



FIA 2018

XI Congreso Iberoamericano de Acústica; X Congreso Ibérico de Acústica; 49º Congreso Español de Acústica -TECNIACUSTICA'18-
24 al 26 de octubre

APLICACIÓN EN MATLAB PARA EL DISEÑO Y PRESENTACIÓN DE PAISAJES SONOROS VIRTUALES

PACS: 43.60.Ac

Valencia, A. ; San Martín, R. ; Ezcurra, A. ; Arana, M.
Universidad Pública de Navarra. Departamento de Física.
Campus de Arrosadía. 31006. Pamplona. Spain
Tel.: 948 168 451
Fax: 948 169 565
E-mail: ana.valencia@unavarra.es

Palabras Clave: paisajes sonoros, aplicación Matlab, escenas sonoras virtuales, tiempo real.

ABSTRACT

The concept of soundscapes refers to the perception and analysis of the sound environment of real acoustic landscapes, composed of sets of sound sources that surround the listener. Through the application presented, it is possible to create virtual soundscapes and reproduce them using headphones or multi-speaker configurations.

This software presents an user interface that allows to activate and move the imported sound sources in real time, play and record the scene. The scene can be presented in binaural format using Head-Related Transfer Functions (HRTFs) or speakers using amplitude panning techniques (Vector-Based Amplitude Panning, VBAP).

RESUMEN

El concepto de paisaje sonoro hace referencia a la percepción y análisis del ambiente sonoro de entornos acústicos reales, compuestos por conjuntos de fuentes sonoras que rodean al oyente. Mediante la aplicación presentada es posible crear de manera virtual paisajes sonoros y reproducirlos a través de auriculares o configuraciones de múltiples altavoces.

La herramienta cuenta con una interfaz de usuario que permite activar y desplazar en tiempo real las fuentes sonoras importadas, reproducir y grabar la escena. Ésta puede presentarse en formato binaural mediante Head-Related Transfer Functions (HRTFs), o altavoces utilizando técnicas de paneo en amplitud (Vector-Based Amplitude Panning, VBAP).

FIA 2018

XI Congreso Iberoamericano de Acústica; X Congreso Ibérico de Acústica; 49º Congreso Español de Acústica -TECNIACUSTICA'18-
24 al 26 de octubre

1. INTRODUCCIÓN

El campo sonoro que nos rodea nos ayuda a percibir de manera tridimensional nuestro entorno, entendiéndolo como un conjunto de fuentes acústicas ubicadas en posiciones concretas a nuestro alrededor. Así, los paisajes sonoros comunican información que nuestro cerebro procesa generando una valoración del ambiente sonoro. Esta valoración está relacionada de manera directa con la calidad acústica del entorno y puede ser descrita y analizada de forma objetiva con el análisis de medidas realizadas *in situ*. No obstante, la recreación controlada de campos sonoros en entornos de laboratorio permite una evaluación sistemática de los mismos, para lo que resulta de gran utilidad el empleo de técnicas de presentación de audio basadas en síntesis binaural (Head-Related Transfer Function) o distribuciones de altavoces (Vector-Based Amplitude Panning [1], Ambisonics Amplitude Panning, Wield Field Synthesis...).

Existen numerosos entornos de audio que permiten desarrollar herramientas de síntesis y composición de audio en tiempo real, como CSound [2], RT-Cmix [3], Pure Data [4], MAX [5] o SuperCollinder [6]. Desde la inclusión de la DSP System Toolbox en Matlab también es posible el streaming de audio en tiempo real en este entorno, gracias a la incorporación del soporte de ASIO.

Existen también herramientas enfocadas a la investigación, como "SoundScape Renderer" (SSR) [7] y "SART3D: 3D Spatial Audio Reproduction Toolbox" [8]. La primera de ellas, SSR, es una herramienta que permite la reproducción de audio espacial en tiempo real mediante el uso de una gran variedad de algoritmos de renderizado: WFS, VBAP [1], AAP, Near-field-corrected Higher-Order Ambisonics... Desarrollado en C++, solo puede ejecutarse en LINUX o MAC OS X.

La herramienta SART3D [8] ha sido desarrollada en Matlab y ofrece la posibilidad de procesar audio espacial en tiempo real de manera intuitiva, desplazando las fuentes virtuales a través de las pantallas de visualización y con la posibilidad de cambiar la configuración del dispositivo de reproducción empleado mediante una modificación sencilla del código desarrollado. Los algoritmos de renderizado se presentan en la herramienta de manera independiente, lo que permite incluir nuevos que se puedan desarrollar.

