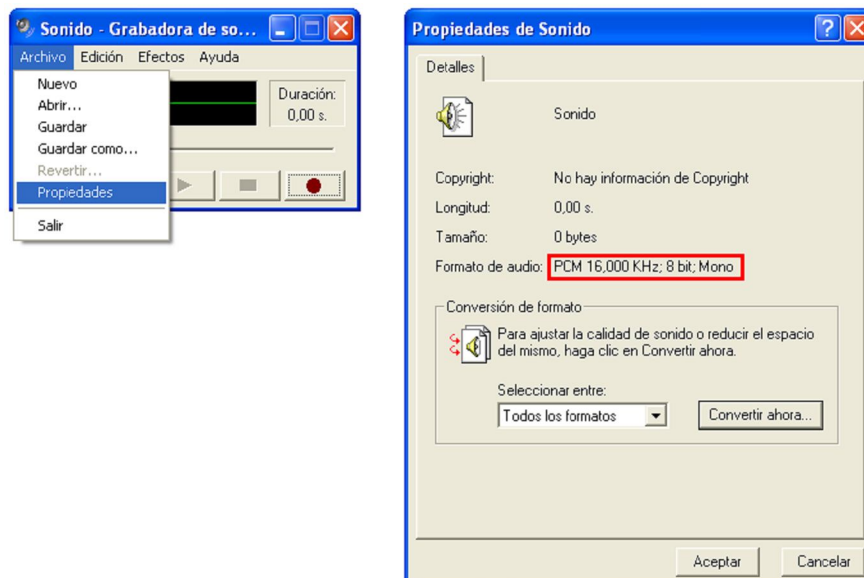
http://www.muse.demon.co.uk/mn_download.html

Para realizar el proceso de transformación de un sonido mono a un sonido 3D seguiremos los siguientes pasos:

1- **Sonido mono.** Primero necesitaremos un sonido mono (sonido.wav), el cual será procesado por el sistema. Este fichero podemos cogerlo directamente de nuestro ordenador, como un sonido de muestra, o podemos generarlos directamente nosotros.

2- **Grabación de un sonido mono.**

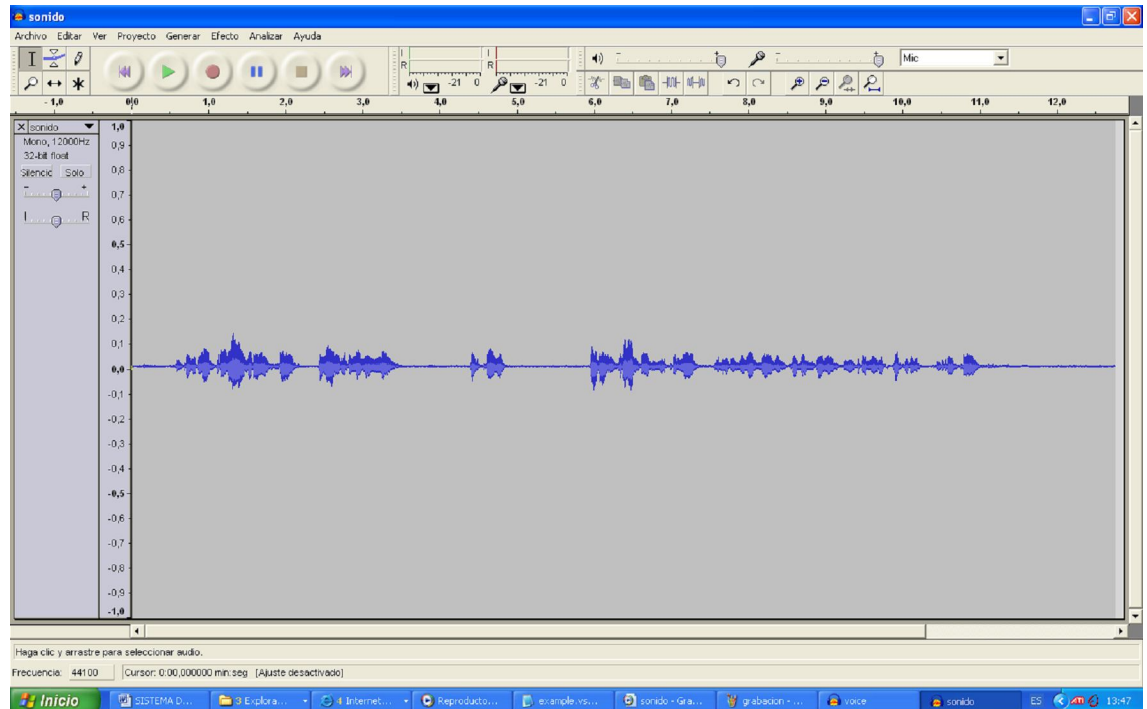
Para obtener “sonido.wav” utilizaremos la grabadora de sonidos de Windows. Por defecto, la grabadora graba en estereo, con lo que tendremos que configurarla en mono.



3- Modificar el audio.

Si la reproducción del archivo de sonido que hemos obtenido no es de calidad, se escucha bajo o tiene ruido de fondo, podemos tratarlo en el programa audacity.

