

AUDIO ESPACIAL

AMBI-ROLL - HERRAMIENTA PARA LA ESPACIALIZACIÓN DE ARCHIVOS DE AUDIO

November 28, 2017

Alejandro Gómez
Santiago González
Esteban Henao
Juan Miguel Jaramillo
Sergio Mejía
Diego Murillo
Andrés Zapata

Contents

1	Introducción	3
2	Descripción del Software	3
2.1	Instalación	3
2.2	Especificaciones	3
2.3	Análisis	3
2.4	Diseño	4
2.5	Implementación	7
2.6	Validación	7
2.6.1	Pruebas de usuario	7
2.6.2	Resultados de la prueba de aceptación	9
3	Estructura de la herramienta	9
3.1	Módulos	10
3.1.1	Espacialización	10
3.1.2	Load HRTFs	11
3.1.3	PLAY/STOP	11
3.1.4	CANALES	11
3.1.5	DELAY SPEAKERS CONTROL	12
3.1.6	Visualización	13
3.1.7	LEVELS	13
3.1.8	MASTER	14

1 Introducción

La realidad virtual y la tecnología 3D en los últimos años ha experimentado un rápido crecimiento en las diferentes aplicaciones y en la aceptación del público que cada vez empieza a demandar más contenido implementando estas tecnologías, en videojuegos, reproducción de vídeo y audio. Ambi-Roll es un aplicativo desarrollado para espacializar ocho archivos de audio por medio de la técnica de espacialización Ambisonics. El aplicativo permite reproducir el contenido por medio de ocho altavoces o puede ser codificado para reproducirse de forma binaural. Igualmente, el aplicativo permite tener control de reverberación, distancia y cobertura de los diferentes archivos.

2 Descripción del Software

2.1 Instalación

El software no requiere instalación ya que se presenta como un aplicativo ejecutable que al abrirlo muestra la interfaz gráfica de usuario.

2.2 Especificaciones

- Lenguaje de programación: Max V7.
- Plataforma: Windows y Mac OSX.
- Denominación de la aplicación: Ambi-Roll
- Versión: 1.0.
- Base teórica: Audio espacial y Ambisonics.

2.3 Análisis

La aplicación Ambi-Roll permite al usuario especificar la procedencia de ocho archivos de audio en cualquier dirección en azimuth, además permite agregar reverberación a los archivos, cada uno con control de dry/wet. el usuario puede modificar la cobertura de cada uno de los archivos de audio a tres ángulos distintos, un control de ganancia y uno que especifica la distancia a la que se encuentra el archivo de audio del receptor, además cuenta con un modulo de retrasos (delays) para ajustar el sistema en un punto de equidistancia.

Table 1: Descripción y requerimientos del proyecto

Proyecto	Ambi-Roll
Descripción	El proyecto tiene como objetivo espacializar audio por medio de la técnica Ambisonics en un arreglo de 8 altavoces en 2 dimensiones usando un orden de codificación de armónicos esféricos de tercer grado. La espacialización se puede realizar tanto en altavoces como en binaural (audífonos), para realizar el "downmix" a binaural se utiliza la base de datos de respuestas al impulso de una dummy head KU100 de Neumann. La aplicación utiliza 16 respuestas al impulso posicionadas en los mismos ángulos que los altavoces, estas IR son convolucionadas por cada uno de los audios que ingresen al sistema. Además de esto la aplicación permite agregar diferentes efectos a los audios ingresados como spread (extensión), reverberación y control de distancia de cada uno de los sonidos.
Requerimientos	
R1	Debe ser un software tipo .exe
R2	Debe permitir ingresar y manipular un máximo de 8 archivos de audio al tiempo
R3	Posibilidad de controlar el nivel de entrada de la señal
R4	Posibilidad de orientar el archivo de audio espacialmente 360° por medio de técnicas Ambisonics
R5	Posibilidad de controlar la distancia virtual del archivo de audio
R6	Posibilidad de modificar la cobertura del archivo de audio
R7	Posibilidad de agregar reverberación al archivo de audio con control de "dry/wet"
R8	Posibilidad de reproducir archivos de audio en formato binaural o Ambisonics de tercer orden

2.4 Diseño

Para determinar la relación entre el usuario y el sistema, se hace una representación de los casos de uso y la funcionalidad del aplicativo. El diagrama de casos de uso define como el conjunto de funciones afines que el sistema debe cumplir para satisfacer los requerimientos de diseño. Este conjunto de funcionalidades son representados por los casos de uso, como se muestra en la figura 1.