La herramienta presentada en este artículo permite recrear escenas sonoras en entornos de laboratorio combinando el procesado de audio espacial que ofrecen otras herramientas con la posibilidad de grabar el resultado que se va generando en tiempo real. Cuenta con una interfaz de usuario que por un lado permite configurar nuevas escenas y distribuciones de altavoces de manera sencilla y, por otro lado, muestra la escena virtual con funciones básicas de reproductor en tiempo real: *play/pausa*, *stop* y *rec*. Además, SCENA da la opción de exportar la escena resultante en diferentes formatos, algo muy útil para poder emplear los entornos generados de manera virtual para investigaciones posteriores.

2 . DESARROLLO DE LA HERRAMIENTA

El software ha sido diseñado con el fin de crear escenas sonoras virtuales a partir de la carga de ficheros de audio, que posteriormente se puedan grabar y exportar en diferentes formatos. Cada una de las pistas cargadas es representada en la escena virtual mediante un icono que puede ser desplazado en tiempo real, además de modificar su estado a activo o inactivo.

El desarrollo total de SCENA ha sido realizado mediante el software Matlab, desde la interfaz de usuario mediante la que se configura y controla la escena virtual hasta la comunicación con el hardware de reproducción de audio, pasando por el procesado de las señales originales cargadas.

FIA 2018

XI Congreso Iberoamericano de Acústica; X Congreso Ibérico de Acústica; 49º Congreso Español de Acústica -TECNIACUSTICA'18-
24 al 26 de octubre

2.1 Entorno De Programación

Matlab cuenta con numerosas toolboxes con funciones y algoritmos aplicables a diversos campos: matemáticas, estadística, sistemas de control, procesamiento de señal... Para desarrollar SCENA se ha recurrido al uso de funciones específicas para el diseño de sistemas de procesamiento de señales de audio en tiempo real y la comunicación con entornos de representación de audio disponibles en DSP System Toolbox.

La estructura del programa puede dividirse en cuatro fases:

FASE 1. Creación y carga de variables de entrada:

Al iniciar SCENA el usuario puede escoger el método de presentación de la escena sonora y diseñar o cargar una configuración de altavoces determinada. Esta información queda almacenada en una estructura tipo *setup*. Actualmente, la herramienta presenta como técnicas de presentación de audio la síntesis binaural mediante HRTF y audio multicanal mediante VBAP (distribución de altavoces). El gran potencial de VBAP es la cantidad de posibilidades de configuración que presenta, debido a su funcionamiento basado en la composición de un vector unidad de dos o tres dimensiones (configuración en plano o esférica, respectivamente) resultantes de la combinación de los vectores de los altavoces más cercanos a la fuente virtual representada.

Tras escoger la configuración del dispositivo de salida, el usuario puede cargar una escena existente o crear una nueva. Como resultado, el programa genera una estructura tipo *scene*. A la hora de inicializar esta estructura se tiene en cuenta la posterior presentación en tiempo real, que requiere que las pistas sean subdivididas en pequeños fragmentos de modo que puedan ser modificados los parámetros de las fuentes virtuales que representan (posición o estado) de uno a otro.

Estructura *setup*:

Contiene campos para la identificación (*Rep* – nombre de la configuración), la definición de la técnica de presentación de audio empleada (*Renderer* – puede ser 'VBAP' o 'HRTF') y la configuración física del dispositivo (*LScoord* – coordenadas esféricas de los altavoces – y *ChannelMapping* – mapeo de los canales de salida).

Estructura *scene* (S):

Contiene subestructuras en cuyos campos se almacena la información de cada una de las pistas: *id* – identificador de pista, *file* – nombre del fichero de la pista, *name* – nombre escogido por el usuario, *icon* – nombre del fichero del icono que representa la fuente virtual, *allpositions* – coordenadas de la fuente virtual para cada uno de los fragmentos de la pista y *active* – vector que indica si la fuente está o no activa en cada uno de los fragmentos.

FASE 2. Inicialización de variables de salida:

SCENA convierte las variables de entrada anteriores en una escena virtual que se reproduce y modifica en tiempo real. Para ello, es necesario contar con dos objetos de presentación que permitan llevar a cabo la renderización de las pistas de entrada y su posterior reproducción a través del dispositivo de salida, todo ello en streaming en tiempo real.