El sonido que he obtenido en mi caso es el siguiente:

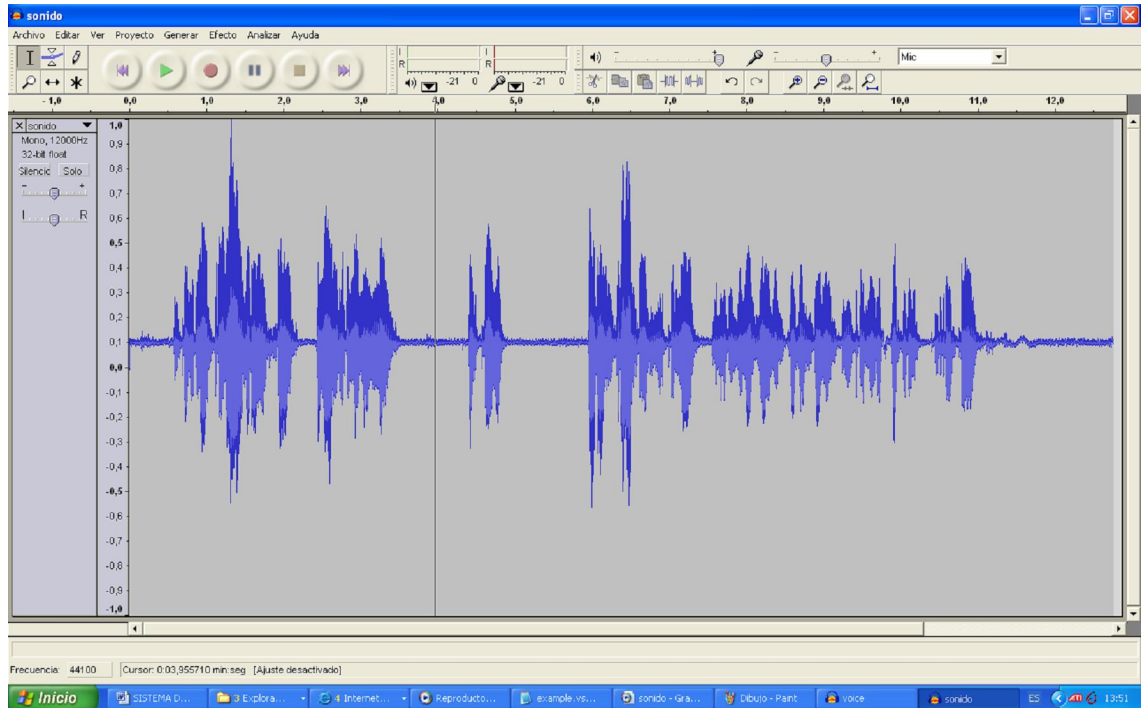


Como se puede observar, el sonido que se ha obtenido de la grabación se puede mejorar amplificando todo el sonido en 16,4dB.

Para amplificar la señal realizaremos lo siguiente:

Editar-> seleccionar todo

Efecto -> amplificar



4- Una vez obtenido el sonido modificado lo guardamos en el directorio donde se encuentre el comando ejecutable vspace.exe y nuestro script.

5- Script

Para crear el script partiremos de un script de ejemplo, en el cual se van comentando los pasos que se van realizando.

El código del script es el siguiente:

```
#####
#####
#
# Ejemplo Script
# -----
#
# General Settings
# -----
#
# At the top of the file general settings are placed.

# Set the early reflection time (see manual). This affects image
# quality and computation time. When you run the program, if you use
# the -v (verbose) flag then you will be told the number of early
# reflection room images are in use. Usually, computation time is
# approximately proportional to this value. For preview purposes early
# reflections can be turned off using the -e flag when you run the
# program. This reduces room images to one.
#
# After the early reflection time the late reflection engine takes
# over. This can be disabled using -l. Various general settings are
```

```
# available to control the late reflection engine.

early reflection time 0.6;

# Set the early reflection minimum gain (see manual). This affects
# image quality and computation time.

early reflection minimum gain 0.005;

# Set the tempo. Times are interpreted as beats relative to this tempo
# (in bpm).

tempo 60;

# Set the units we are to use below.

distance unit metre;

# There are a variety of other options available, particularly
# relating to late reflection reverb settings.
#
#####
#####
#
# Room Settings
# -----
#
# In this section we describe the room in which our recording will take
# place.

room {
    # Set the dimensions of the room. The room is assumed to
    # be box shaped with walls parallel to the X, Y and Z axes.
    # In order, we determine range of valid X, Y and Z values.
    # (e.g. -26 < X < 33.)
    #
    # Note that positive X is generally considered `forward',
    # positive Y is generally considered `left' and Z `up.'

    dimensions -26, 57, -35, 48, -1, 23;

    # Determine how reflective each wall is. A value of 1 indicates
    # that the wall is full reflective, 0 indicates total
    # absorption.

    reflections 0.7, 0.7, 0.7, 0.7, 0.3, 0.9;
}

#####
#####
```

```
#
# Recording Devices
# -----
#
# In this section we describe recording devices within the
# room. Normally only one is needed, however a number are included
# here to demonstrate the different types available. This increases
# calculation time hugely. Recordings are written to file.

# A recording device consisting of two coincident cardioid microphones
# at 90 degrees to each other. Output is written to file
# "stereo.wav". The microphone is located at the origin (<0,0,0>).
#
# Individual cardioid microphones are also available and they can be
# configured from cardioid to figure-of-eight or omnidirectional
# response or anywhere in between.

recording stereo.wav {

    device cardioid pair;

    # Specify the location of the two microphones (they don't have
    # to be at the origin and don't have to be coincident.

    location <0, 0, 0>, <0, 0, 0>;

    # Specify vectors giving the directions the microphones are
    # pointing in. They are pointing forward left and forward
    # right.

    direction <1, 1, 0>, <1, -1, 0>;
}

# A recording device consisting of a second-order Ambisonic
# microphone. The output file ('ambisonic.fmh') is written out in FMH
# format. These files can be played back of arrays of speakers using
# programs such as the AmbisonicPlayer available from
# http://www.muse.demon.co.uk/mn\_download.html. B-Format (first order)
# Ambisonics is also available.

recording ambisonic.fmh {

    device second order ambisonic;

    # Specify the location of the two microphone. Ambisonic
    # microphones are not directional.

    location <0, 0, 0>;
```

```

# Optionally, specify a core radius (see manual) to help deal
# with sounds that pass close to the microphone.

core radius 1;

}

# An omnidirectional microphone in a far corner of the room picking up
# a rather different balance of sound to the others.
#
# Note that if this microphone was just be preview purposes then a
# `simple' microphone could be used instead. These use much less
# computation time but do not use a delay/Doppler model. In this
# context early reflections should be switched off.

recording distant.wav {

    device omnidirectional;

    # Put the microphone far away (it must be in the room).

    location <50, 42, 1>;

    # Optionally, turn up the gain on the microphone to make the image
    # clearer.

    gain 5;

}

#####
#####
#
# Fixed Location Sound
# -----
#
# In this section a track is playing at a single point to the left of
# the listener and then jumps to the distant right. A `track' consists
# of a list of sounds to mix together and instructions on how this mix
# of sounds should be placed in space.

track {
    mix {

        # Read in the file `voice.wav' from disk. Mix in the
        # entire file. Start playing at time zero (the
        # beginning). This is a recording of the word `VSpace.'

        0 voice.wav;
    }
}

