

LIBRO DE LA AUDIO ENGINEERING SOCIETY SECCIÓN COLOMBIA

AVANCES DEL AUDIO EN LATINOAMÉRICA.



Audio Engineering Society
Sección Colombia

Acústica / Electroacústica /
psicoacústica

Producción (Grabación,
mezcla y mastering)

Refuerzo Sonoro /
Sonido en Vivo

Postproducción (Cine,
Broadcast, TV, Videojuegos)





Audio Engineering Society

Sección Colombia

AUDIO ENGINEERING SOCIETY

Sección Colombia

LIBRO DE LA AUDIO ENGINEERING SOCIETY SECCIÓN
COLOMBIA

Avances del audio en Latinoamérica

2016

ISBN- 978-958-56264-0-9

Publicación dirigida a la comunidad científica y profesional del audio en Latinoamérica

AES REGIÓN LATINOAMÉRICA

Vice Presidenta

VALERIA PALOMINO

COMITÉ AES COLOMBIA

Presidente

ANDRÉS MILLÁN

Vice Presidente

RICARDO ESCALLÓN

Secretario

CARLOS CABALLERO

Tesorero

FABIO NICHOLLS

Vocales

RAJIV ARORA

WILLIAM FLÓREZ

ANDREA LIÉVANO

ANDRÉS NÚÑEZ

JORGE VALENCIA

JOHNATAN MONTENEGRO



Comité Científico

PhD. Mauricio Ardila

PhD. Ana María Jaramillo Jaramillo

PhD. José Ricardo Zapata

Mg. Daniel Vásquez

Mg. Ricardo Escallón

Mg. Jorge Mario Valencia Upegui

Mg. Carlos Andrés Caballero Parra

Mg. Luis Fernando Hermida

Mg. Raúl Enrique Rincón

Mg. Luis Jorge Herrera

Ing. Manuel Fernando Torres Cifuentes

Ing. Belman Rodriguez

Hugo Villegas

Compilador

Carlos Andrés Caballero Parra

Diseño Gráfico

Shapes | Comunicación Visual

Publicación Digital

Editado por:

Instituto ENE Audio

Los contenidos de ésta publicación son responsabilidad de cada autor, por lo tanto, no comprometen la opinión del Instituto ENE Audio y tampoco de la Sección AES Colombia.

El Comité AES Colombia autoriza la reproducción parcial o total de los contenidos sólo con fines académicos, con la solicitud expresa de mencionar la fuente.

Audio Engineering Society

Sección Colombia

Cel: +57 (310) 768-3433

Contacto: facebook.com/aescolombia

Web: www.aes-colombia.org

Bogotá - Colombia

CONTENIDO



1. Presentación

Pag. 4

2. Diseño y construcción de un preamplificador a válvulas con emulación de calidad tonal de dos amplificadores clásicos de guitarra.