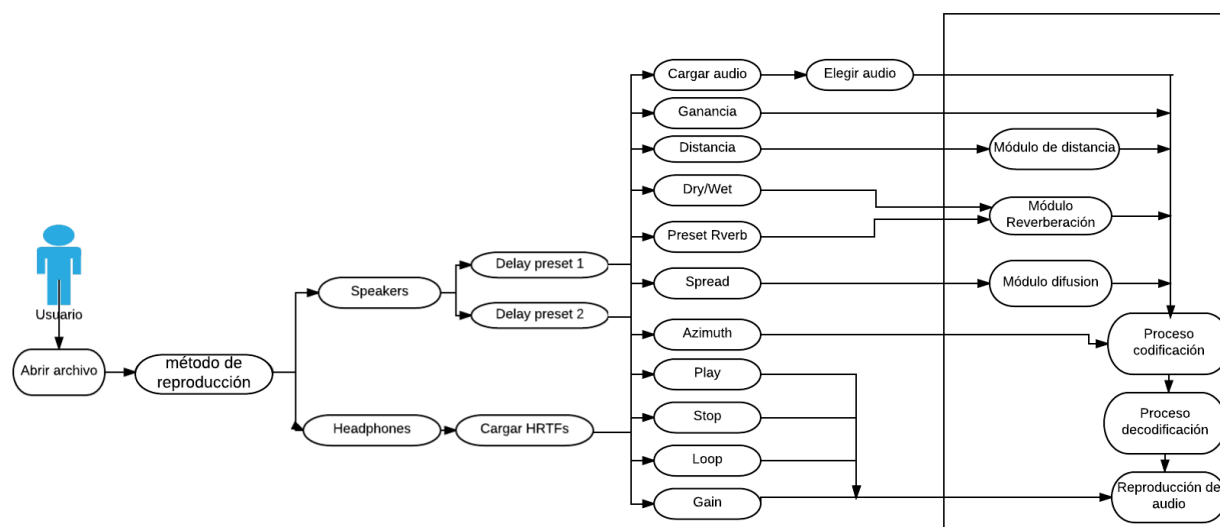


Figure 1: Diagrama de caso de uso

En las tablas 2 y 3 se hace una descripción detallada de cada uno de los casos de uso operados por el usuario y se hace referencia a los requerimientos consignados en la tabla 1.

Table 2: Descripción de los casos de uso

Ambi-Roll DESCRIPCIÓN DE CASOS DE USO	
Nombre:	Abrir la aplicación.
Actor:	Usuario.
Función:	Permitir utilizar la aplicación.
Descripción:	El usuario debe dar doble click en el ejecutable de la aplicación.
Referencia:	R1
Nombre:	Selección del método de espacialización.
Actor:	Usuario.
Función:	El usuario puede elegir entre reproducir el archivo de audio por medio de Ambisonics (8 canales) o binaural.
Descripción:	El usuario debe dar click en el botón que diga el método que desee, en caso de ser Binaural, debe cargar las respectivas HRTFs dando click en el botón Cargar HRTF's y cargando la carpeta adjunta en la carpeta donde se encuentra el ejecutable.
Referencia:	R8

Table 3: Continuación de la descripción de los casos de uso

Ambi-Roll DESCRIPCIÓN DE CASOS DE USO	
Nombre:	Cargar uno o varios archivos de audio.
Actor:	Usuario.
Función:	Introducir en la aplicación el archivo de audio que va a ser procesado y cuadrar el o los respectivos niveles.
Descripción:	El usuario debe dar click en el botón Cargar audio.
Referencia:	R2 y R3
Nombre:	Espacializar el archivo de audio.
Actor:	Usuario.
Función:	Dar la dirección de procedencia deseada para el archivo de audio.
Descripción:	El usuario debe por medio de la interfaz gráfica ubicar el archivo de audio donde lo desee.
Referencia:	R4
Nombre:	Reverberación.
Actor:	Usuario.
Función:	Aplicar reverberación al archivo procesado espacialmente.
Descripción:	El usuario por medio de un control Gráfico de Dry/Wet podrá añadir a la señal reverberación.
Referencia:	R7
Nombre:	Control SPREAD.
Actor:	Usuario.
Función:	Control de cobertura del archivo de audio.
Descripción:	El usuario podrá elegir entre tres tipos de cobertura para el archivo de audio, el primero definido, disperso y difuso.
Referencia:	R6
Nombre:	Control de distancia.
Actor:	Usuario.
Función:	Control de distancia del archivo de audio.
Descripción:	El usuario podrá elegir por medio de un control de ganancia la distancia en que se encuentra el archivo de audio espacialmente, es decir, donde se encuentra el archivo con respecto a quien escucha.
Referencia:	R5
Nombre:	Amplitud.
Actor:	Usuario.
Función:	Control de amplitud de los archivos de audio.
Descripción:	El usuario podrá hacer una mezcla independiente controlando la amplitud de la señal de cada canal y un control de amplitud Master de la señal.
Referencia:	R8