Objeto virtual *icon* (VI):

En las propiedades del objeto se almacena la información de cada una de las pistas cargadas por el usuario (nombre, posición y estado), el objeto de representación virtual de la fuente con el que interactúa el usuario, un segundo objeto con los datos correspondientes al método de renderizado y una estructura que contiene los canales activos en cada momento y los filtros que se deben aplicar a las señales de entrada (fragmentos de las pistas de audio).

FIA 2018

XI Congreso Iberoamericano de Acústica; X Congreso Ibérico de Acústica; 49º Congreso Español de Acústica -TECNIACUSTICA'18-
24 al 26 de octubre

Como objeto, *VI* cuenta con una serie de métodos que le permiten consultar y modificar sus propiedades, de manera que la interacción del usuario con la fuente virtual se traduce en una actualización del objeto. Por ejemplo, en caso de que se traslade la fuente, cambiará la posición del objeto y se actualizarán los filtros de renderizado.

Objeto *audio_out*.

DSP Toolbox ofrece la posibilidad de crear un objeto *dsp.AudioPlayer* capaz de reproducir audio a través de un dispositivo de reproducción conectado. Contiene los parámetros necesarios para la reproducción: *DeviceName* – nombre del dispositivo conectado, *SampleRate* – frecuencia de muestreo, *BufferSize* – tamaño del buffer de salida (longitud de los fragmentos en los que se dividen las pistas de entrada para su presentación) y *ChannelMapping* – mapeo de los canales de salida.

El método *step* permite reproducir la señal de salida según la longitud de buffer escogida.

FASE 3. Puesta en marcha del reproductor:

A partir de las variables de entrada y salida el software es capaz de reproducir en tiempo real una escena. Para ello, fragmenta la señales audio de entrada (Figura 1), tomando en cada iteración un número de muestras igual al tamaño de cola definido (*audio_out.BufferSize*).



Figura 1. División de las pistas de audio (*T*) en fragmentos (*F*) de tantas muestras como la longitud del buffer (*N*). La longitud de las pistas es *L*.

Para cada fragmento, la función de reproducción programada lleva a cabo las siguientes acciones (ver Figura 2):

- Comprobación del estado de las fuentes virtuales.
Las fuentes pueden estar activas o inactivas, y su estado puede cambiar de un fragmento a otro por la interacción del usuario. Así, es necesario identificar cuáles de ellas están activas para procesarlas posteriormente o sencillamente omitirlas. Además, si la escena cargada ya ha sido grabada anteriormente, el programa debe leer el estado almacenado puesto que los cambios en el mismo pueden deberse a esa grabación previa en vez de a la interacción del usuario en el momento de la reproducción y, por lo tanto, las actuaciones posteriores serán diferentes.
- Comprobación de la posición de las fuentes virtuales.
Una vez terminado el “recuento” de fuentes activas, se procede con la búsqueda de posibles cambios en la posición de las fuentes. De nuevo, estos cambios se pueden deber a la interacción del usuario en el momento o bien a grabaciones anteriores de la escena. En cada caso, el programa actúa en consecuencia: si es por interacción, se actualiza la posición del objeto *VI* y con ello los filtros de renderizado que serán aplicados y, en caso de estar activada la opción de grabado, se actualiza la variable *scene*; si se debe a grabaciones previas, se actualiza la posición del objeto de fuente virtual almacenado en *VI*.
- Aplicación de los filtros de renderizado y reproducción.
El programa genera una señal de salida en cada iteración. Esta señal es la suma de los fragmentos correspondientes de cada pista filtrados según el método de renderizado

FIA 2018

XI Congreso Iberoamericano de Acústica; X Congreso Ibérico de Acústica; 49º Congreso Español de Acústica -TECNIACUSTICA'18-
24 al 26 de octubre

seleccionado. Así, se procede a aplicar en bucle a cada una de las pistas los filtros cargados previamente según su posición (un filtro por cada canal activo). Además, en el caso de que la posición de la fuente varíe entre dos iteraciones la función aplica cross-fading, un método que permite suavizar los cambios repentinos de amplitud de la señal mediante la suma de la señal filtrada en la posición previa y la filtrada en la actual, ambas con un enventanado tipo Hanning.