```

```

# Play the sound again after two beats. Apply a gain
# to make the sound louder to compensate for the
# distance (see below). The sound will be more
# reverberant for the microphones at <0, 0, 0>.

    2    voice.wav    gain 10;
}

motion {

    # Start out static, just to the left.

    0    fixed <0, 1, 0>;

    # Second time, jump to the distant right (slightly
    # in front).

    2    fixed <5, -20, 0>;
}
}

#####
#####
#
# Moving Sounds 1
# -----
#
# In this section two sounds are mixed together and moved from left to
# right in front of the origin.

track {
    mix {

        # Read in ten seconds of sound from the file
        # `csaudio.wav', playing for 10 beats (10 seconds at this
        # tempo), 5 seconds into the performance. Apply a gain
        # to make it louder.

        5    csaudio.wav    gain 1.5 from 0 to 10;

        # Say `VSpace' three times.

        5    voice.wav    gain 1.5;
        10   voice.wav    gain 1.5;
        15   voice.wav    gain 1.5;
    }

    motion {

        # Move the sound from left to right, passing a few metres

```

```

# in front.

5 to 15    line    <3, 10, 0> to <3, -10, 0>;

    }
}

#####
#####
#
# Moving Sounds 2
# -----
#
# In this section a sound rotate around the origin. Sounds become
# directional.

track {
  mix {
    18    csaudio.wav    from 0 to 10;
  }

  motion {

    # Rotate the sound around the origin starting ten
    # metres in front and moving around to the left in a
    # half second. A complete circle therefore takes two
    # seconds.

    18 to 28    arc    centre    <0, 0, 0>
                  start <6, 0, 0>
                  to    <0, 6, 0>
                  in    0.5;

  }

  direction {

    # Start out with the usual (omnidirectional) profile,
    # but switch to being highly directional towards the
    # left. This means the sound is louder when it is to
    # the right of the microphone.

    23 to 23.5  line    <0, 0, 0> to <0, 1, 0>;

  }
}

#####
#####
#
# Lots of Sounds at Once

```

```

# -----
#
track {
  mix {
    30  csaudio.wav    from 15 to 20;
    33  csaudio.wav    from 10 to 20;
    40  csaudio.wav    from 10 to 15;
  }
  motion {

    # The `magnitude' keyword changes the length of a
    # vector without changing its direction.

    0    fixed <6, 4, 0> magnitude 4;
  }
}
track {
  mix {
    30  csaudio.wav    from 10 to 20;
    35  csaudio.wav    from 10 to 20;
    40  csaudio.wav    from 15 to 20;
  }
  motion {
    0    fixed <-7, -1, 0> magnitude 4;
  }
}
track {
  mix {
    34  csaudio.wav    from 10 to 20;
    30  csaudio.wav    from 17 to 20;
    38  csaudio.wav    from 10 to 17;
  }
  motion {
    0    fixed <2, -7, 0> magnitude 4;
  }
}

#####
#####
# eof

```


6- Ejecutar el script



```
C:\WINDOWS\system32\cmd.exe
C:\ambisonic>uspace.exe ejemplo2.usp
Seconds of audio completed: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16.
C:\ambisonic>
```

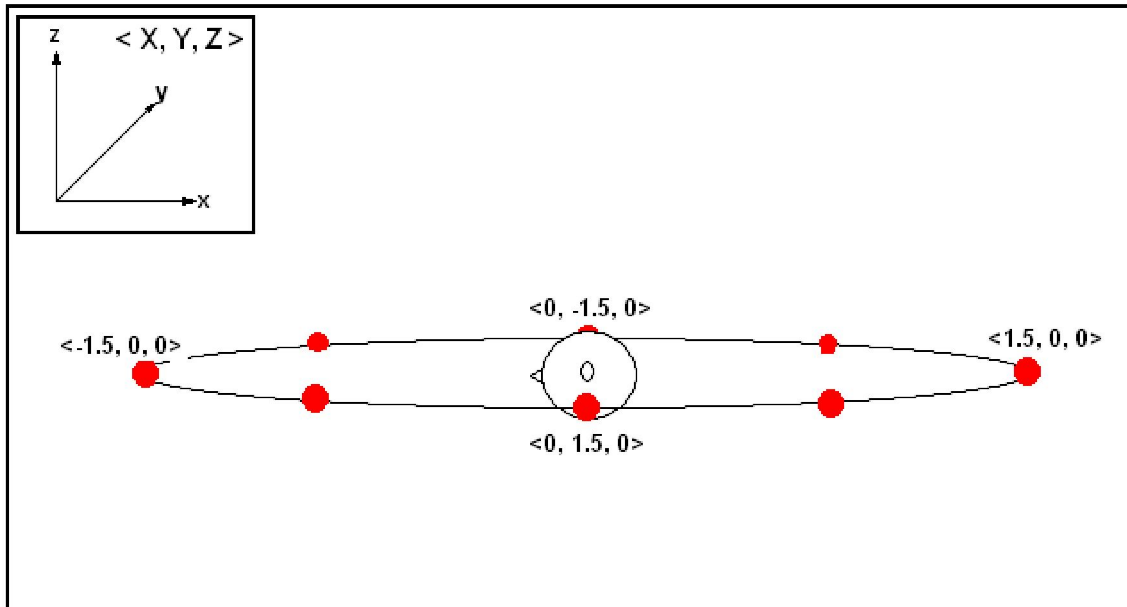
Después de ejecutar el script, muestra por pantalla la duración de la pista. En nuestro caso dura 16 sg.

