

AUDIO ESPACIAL

3DX SPATIALIZER - PLUGIN PARA LA ESPACIALIZACIÓN DE ARCHIVOS DE AUDIO

31 de octubre de 2017

Diego Mauricio Murillo Gómez
Angélica Londono Borja
Thomas Viana Alzate
Ing. de Sonido

Índice general

0.1. Introducción	2
0.2. Descripción del Software	2
0.2.1. Instalación	2
0.2.2. Especificaciones	2
0.2.3. Análisis	3
0.2.4. Diseño	3
0.2.5. Implementación	7
0.2.6. Validación	8
0.2.6.1. Pruebas de usuario	8
0.3. Estructura del software	10
0.3.1. Módulos	11
0.3.1.1. LOAD IR:	11
0.3.1.2. DRY / WET:	12
0.3.1.3. DISTANCE:	12
0.3.1.4. AZIMUTH / ELEVATION:	12
0.3.1.5. INPUT / OUTPUT:	13

0.1. INTRODUCCIÓN

En la actualidad hay un acelerado crecimiento en la industria multimedia, la cual busca incansablemente desarrollar formatos de reproducción visual y auditiva que le permitan al usuario tener una experiencia inmersiva. Todo lo relacionado con el audio espacial se ha convertido en un tema atractivo para muchos amantes de los videojuegos y el contenido social mediático. El *3DX Spatializer* es un plugin desarrollado para espacializar archivos de audio en tiempo real y recrearlos en un recinto específico permitiendo crear experiencias sonoras inmersivas en proyectos musicales y audiovisuales. El plugin está desarrollado bajo la implementación de convoluciones en tiempo real con respuestas al impulso y HRTFs (*Head Related Transfer Functions*).

0.2. DESCRIPCIÓN DEL SOFTWARE

0.2.1. Instalación

Para que el DAW (Digital Audio Workstation) pueda reconocer el plugin es indispensable pegar el archivo .VST en la carpeta de plugins definida por el DAW.

0.2.2. Especificaciones



- Lenguaje de programación: Cabbage/cSound.
- Plataforma: macOSX.
- Denominación de la aplicación: 3DX Spatializer
- Versión: 1.0.
- Base teórica: Convolución en tiempo real y síntesis binaural.

0.2.3. Análisis

El plugin 3DX Spatializer es una herramienta que le permite al usuario especificar la orientación de procedencia en términos de azimuth y elevación de archivos de audio y recrearlos en un recinto determinado. Además de tener estas opciones, el usuario tiene la posibilidad de determinar la distancia a la cual desea ubicar el archivo de audio, modificar el nivel de entrada y salida de la señal y automatizar parámetros.

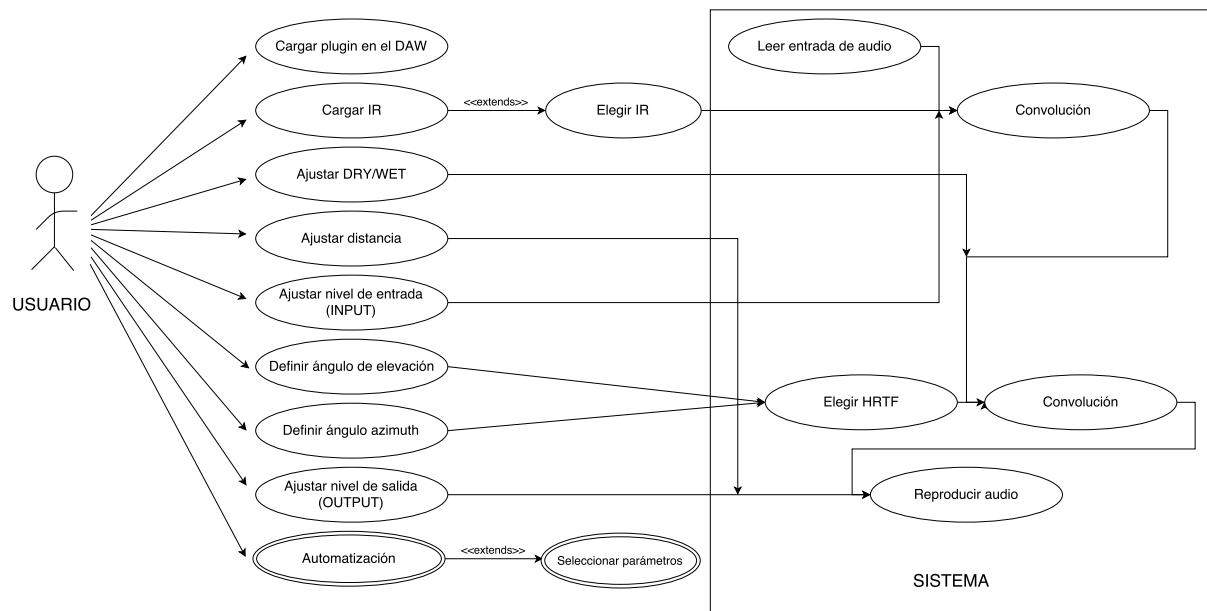
En la tabla 1 se especifican los requerimientos de diseño y una descripción del proyecto.