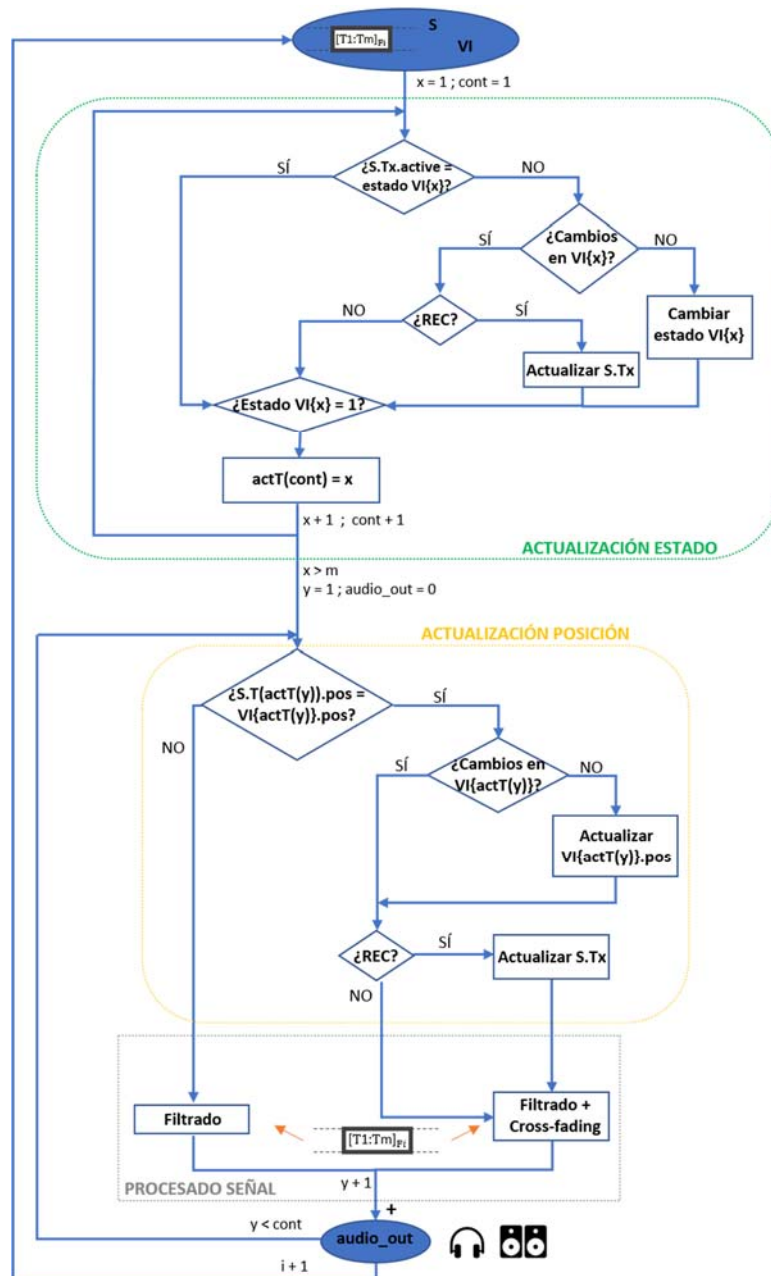


Figura 2. Diagrama de flujo que muestra el funcionamiento del reproductor en sus tres fases: actualización de estado, actualización de posición y procesado de señal.

FASE 4. Exportación de la escena creada:

FIA 2018

XI Congreso Iberoamericano de Acústica; X Congreso Ibérico de Acústica; 49º Congreso Español de Acústica -TECNIACUSTICA'18-
24 al 26 de octubre

El reproductor permite activar la opción de grabado. En ese caso, se actualiza la variable *scene*, de manera que su matriz de posiciones y su vector de estado reflejan los cambios de las fuentes virtuales en el tiempo. Estas escenas, además, se guardan en un fichero .mat al finalizar o parar su reproducción, por lo que se pueden cargar posteriormente para su reproducción o modificación. Este es el verdadero potencial de la herramienta, su capacidad de almacenar escenas de manera eficiente a través de una variable estructural.

A partir de esta variable se puede obtener un fichero de audio .wav en pocos segundos gracias al método de exportación implementado en SCENA. El programa requiere tan solo de la variable *scene* y la información sobre el tipo de renderizado escogido por el usuario para generar una versión .wav de la escena creada preparada para ser reproducida a través del dispositivo de salida descrito durante la inicialización de *setup*. El funcionamiento de la exportación es básicamente el mismo que el de la reproducción, con la diferencia de que no es necesario ralentizar la ejecución de manera que el audio se reproduzca en streaming, aprovechando así al máximo el potencial de procesamiento que ofrece Matlab.