2.5 Implementación

A continuación se define el flujo de la señal mediante el diagrama de bloques presentado en la figura 2. Este diagrama, muestra además, las diferentes etapas que componen el aplicativo, teniendo en primera instancia un control de reproducción con la posibilidad de reproducir en bucle (loop), posteriormente la señal pasa por un control de nivel de entrada, para seguir a un proceso de control de nivel por distancia, luego, la señal es afectada por un módulo de reverberación que tiene control de "Dry/wet" para ajustar la cantidad de reverberación deseada, la opción de "Spread" permite al usuario dar un efecto de dispersión a la señal deseada en caso que no se desee una ubicación espacial puntual. Para dar una ubicación específica, se implementa el control "Azimuth", el cual permite ubicar la fuente en una dirección dada por el usuario. Toda esta información es utilizada por el módulo de codificación "Bencoder", el cual me codifica la señal para una posición en términos de armónicos esféricos, esta señal codificada ingresa al decodificador "Decoder" que me permite la reproducción del campo sonoro en un sistema multicanal (8 altavoces). Además es posible escoger si la reproducción se realiza de forma binaural.

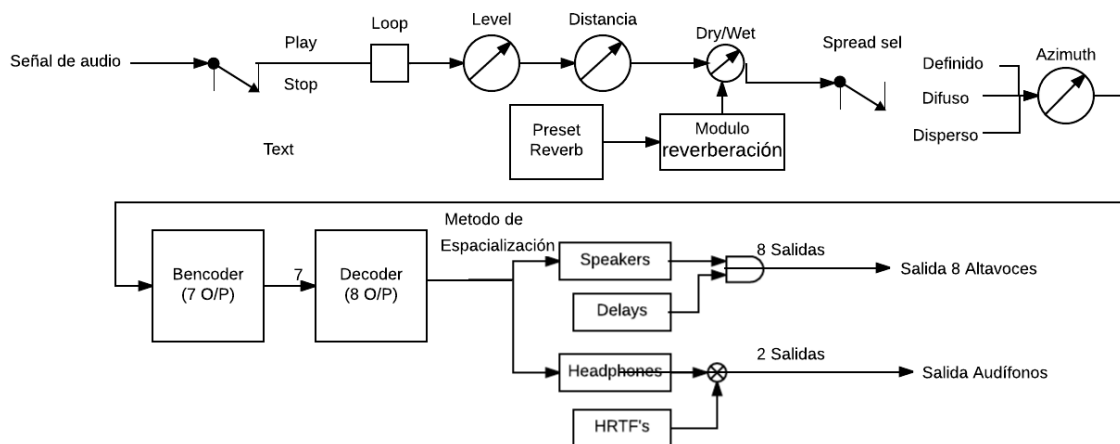


Figure 2: Diagrama de bloques del aplicativo

2.6 Validación

2.6.1 Pruebas de usuario

Se realizó una prueba de aceptación de usuario con el objetivo de validar el cumplimiento de los requerimientos de diseño de la aplicación. A fin de reconocer la aceptación de la aplicación por parte de diferentes usuarios, se realizó una encuesta con preguntas concretas relacionadas a los requerimientos de funcionamiento. Cada uno de los usuarios encuestados, tuvieron la libertad de verificar el debido funcionamiento de cada uno de los controles variables de la aplicación antes de responder las preguntas que se muestran a continuación.

Encuesta para la prueba de aceptación del usuario

Las siguientes preguntas se responderán en una escala del 1 al 5, siguiendo los siguientes apartados.

1. Totalmente insatisfecho/Totalmente en desacuerdo.
2. Insatisfecho/ En desacuerdo.

3. Indiferente.
4. Satisfecho/De Acuerdo.
5. Totalmente satisfecho/Totalmente de acuerdo.

-
1. Nombres y apellidos.
 2. Documento de identidad.
 3. Profesión.
 4. Ocupación

- Docente
- Estudiante

5. ¿La interfaz de usuario es amigable?