Tabla 1: Descripción y requerimientos del proyecto

PROYECTO	VST 3DX Spatializer
DESCRIPCIÓN	El proyecto tiene como objetivo espacializar archivos de audio en el espacio tridimensional recreándolos en un recinto deseado mediante la convolución con una respuesta al impulso que el usuario puede cargar.
REQUERIMIENTOS	<p>R1 Debe ser un software tipo plugin con formato .VST.</p> <p>R2 Posibilidad de cargar diferentes respuestas al impulso monofónicas.</p> <p>R3 Posibilidad de controlar la influencia de la respuesta al impulso sobre el archivo de audio.</p> <p>R4 Posibilidad de controlar la distancia virtual del archivo de audio.</p> <p>R5 Posibilidad de controlar el nivel de entrada y salida de la señal.</p> <p>R6 Posibilidad de orientar el archivo de audio en cualquier dirección del espacio tridimensional.</p> <p>R7 Procesamiento en tiempo real de la espacialización.</p> <p>R8 Posibilidad de automatizar parámetros.</p>

0.2.4. Diseño

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

**Figura 1:** Diagrama de casos de uso

En las tablas 2, 3 y 4 se hace una descripción detallada de cada uno de los casos de uso y subcasos de uso operados por el usuario y se hace referencia a los requerimientos consignados en la tabla 1, con los cuales tiene relación.

Tabla 2: Descripción de los casos de uso

3DX Spatializer**DESCRIPCIÓN DE CASOS DE USO**

Nombre:	Cargar plugin en el DAW
Actor:	Usuario
Función:	Permitir el funcionamiento del plugin en el DAW.
Descripción:	El usuario debe cargar el plugin tipo VST en un canal de audio en el DAW.
Referencia:	De requerimientos: R1
Nombre:	Cargar IR, Elegir IR
Actor:	Usuario
Función:	Permitir elegir una respuesta al impulso para cargar al plugin.
Descripción:	Una vez se decide cargar la respuesta al impulso (IR), el usuario puede elegir entre diferentes IR monofónicas (Sala 1, Sala 2, Sala 3, Sala 4, Capilla).
Referencia:	De requerimientos: R2, R7
Nombre:	Ajustar DRY/WET
Actor:	Usuario
Función:	Controlar la influencia de la IR sobre el archivo de audio.
Descripción:	El usuario puede determinar, mediante un control DRY/WET, qué tan “seco” o “mojado” va a estar el audio procesado. Entendiendo por “seco” el archivo de audio sin influencia de la IR, y por “mojado” el archivo completamente influenciado por la IR.
Referencia:	De requerimientos: R3, R7
Nombre:	Ajustar distancia
Actor:	Usuario
Función:	Controlar la distancia a la que se desea ubicar el archivo de audio.
Descripción:	El usuario puede determinar mediante este control la lejanía, o bien, la cercanía del archivo de audio en el espacio.
Referencia:	De requerimientos: R4, R7

Tabla 3: Continuación Descripción de los casos de uso

3DX Spatializer**DESCRIPCIÓN DE CASOS DE USO**

Nombre: Ajustar nivel de entrada

Actor: Usuario

Función: Controlar el nivel de entrada.

Descripción: El usuario puede determinar el nivel de entrada de la señal antes de ser procesada.

Referencia: De requerimientos: R5

Nombre: Definir ángulo de elevación.

Actor: Usuario

Función: Determinar el ángulo de elevación del archivo de audio.

Descripción: El usuario puede especificar el ángulo de elevación en un rango de -80° a 80°, dependiendo de la orientación del archivo de audio deseada en el eje de elevación. Entendiendo -80° como el punto más bajo y 80° como el más alto.

Referencia: De requerimientos: R6, R7

Nombre: Definir ángulo azimuth.

Actor: Usuario

Función: Determinar el ángulo azimuth del archivo de audio.

Descripción: El usuario puede especificar el ángulo azimuth en un rango de -180° a 180°, dependiendo de la orientación del archivo de audio deseada en el eje cenital. Entendiendo ángulos negativos como la parte izquierda de la cabeza y los ángulos positivos como la derecha.

Referencia: De requerimientos: R6, R7

Tabla 4: Continuación Descripción de los casos de uso**3DX Spatializer****DESCRIPCIÓN DE CASOS DE USO**

Nombre: Ajustar nivel de salida

Actor: Usuario

Función: Controlar el nivel de salida.

Descripción: El usuario puede determinar el nivel de la señal luego de ser procesada.