Para ver la potencia de este programa, configuraremos diferentes escenarios para comprobar los diferentes efectos que se pueden conseguir. Realmente, Cuantos mas efectos introduzcamos puede ser una carga para el ordenador.

ESCENARIO 1:

El sonido se mueve alrededor del oyente.

Crearemos a partir de un sonido grabado en mono una imagen en el receptor, que simule que el sonido se traslada alrededor de él.
En este caso utilizaré un tono de una frecuencia de 440 Hz.



Configuración de la habitación donde se va a reproducir el sonido.

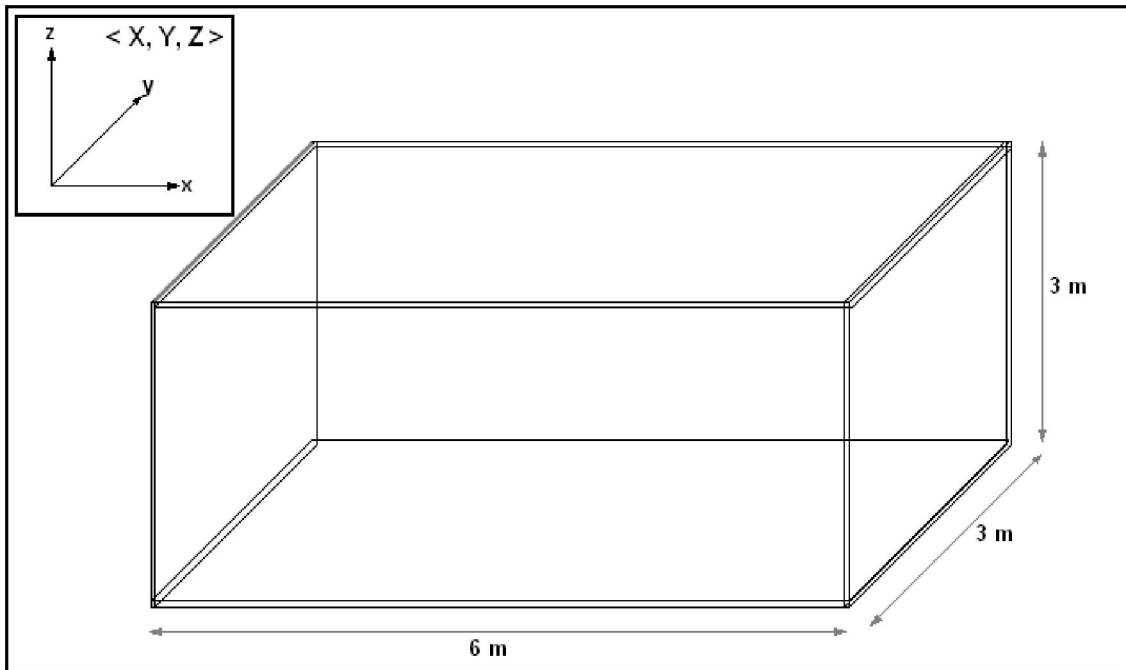
```
#####
#####
#
# Room Settings
# -----
room {

    # dimensiones de la habitación donde se realiza la grabación
    dimensions -3, 3, -1.5, 1.5, -1.5, 1.5;

    # coeficiente de reflexión para cada una de las paredes
    reflections 0.7, 0.7, 0.7, 0.7, 0.2, 0.3;
}

#####
#####
```

Representación en un plano de la habitación virtual:



Vamos a simular que los micrófonos se sitúan a cada lado de los oídos. Con esto conseguiremos mayor realismo.

```
#####
#####
#
# Recording Devices
# -----
recording escenario.wav {

    device cardioid pair;

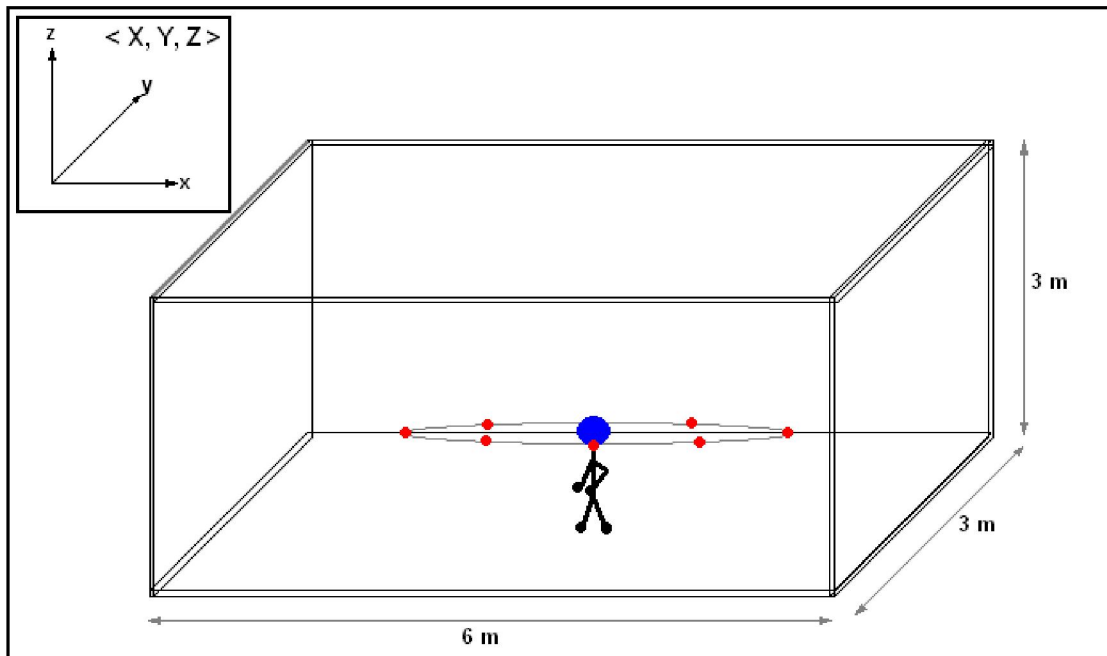
    # con esta configuración estamos poniendo que la distancia entre
    # micrófonos es de 8cm, mas o menos la distancia entre nuestros oídos.

    location <0, 0.04, 0>, <0, -0.04, 0>;

    direction <0, 0.04, 0>, <0, -0.04, 0>;

}
```