Además, de manera paralela el programa genera un fichero de audio en B-format. De este modo se compatibiliza el resultado de SCENA con plataformas populares que actualmente están capacitadas para reproducir videos 360º con audio espacial mediante la tecnología Ambisonics, como YouTube o Facebook.

2.2 Interfaz De Usuario

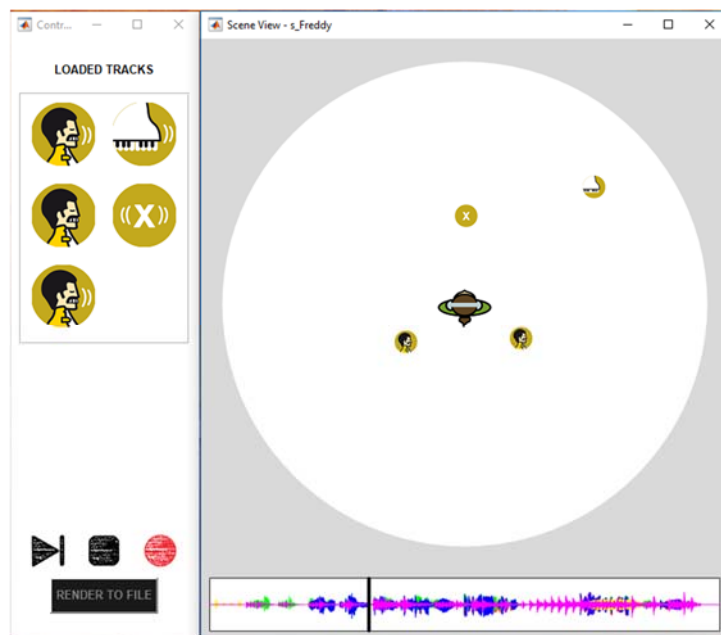


Figura 3. Interfaz de usuario que muestra la escena virtual (derecha) junto al panel de control (izquierda). En la escena cargada una de las fuentes se muestra inactiva.

En la Figura 3 se muestra una escena virtual presentada mediante síntesis binaural, no obstante también se pueden diseñar configuraciones de altavoces si se va a emplear VBAP. Esta técnica permite emplear distribuciones tanto bidimensionales como tridimensionales. A través de la pantalla de setup el usuario puede crear y cargar configuraciones cúbicas, esféricas, circunferencias centradas en el oyente (caso 2D), configuraciones surround como 7.1, 5.1...

FIA 2018

XI Congreso Iberoamericano de Acústica; X Congreso Ibérico de Acústica; 49º Congreso Español de Acústica -TECNIACUSTICA'18-
24 al 26 de octubre

Una vez cargadas la configuración del dispositivo de salida y la escena el usuario puede interactuar con la escena virtual a través del reproductor. Además de las funciones básicas (play, pausa, stop, rec, adelantar y atrasar), es posible cambiar el estado de las fuentes y desplazarlas. También se puede acceder a la función de exportación mediante el botón "RENDER TO FILE".

3 . PUESTA EN MARCHA

Una vez desarrollada la herramienta, se diseñaron varias escenas para comprobar su correcto funcionamiento. Las escenas eran presentadas mediante las dos técnicas, HRTF y VBAP, constatando que la escena virtual mostrada en la pantalla se correspondía con la percibida.

Para estas pruebas se emplearon recreaciones musicales, conjuntos de pistas correspondientes a los distintos instrumentos que integran una composición musical. Así, además de verificar el correcto funcionamiento de la herramienta se exploró su funcionalidad como software musical capaz de mezclar y exportar audio, añadiendo como novedad la percepción tridimensional de la escena musical resultante y su control en tiempo real. Esto abre la posibilidad de profundizar en aplicaciones para la herramienta que aporten nuevas funcionalidades a la edición de audio ofertada por softwares musicales convencionales (Pro Tools, Audacity, Garageband, FL Studio...).

Las escenas se sometieron a testeo en una sala preparada para la renderización de audio espacial, situada en el Laboratorio de Acústica de la Universidad Pública de Navarra (Figura 4).



Figura 4. Aplicación SCENA en la sala de renderización de audio espacial del Laboratorio de Acústica de la Universidad Pública de Navarra.