1. 2. 3. 4. 5.

6. ¿La aplicación le otorga la posibilidad de cargar archivos de audio (Máximo 8)?

1. 2. 3. 4. 5.

7. ¿La aplicación le permite escoger un método de reproducción entre Ambisonics o Binaural

1. 2. 3. 4. 5.

8. ¿Considera usted que la aplicación le permite modificar la dirección de procedencia del archivo de audio?

1. 2. 3. 4. 5.

9. ¿Considera que la reverberación cumple con las características de este efecto?

1. 2. 3. 4. 5.

10. ¿Considera que el control de Spread si permite percibir la diferencia del archivo de audio cuando es definido, disperso o difuso?

1. 2. 3. 4. 5.

11. ¿Considera que Se puede percibir el cambio en la ubicación del archivo de audio modificando el parámetro de distancia?

1. 2. 3. 4. 5.

12. ¿Considera que la aplicación permite tener un control de nivel de la señal adecuado?

1. 2. 3. 4. 5.

2.6.2 Resultados de la prueba de aceptación

Los resultados de la prueba de aceptación son los siguientes

- el 100% de los usuarios estuvieron totalmente satisfechos con la facilidad del uso de la interfaz de usuario.
- el 100% de los usuarios estuvieron totalmente satisfechos con la posibilidad de procesar los ocho archivos de audio.
- el 100% de los usuarios estuvieron totalmente satisfechos con la posibilidad de escoger un método de reproducción (Sistema ambisonics o binaural).
- el 80% de los usuarios estuvieron totalmente satisfechos y el 20% restante estuvieron satisfechos con que la aplicación permite modificar la dirección de procedencia del los archivos de audio.
- el 80% de los usuarios estuvieron totalmente satisfechos y el 20% restante estuvieron satisfechos con las características de la reverberación.
- el 90% de los usuarios estuvieron totalmente satisfechos con la percepción que genera la el control de Spread, mientras el 10% estuvieron satisfechos.
- el 100% de los usuarios estuvieron totalmente satisfechos con la percepción que genera la modificación del parámetro distancia.
- el 80% de los usuarios estuvieron totalmente satisfechos y el 20% restante estuvieron satisfechos con el control de nivel general.

3 Estructura de la herramienta

La aplicación contiene un panel principal subdividido en diferentes módulos que se encargan de las diferentes funciones. Existen 3 secciones principales: la primera sección se encuentra en la parte superior donde están los controles de entrada y el método de espacializcion, la segunda se encuentra en la parte media y contiene los canales de audio, y en la tercera se encuentran las opciones de salida y los medidores.