Ahora simularemos que el sonido se reproduce alrededor del oyente, como muestra la figura:



Código para localizar el sonido:

```
#####
#
# Fixed Location Sound
# -----

track {
  mix {
    0    tono440.wav;
  }

  motion {

    # limitamos la duración del sonido a 10 sg. Mediante esta función
    # haremos que la fuente de sonido, viaje alrededor del oyente.

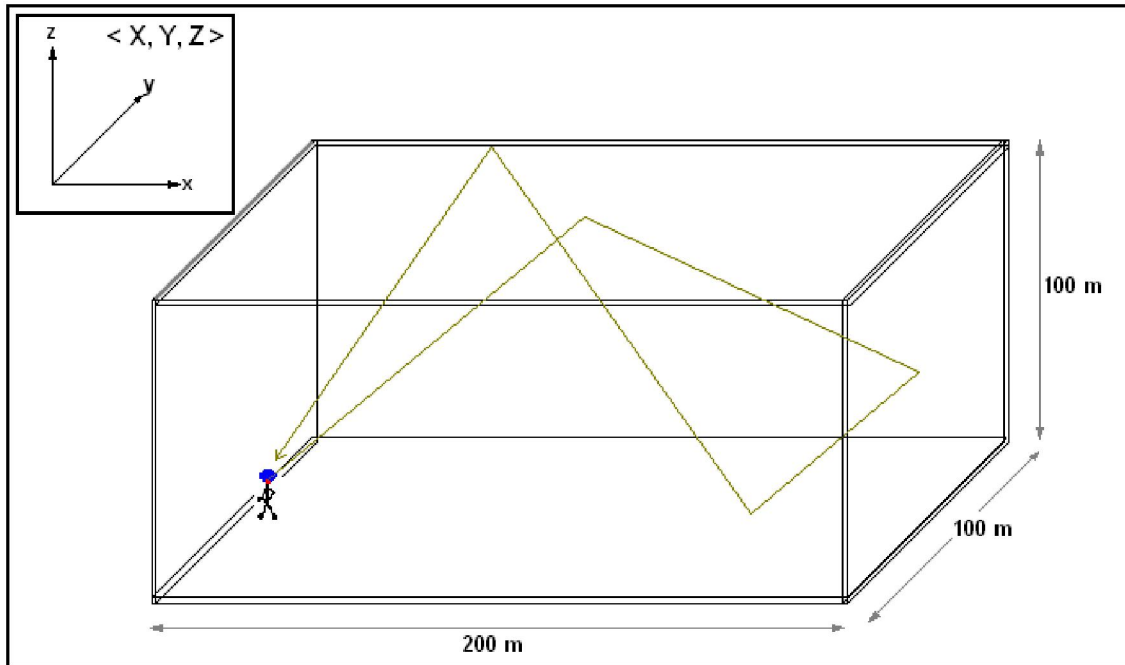
    0 to 10    arc    centre    <0,0,0>
                start <0,1.5,0>
                to <1.5,0,0>
                in 0.9;

#####
# eof
```

ESCENARIO 2

Recepción de nuestra propia voz por el efecto eco

Primero, vamos a crear una habitación virtual que pueda producirse este efecto. Configuraremos una sala de 200x100x100 con unos coeficientes de reflexión altos.



Código de nuestro script:

```
#####
#####
#
# General Settings
#
early reflection time 0.6;
early reflection minimum gain 0.005;
tempo 60;
distance unit metre;
#####
#####
#
# Room Settings
# -----
#
```

```
# In this section we describe the room in which our recording will take
# place.

room {

    dimensions -100, 100, -100, 100, -100, 100;

    reflections 0.9, 0.9, 0.9, 0.9, 0.9, 0.9;
}

#####
#####
#
# Recording Devices
# -----

recording eco.fmh {

    device second order ambisonic;

    location <0, 0, 0>;

    core radius 1;
}

recording eco.wav {

    device cardioid pair;

    location <-98, 0.04, 0>, <-98, -0.04, 0>;

    direction <0, 0.04, 0>, <0, -0.04, 0>;
}

#####
#####
#
# Fixed Location Sound

track {
    mix {

        0 hola.wav;

    }

    motion {
```

```
0    fixed <-100, 0, 0> magnitude 10;  
    }  
}  
  
#####  
# eof
```