4 . CONCLUSIONES Y LÍNEAS FUTURAS

Se ha desarrollado una herramienta en Matlab capaz de reproducir, modificar y grabar entornos sonoros en tiempo real. Las escenas resultantes se pueden exportar en formato .wav y B-format (Ambisonics), lo que hace que la aplicación sea compatible con otros entornos que integran el audio espacial, como YouTube. La interacción en tiempo real del usuario con la escena a través de su interfaz ayuda a comprender de manera sensorial e intuitiva el concepto de audio espacial.

FIA 2018

XI Congreso Iberoamericano de Acústica; X Congreso Ibérico de Acústica; 49º Congreso Español de Acústica -TECNIACUSTICA'18-
24 al 26 de octubre

Este software es el punto de partida para el desarrollo de nuevas aplicaciones con funcionalidades más específicas. Una de ellas es el análisis de paisajes sonoros desde el punto de vista de la psicoacústica. Existe un conjunto de parámetros psicoacústicos, como loudness, roughness o sharpness, que analizados en conjunto permiten determinar el grado de agrado de un entorno. Este análisis se puede realizar sobre los entornos recreados mediante SCENA, siendo posible además llevar a cabo una comparativa entre los resultados volcados por la psicoacústica y las percepciones reales de individuos sometidos a un test subjetivo.

Otra aplicación futura del software es la recreación de escenas musicales. Actualmente no es posible actuar sobre las pistas cargadas de manera directa, la herramienta se limita a aplicarles filtros de renderizado dependientes de su posición. Resultaría útil tener la opción de modificarlas por separado mediante efectos que pueden ser tan simples como un filtrado en amplitud que permita simular los cambios de intensidad ejecutados por el instrumento. Esto podría ser añadido a SCENA, que resultaría la solución más cómoda para el usuario, o se podría buscar la compatibilidad del software con los resultados de otros programas que ya cuentan con esas funcionalidades.

Por último, en un futuro la herramienta contará con la opción de aplicar codificación de audio direccional (Directional Audio Coding, DirAC), una técnica no lineal de procesamiento y síntesis de audio espacial basada en que los mecanismos de percepción del sonido directo son diferentes a los del sonido difuso. Una escena se descompondrá por tanto en su componente directa y su componente difusa, y ambos grupos de señales se podrán sintetizar con técnicas diferentes. De ese modo, la herramienta ganará en realismo y además se abrirá la puerta a nuevas aplicaciones que tengan en cuenta las características acústicas del entorno simulado.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el Ministerio de Economía y Competitividad a través del proyecto de Investigación I+D+I con referencia BIA2016-76957-C3-2-R

REFERENCIAS

- [1] V. Pulkki (1997). *Virtual Sound Source Positioning Using Vector Base Amplitude Panning*. Journal of the Audio Engineering Society 45(6): 456-466.
- [2] B. Vercoe and D. Ellis (1990). *Real-Time CSOUND: Software Synthesis with Sensing and Control*. ICMC Glasgow 1990 Proceedings.
- [3] B. Garton and D. Topper (1997). *RTcmix -- Using CMIX in Real Time*. ICMC Thessaloniki 1990 Proceedings.
- [4] M. Puckette (1996). *Pure Data: another integrated computer music environment*. Proceedings of the second intercollege computer music concerts 37-41.
- [5] M. Puckette, T. Apel and D. Zicarelli (1998). *Real-time audio analysis tools for Pd and MSP*. ICMC Michigan 1998 Proceedings.
- [6] J. McCartney (2002). *Rethinking the computer music language: SuperCollider*. Computer Music Journal, 26(4): 61-68.
- [7] W. J. Davies (2015). *Cognition of soundscapes and other complex acoustic scenes*. Proceedings of the 44th International Congress and Exposition on Noise Control Engineering, San Francisco.



FIA 2018

XI Congreso Iberoamericano de Acústica; X Congreso Ibérico de Acústica; 49º Congreso Español de Acústica -TECNIACUSTICA'18-
24 al 26 de octubre

[8] G. Moreno, M. Cobos, J. Lopez-Ballester, P. Gutierrez-Parera, J. Segura and A. M. Torres (2015). *On the Development of a MATLAB-based Tool for Real-time Spatial Audio Rendering*. Proceedings of the 138th Convention of the Audio Engineering Society, Warsaw.