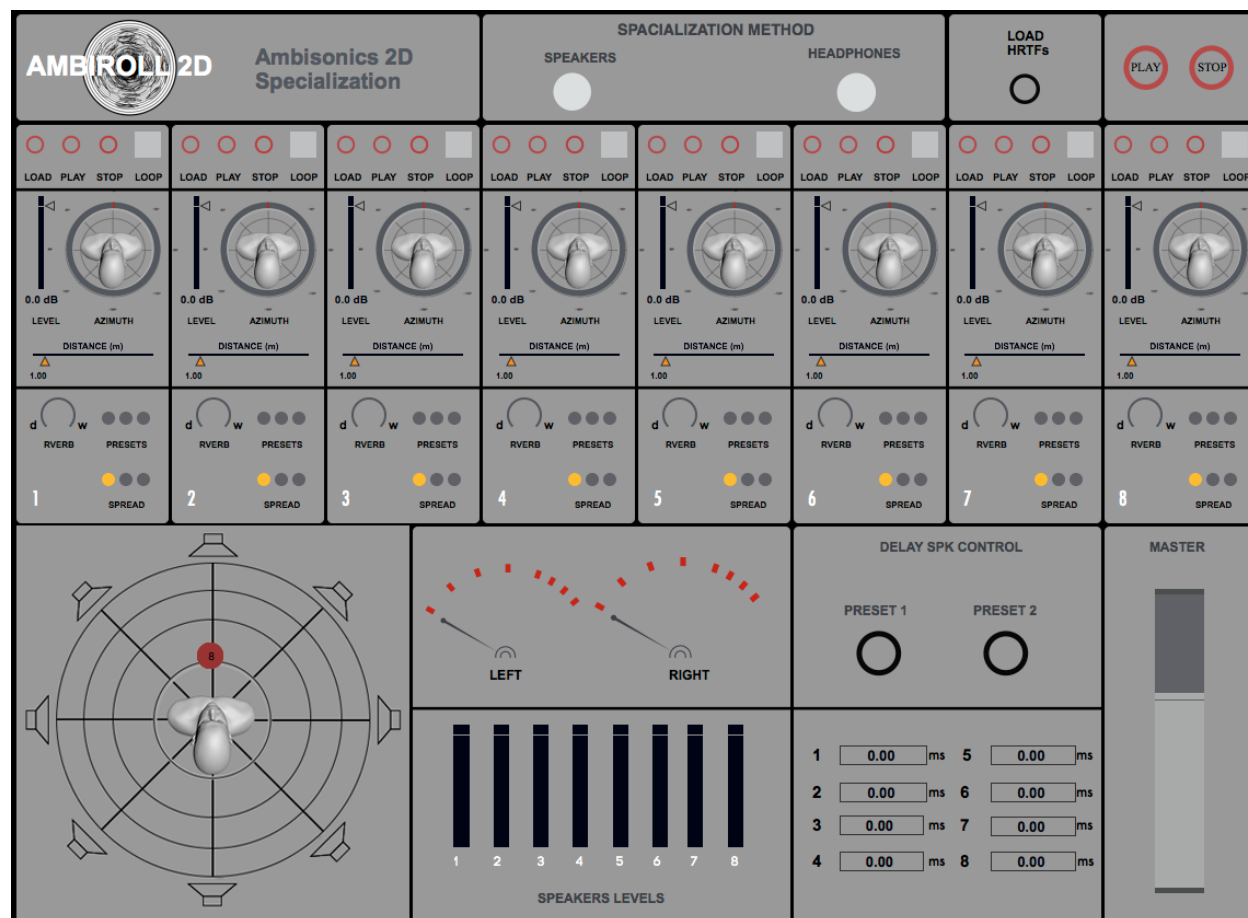


Figure 3: Interfaz gráfica de usuario

3.1 Módulos

3.1.1 Espacialización

Este módulo permite escoger el método de espacialización con el que se va a usar el sistema, el modo speakers utiliza ambisonics para espacializar mediante 8 altavoces, mientras que el modo headphones utiliza ambisonics y respuestas al impulso binaurales para espacializar mediante unos audífonos.



Figure 4: Modulo de elección de método de espacialización

3.1.2 Load HRTFs

En este módulo se puede cargar un banco de 16 HRTFs, la aplicación incluye un paquete de una cabeza binaural genérica de referencia KU100 de Neumann. El usuario puede cargar cualquier otro banco de HRTFs correspondientes a las posiciones de los altavoces.

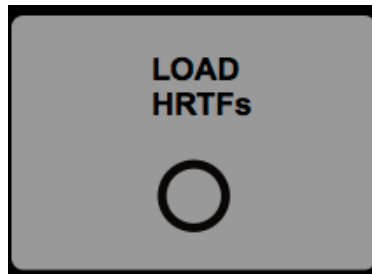


Figure 5: Módulo para cargar las HRTFs

3.1.3 PLAY/STOP

Los botones play y stop permiten reproducir o parar todos los canales al tiempo



Figure 6: Módulo PLAY/STOP

3.1.4 CANALES

El modulo de los canales permiten controlar diferentes parámetros que actúan sobre el audio, en la región superior se encuentran los botones de reproducción y carga del audio, en la parte central el fader level ajusta el nivel de salida de la pista, el control azimuth espacializa el audio en el eje azimuthal desde 0° a 360° . Por otro lado esta el control Distance que utiliza la ley del inverso cuadrado para dar un efecto de distancia a cada sonido y esta dado en metros. En la parte inferior se encuentran los efectos, el primero es una reverberación que contiene 3 diferentes presets, el segundo un control de spread que cambia la distribución del sonido en los altavoces 15° , 30° o 45° . Ambi-Roll tiene 8 canales donde se pueden modificar diferentes sonidos al tiempo.