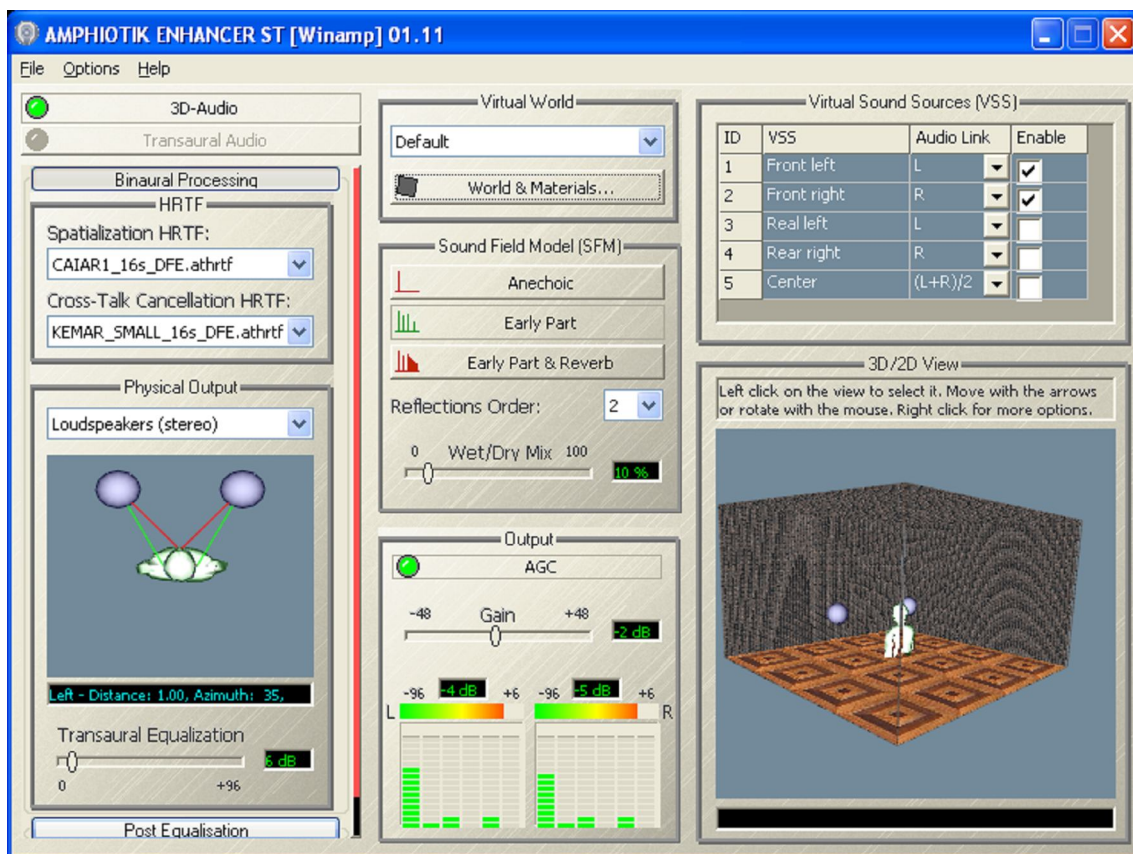
5.2. PROGRAMA WINAMP CON EL PLUGIN AMPHIOTIK ENHACER ST

Otro programa para crear sonido 3D en un ambiente virtual es este plugin para Winamp.

Mediante esta plugin, podemos modificar el sonido de reproducción en un sonido que se adecue a un entorno virtual que nosotros configuremos.

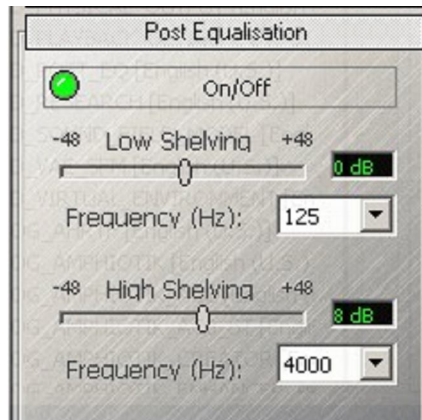
Amphiotik Enhacer es un potente sistema de audio 3D, el cual procesa un sonido estereo y lo convierte según las características de entorno, posicionamiento del oyente en sonido ambiental.

El aspecto que muestra el programa es el siguiente:

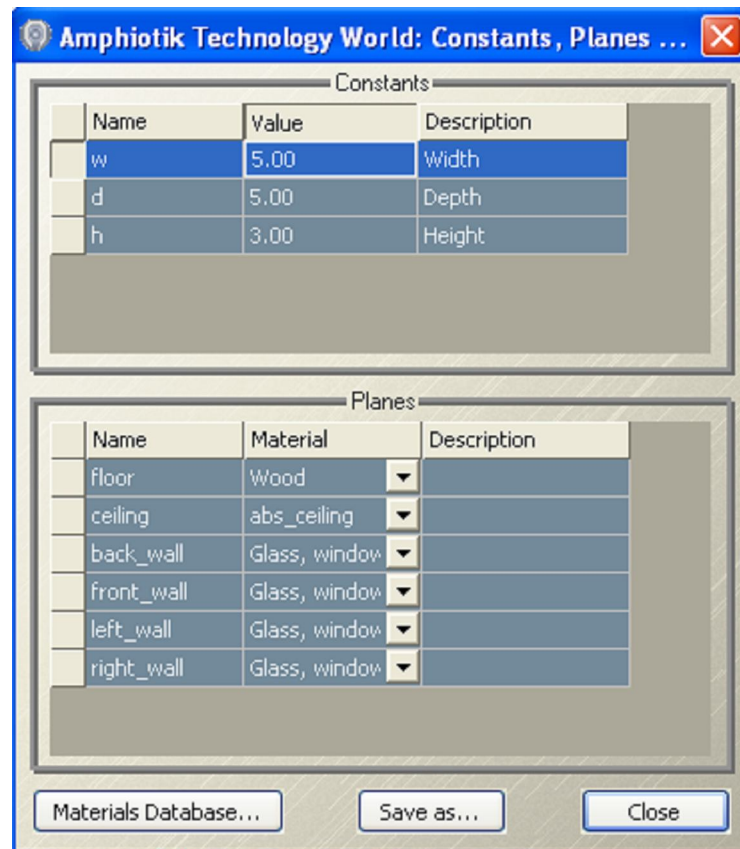


Vamos a comentar cada una de las opciones que nos ofrece este programa:

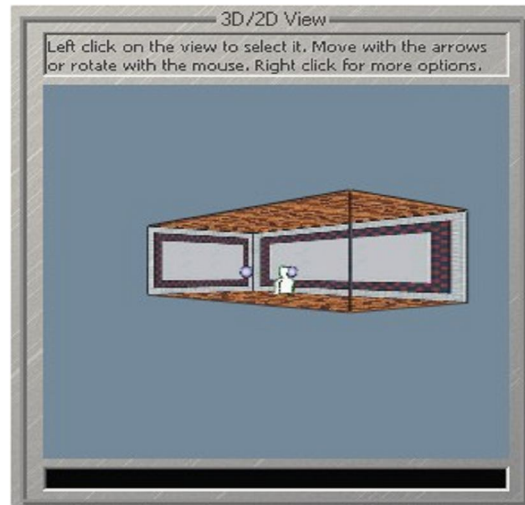
- **Binaural Processing (HRTF).** Función de transferencia HRTF, este parámetro es ofrecido por los laboratorios de sonido especializados, los cuales han conseguido recopilar en los archivos *.athrtf las funciones de transferencia de cada uno de sus muñecos de simulación.
- **Post Equalization.** El módulo ecualizador soporta dos tipos de filtros: Low Shelving y High Shelving.



- Physical Output (posición de los altavoces).** Aquí se tiene que configurar primero, que tipo de altavoces vamos a utilizar para reproducir el sonido. Las dos opciones que ofrece son, un par de altavoces o auriculares.
 En el caso de seleccionar los altavoces, tenemos la posibilidad de cambiar la distancia y ángulo (azimut) respecto el oyente.
- Virtual World.** Aquí configuraremos las dimensiones de la sala en la que se va a encontrar el oyente, así como del material que están fabricados cada una de las paredes.

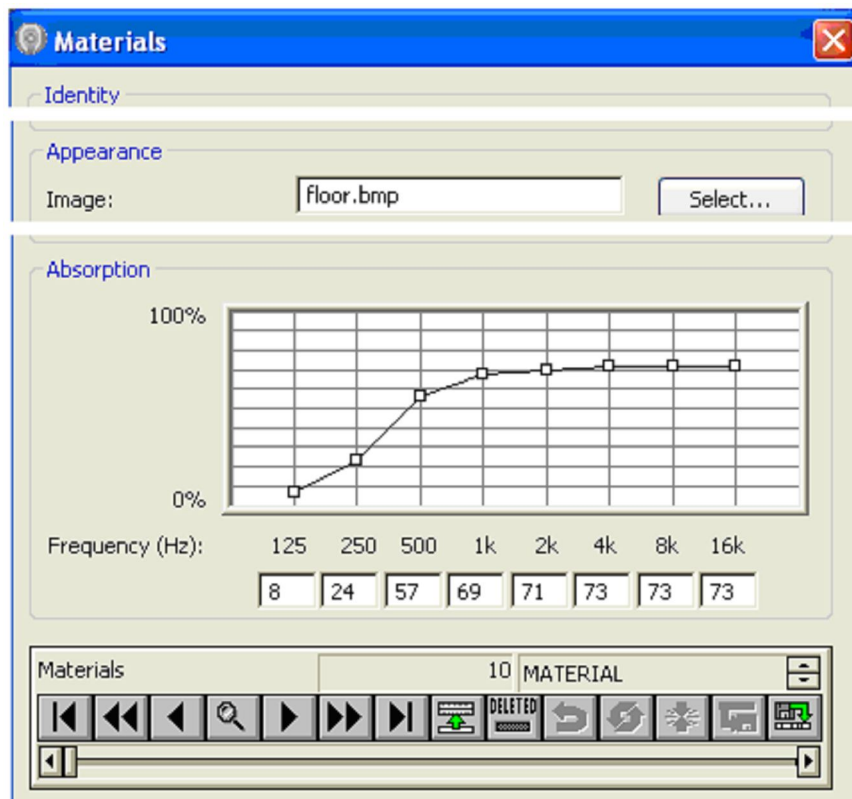


Pantalla de configuración



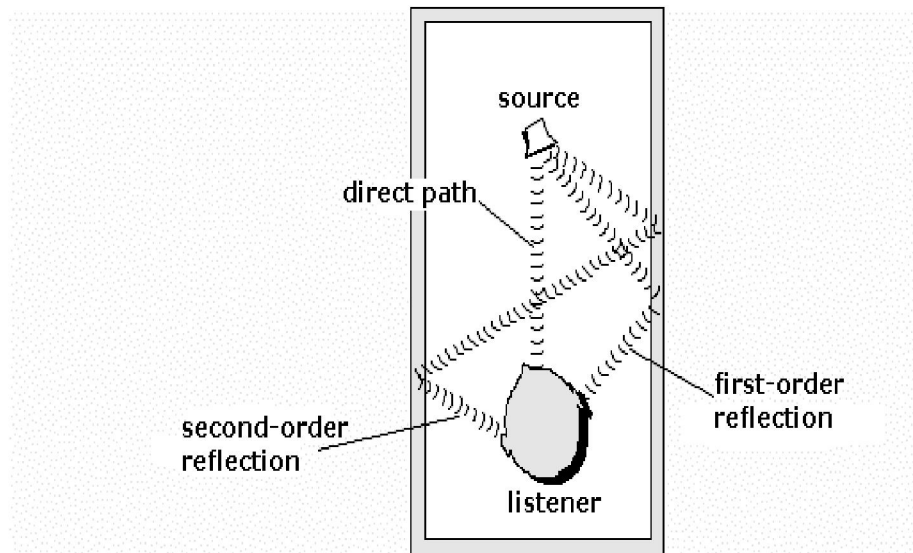
Vista de las dimensiones de la sala

También nos ofrece la posibilidad de crear nosotros mismos un material. Para ello, como podemos ver en la figura siguiente tendremos que graduar el nivel de absorción a cada una de las frecuencias. Para crear un tipo de estructura podemos mirar en revistas especializadas de sonido, para ver los valores de absorción de los diferentes materiales y trasladarlo aquí.