Referencia: De requerimientos: R5, R7

Nombre: Automatización

Actor: Usuario

Función: Permitir la automatización de parámetros.

Descripción: El usuario tiene la posibilidad de automatizar cada uno de los parámetros variables del plugin directamente desde el DAW.

Referencia: De requerimientos: R8, R7

Nombre: Automatización, Seleccionar parámetros.

Actor: Usuario

Función: Determinar los parámetros a automatizar.

Descripción: El usuario debe especificar en el DAW cada uno de los parámetros que desee automatizar.

Referencia: De requerimientos: R8, R7

0.2.5. Implementación

Con el propósito de satisfacer cada uno de los requerimientos, antes de comenzar a desarrollar el código en el entorno de programación de Cabbage, se define el flujo de señal mediante un diagrama de bloques. La señal de audio entra al plugin y, en primera instancia, se ve afectada por un control de nivel antes de ser procesada; luego, una vez el usuario carga una respuesta al impulso, esta señal de audio es convolucionada con dicha respuesta al impulso. Se establece un control DRY/WET que permite hacer una mezcla entre el archivo de audio sin convolucionar y el resultado de la convolución, para así determinar la influencia que se tiene de la respuesta al impulso sobre el archivo de audio. Luego se realiza la convolución con las HRTFs (Head Related Transfer Functions) determinadas por la coordenada que especifique el usuario mediante los controles de azimuth y elevación. Finalmente la señal de audio es

afectada por un control de distancia basado en la ley del inverso cuadrado y por un control de nivel de salida, como se observa en la figura 2.

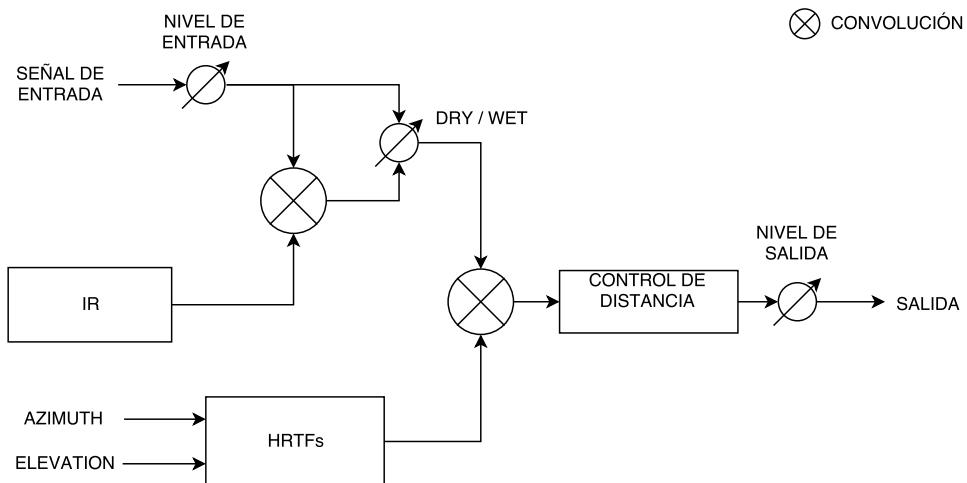


Figura 2: Diagrama de bloques del flujo de señal

0.2.6. Validación

0.2.6.1. Pruebas de usuario

Se realizó una prueba de aceptación de usuario diseñada siguiendo las recomendaciones del estándar de la National Bureau Standards (NBS) Special Publication 500-136 con el objetivo del desarrollo de validar el cumplimiento de los requerimientos de diseño del software desarrollado. A fin de reconocer la aceptación del software por parte de diferentes usuarios, se realizó una encuesta con preguntas concretas relacionadas a los requerimientos de funcionamiento del plugin. La prueba manual del funcionamiento del software se realizó en actividad con el DAW Reaper. Cada uno de los usuario encuestados, tuvieron la libertad de verificar el debido funcionamiento de cada uno de los controles variables del plugin antes de responder las preguntas que se muestran a continuación.

Encuesta para la pueba de aceptación de usuario

Antes de comenzar, verifique junto con el encuestador, que el software haya sido cargado correctamente en el DAW. Se realizarán preguntas concretas con respecto al funcionamiento del software, orientadas a la verificación de los requerimientos de diseño propuestos.

1. ¿Es el software evaluado, un software tipo plug-in en formato .VST?
Sí No
¿Es otro formato?, ¿cuál?.
2. ¿El software le otorga la posibilidad de cargar diferentes respuestas al impulso monofónicas?
Sí No
¿Por qué no?
3. ¿El software le permite controlar la influencia de la respuesta al impulso sobre el archivo de audio?
Sí No
¿Por qué no?
4. ¿Considera que el software evaluado ofrece la posibilidad de alejar virtualmente el archivo de audio?
Sí No
¿Por qué no?.
5. ¿Piensa usted que el software permite orientar el archivo de audio hacia cualquier dirección de un espacio tridimensional?
Sí No
¿Por qué no?
6. ¿Considera que el software realiza la espacialización del archivo de audio en tiempo real?
Sí No
¿Por qué no?
7. Teniendo en cuenta la anterior evaluación, ¿acepta usted el desarrollo del software, indicando y verificando que cumple con las funciones para el cual está diseñado?
Sí No
¿Por qué no?