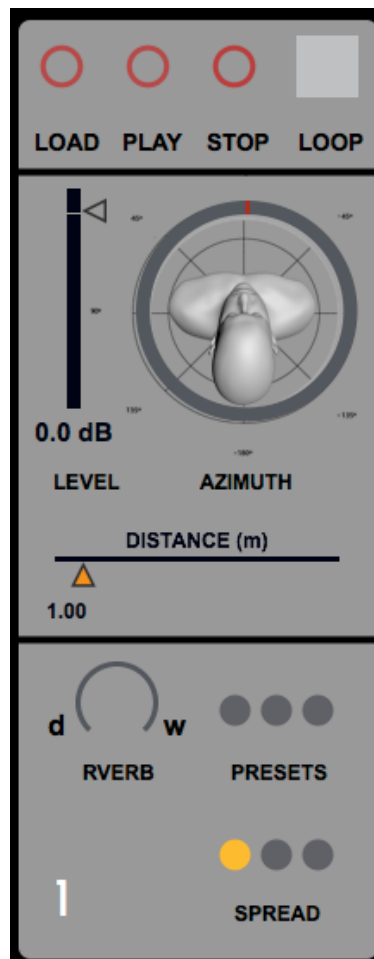


Figure 7: Módulo de canal

3.1.5 DELAY SPEAKERS CONTROL

El control de delays permite modificar en tiempo real el retraso de las señales de los altavoces, esto con el fin de poder acceder a diferentes distribuciones de los altavoces sin interferir con la reconstrucción del campo sonoro generado por ambisonics. Este control cuenta con dos presets.

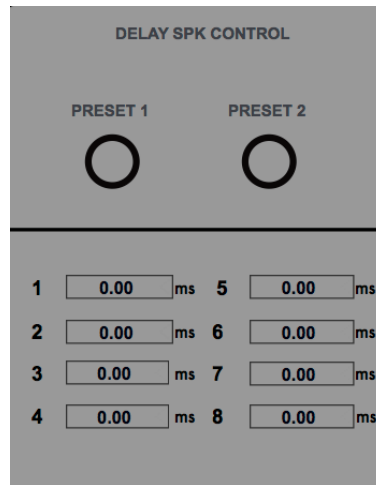


Figure 8: Módulo de control de delays

3.1.6 Visualización

El módulo de visualización muestra la distribución de los sonidos de cada canal a lo largo del eje azimutal y la pérdida por distancia aplicada. Cada canal está nombrado con su respectivo número.

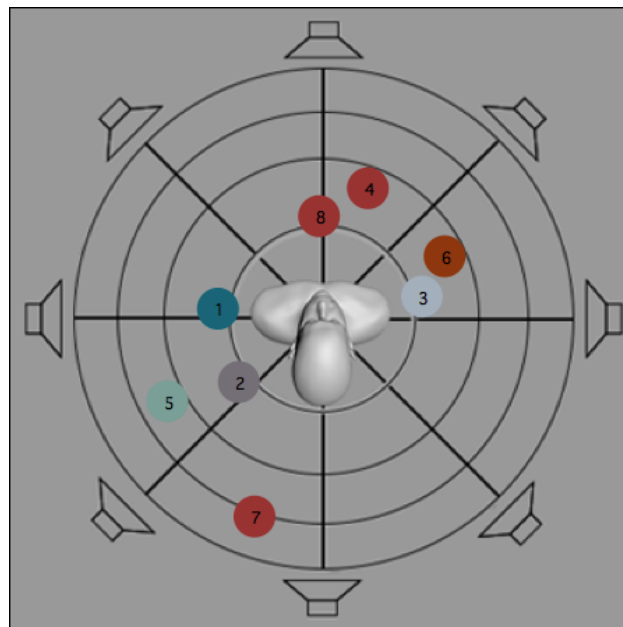


Figure 9: Módulo de visualización

3.1.7 LEVELS

Este módulo permite observar los niveles RMS de las señales tanto de los altavoces como los audífonos.

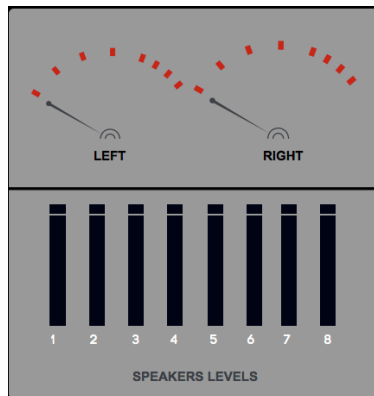


Figure 10: Módulo de niveles

3.1.8 MASTER

Este módulo controla el nivel de salida de la señal tanto como para los altavoces como para la señal binaural.

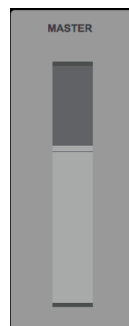


Figure 11: Módulo Master

El usuario tiene la posibilidad de controlar todos los parámetros descritos.