- **Sound Field Model (SFM).** Los SFM que pueden aplicarse en un auditorio virtual son estos:
 - *Anechoic.* El anechoic SFM asume que las paredes de la sala virtual son totalmente absorbentes, consecuentemente no existirán reflexiones.

- *Early Part*. La early part SFM calcula la primera reflexión en la sala virtual. El orden de reflexión se puede ajustar mediante el campo "Reflection Order".



- *Early Part & Reverb*. Lo mismo que en el apartado anterior, pero aquí se incluye la función de reverberación. El parámetro Wet/Dry mostrado debajo ajusta la reverberación de la sala. El valor recomendado se encuentra entre el 5% y el 20%.

Tiempos óptimos de reverberación en diferentes salas:

Los tiempos óptimos de reverberación dependen de la función del recinto. Un salón de clase requerirá un tiempo de reverberación corto (por ejemplo, $T=0.7-0.8$ s), mientras que una sala destinada a la actividad musical requerirá T mayores.

T óptimo dependerá también del género musical. La música de cámara requiere T menores (quizás $T=1.25-1.5$ s), mientras que la música de orquesta exige T mayores (hasta $T=2.5$ s). La música de órgano necesita T mucho mayores ($T=3-4$ s).

6. PRACTICAS A DESARROLLAR

Una práctica que no entra en este trabajo, pero que se podría proponer a otros alumnos como una práctica muy interesante, sería utilizar el programa ODEON para realizar simulaciones acústicas sobre salas virtuales.

El programa ODEON es un sistema de simulación acústica que facilita el diseño de salas. A partir de una descripción de una sala, permite calcular múltiples parámetros acústicos a la hora de ver el comportamiento que tiene la sala sobre las posibles fuentes de sonido.

El único inconveniente que podemos encontrar, es que se tendría que trabajar sobre una versión demo, ya que se trata de una aplicación de pago.

Para ver un poco las posibilidades que ofrece este programa se van a detallar unas cuantas:

(Como se trata de una versión demo, las salas donde realizaremos las pruebas estarán limitadas a unas pocas).

- Ofrece la posibilidad de representar la sala de forma tridimensional. Con ello permite observar la geometría de la sala.
- Definir fuentes y receptores puntuales en el entorno de la sala
- Asignación de materiales a las superficies de la sala
- Cálculos estimados de tiempos de reverberación en la sala.
- Auralización, permite escuchar sonido en la sala virtual que hayamos creado.

Para ver un poco la pinta que tienen, aconsejo pasarse por la web de ODEON:
<http://helmholtz.oersted.dtu.dk/~odeon/>

7. REFERENCIAS

[1] Localización espacial de sonido
<http://www.sacom.org.ar/secciones/segunda/Sesiones%20Tem%C3%A1ticas/Di%20Liscia.htm>

[2] Audio virtual 3D para auriculares
<http://www.sensaura.com/whitepapers/pdfs/devpc007.pdf>

[3] Tutorial sonido 3D
http://www.tav.net/3d/docs_sonido3d/index.htm

[4] Ingeniería acústica (referencias)
<http://www.gts.tsc.uvigo.es/audio/iac/index.php>

HRTF Head – Related Transfer Function

[5] HRTF - Head Related Transfer Function
http://interface.cipic.ucdavis.edu/CIL_tutorial/3D_HRTF/3D_HRTF.htm

[6] HRTF
http://www.absoluteastronomy.com/encyclopedia/H/He/Head-related_transfer_function.htm

[7] Sistemas basados en HRTF
http://interface.cipic.ucdavis.edu/CIL_tutorial/3D_sys2/3D_sys2.htm

[8] Mediciones HRTF para el dummy KEMAR
<http://sound.media.mit.edu/KEMAR.html>

[9] Base de datos HRTF. Mediciones para 45 sujetos.
http://interface.cipic.ucdavis.edu/CIL_html/CIL_HRTF_database.htm

[10] Construcción de una tarjeta de sonido implementando las funciones HRTF.
<http://www.arabion.net/gproject.html>

ARTÍCULOS:

[11] Acústica, psicoacústica, Simulación acústica, acústica electrónica, sonido 3D en auriculares,
<http://www.headwize.com/tech/>

[12] Web especializada en sonido
<http://www.hispasonic.com/>

Grabaciones binaurales

[13] Grabaciones binaurales
http://interface.cipic.ucdavis.edu/CIL_tutorial/3D_sys1/binaural.htm

[14] Material para realizar grabaciones binaurales
<http://www.sonicstudios.com/>

[15] Canciones Grabadas con el sistema binaural
<http://www.sonicstudios.com/mp3.htm>

Programas para crear sonido 3D

[16] VSPACE
<http://www.muse.demon.co.uk/vspace/vspace.html>

[17] Soundmania 1.0
<http://www.radgametools.com/mssdown.htm>

[18] 3Dsoundstore
<http://www.3dsoundstore.com/www-3dsoundstore-com/product-qcreator.asp>

[19] maven3D
<http://maven3d.com/en/download.asp?PID=1>

Acústica de salas

[20] Modelado de Acústica de salas. Auralización, técnicas binaurales
<http://www.at.oersted.dtu.dk/~clc/>

[21] Programa que simula diferentes escenarios. ODEON room
<http://www.dat.dtu.dk/%7Eodeon/Sounds.htm>