Resultados de la prueba de aceptación:

Los resultados obtenidos después de la realización de la prueba fueron los siguientes:

- 100% de los encuestados estuvieron de acuerdo con que el formato del plugin es .VST.
- 100% estuvieron de acuerdo con que otorga la posibilidad de cargar diferentes respuestas al impulso monofónicas.
- 100% afirmaron que el plugin permite controlar la influencia de la respuesta al impulso sobre el archivo de audio mediante el control DRY/WET.
- 90% de los encuestados acordaron que el software evaluado ofrece la posibilidad de alejar virtualmente el archivo de audio. El otro 10% argumentan que no solo basta con controlar la atenuación por distancia, sino que se podrían considerar filtros por absorción del aire y cambios en cuanto a las reflexiones del recinto, lo cual computacionalmente implicaría cargar un tren de respuestas al impulso convirtiéndose en un plugin computacionalmente muy pesado y robusto.
- 90% estuvieron de acuerdo con que el plugin permite orientar el archivo de audio hacia cualquier dirección de un espacio tridimensional. El otro 10% estuvieron dudosos en cuanto a algunos ángulos de procedencia entendiendo que el método de espacialización puede presentar inconsistencias dependiendo de la morfología de la persona dado que el banco de HRTFs utilizado corresponde a la respuesta de una cabeza genérica.
- 100% de los encuestados consideraron que el software realiza la espacialización del archivo de audio en tiempo real.
- 100% aceptaron el desarrollo del software, indicando y verificando que cumple con las funciones para el cual está diseñado.

Todos los usuarios encuestados estuvieron satisfechos con la interfaz de usuario.

0.3. ESTRUCTURA DEL SOFTWARE

Desde que el plugin es cargado sobre el canal de DAW, automáticamente carga una respuesta al impulso por defecto con el fin de que no sea un requisito indispensable para utilizar el módulo de especialización. Permite cargar una HRTF mediante un botón y tiene seis(6) controles variables para modificar cada uno de los parámetros.



Figura 3: Interfaz gráfica de usuario

0.3.1. Módulos

El plugin cuenta con tres módulos implícitos, uno para controlar la reverberación por convolución con la respuesta al impulso, uno para espacializar el archivo de audio, y otro para manejar la estructura de ganancia con controles de nivel de entrada y salida de la señal.

0.3.1.1. LOAD IR:

Permite al usuario elegir entre diferentes respuestas al impulso monofónicas. El software viene por defecto con cinco respuestas al impulso, pero el usuario tiene la posibilidad de cargar cualquiera que tenga a su disposición.



0.3.1.2. DRY / WET:

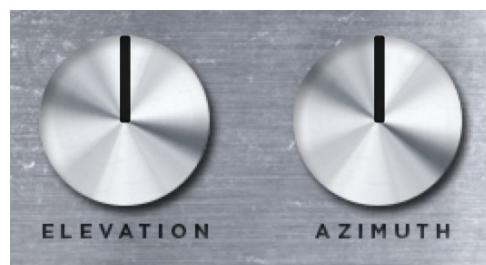
Permite al usuario determinar cuán influyente será la respuesta al impulso sobre el archivo de audio.

**0.3.1.3. DISTANCE:**

Con un algoritmo basado en la pérdida de nivel por distancia y ley del inverso cuadrado, permite al usuario alejar la fuente virtualmente hasta 30 m.

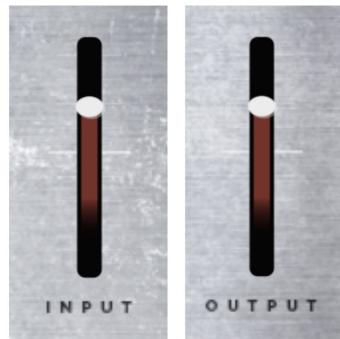
**0.3.1.4. AZIMUTH / ELEVATION:**

Permiten definir la orientación de procedencia del archivo de audio especificando una coordenada esférica a una resolución de 10°. ELEVATION toma valores entre -80° (abajo) y 80° (arriba) en el eje cenital y AZIMUTH entre -180° (izquierda) y 180° (derecha) en el eje azimutal.



0.3.1.5. INPUT / OUTPUT:

Permite controlar la el nivel de entrada del audio sin procesar y el nivel de salida de la señal procesada.



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