Johnatan López Pilco

Pag. 5

3. Metodología para la medición de directividad en tiple colombiano

Fredy Alzate Arias

José Julián Cadavid Sierra

Pag. 11

4. Comparación entre diferentes frecuencias de muestreo y profundidad de bits en el desarrollo de una producción musical

Sebastián Lopera Gómez

Pag. 19

5. Mezcla de audio binaural mediante interfaz tangible

Juan Esteban Herrera

Ricardo Moreno

Pag. 27

6. Desarrollo de algoritmo de interpretación de respuestas al impulso en tres dimensiones

Federico Nahuel Cacavelos

Augusto Bonelli Toro

Alejandro Bidondo

Pag. 33

7. Audiojuego con implementación de audio binaural

Andrés Carmona

Diego Sierra

Ricardo Moreno

Pag. 39

8. Métodos alternativos de espacialización compuesta. Parte 1

Juan David Sierra

Pag. 47

9. Métodos alternativos de espacialización compuesta. Parte 2

Juan David Sierra

Pag. 55

10. Criterios: medición de respuesta al impulso y obtención de parámetros acústicos

Yonny Quiceno

Alejandro Mejía

Juan F. Londoño

Antonio Escamilla

Pag. 61

11. Conferencia Latinoamericana
Resumen

Momentos destacados del evento

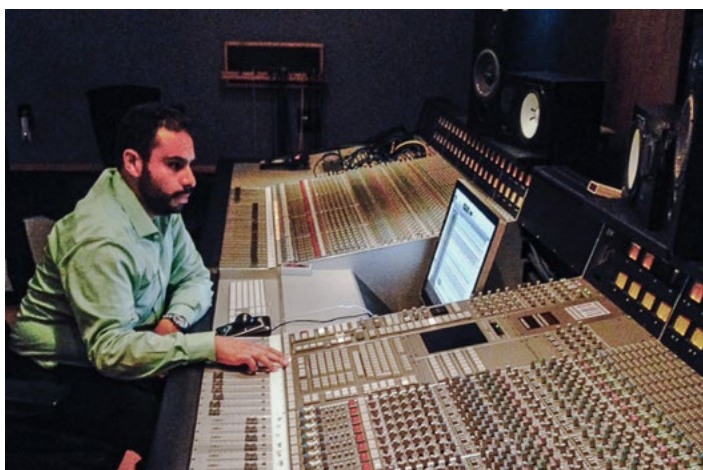
Pag. 69-70

12. Comité AES Colombia

Pag. 74

13. Próximos eventos

Pag. 75



PRESENTACIÓN

La Conferencia Latinoamericana de la Audio Engineering Society 2016 fue uno de los eventos regionales más importantes y destacados de la asociación. El impacto que marcó a lo largo de la comunidad de audio a nivel Latinoamericano fue importante y cumplió con los estándares de calidad establecidos por la AES. Para esta oportunidad, recibimos como anfitriones asistentes de Ecuador, Perú, Costa Rica, Estados Unidos, México, Argentina y diversas ciudades de Colombia, a quienes ofrecimos contenidos muy nutridos e impartidos por algunos de los expertos nacionales e internacionales más importantes y reconocidos de las diferentes ramas del sonido.

A diferencia de conferencias anteriores, para esta oportunidad decidimos presentar un evento que girara alrededor de cuatro ejes temáticos: Acústica/Electroacústica/Psicoacústica, Refuerzo Sonoro/Diseño de Sistemas de Sonido, Producción (Grabación, Mezcla y Mastering) y PostProducción de Audio/Videojuegos. A lo largo del evento se presentaron conferencias, talleres, paneles de discusión demostraciones, competencias de diseño y grabación, tours técnicos, críticas de grabación, ponencias y posters de investigación por estudiantes y profesionales nacionales e internacionales, haciendo de la Conferencia el evento más grande en audio efectuado en el país.

Adicionalmente, participó Sennheiser con la presentación y demostración de su sistema AMBEO 3D, Technik S.A. con audiometrías gratuitas para los asistentes y más de 25 compañías que presentaron sus productos y servicios en la zona de exhibición comercial. Este magno evento fue patrocinado por Electro Voice, Sennheiser, Olson Sound Design, Meyer Sound, InfoComm International, AVID y AFMG (Ahnert Feistel Media Group) y fue posible gracias al apoyo de la Universidad de San Buenaventura, el Instituto Ene Audio y Diffusion Magazine.

Los resultados de la Conferencia Latinoamericana 2016 conducen a la periodicidad de un evento anual de la asociación, a mantener una constante convocatoria de proyectos de investigación y al fortalecimiento de los contenidos presentados en las competencias estudiantiles de diseño y grabación. Ahora bien, las memorias de esta convocatoria inician el capítulo de un compromiso de incentivar a investigadores, científicos, docentes y estudiantes a fortalecer sus procesos de investigación y ser partícipes de futuras convocatorias lideradas por la AES dentro del marco de sus eventos locales e internacionales. Ing.

ING. ANDRÉS FELIPE MILLÁN SILVA
Presidente AES Colombia

AVANCES DEL AUDIO EN LATINOAMÉRICA.



Audio Engineering Society
Sección Colombia



DISEÑO Y CONSTRUCCIÓN DE UN PREAMPLIFICADOR A VÁLVULAS CON CALIDAD TONAL DE DOS AMPLIFICADORES CLÁSICOS DE GUITARRA

Jhonatan López Pilco

Ing. Jhonatan López Pilco
Ingeniería en Sonido y Acústica, Universidad de las Américas
 Quito, Ecuador
 jlopez@udlanet.ec
 jhonyg_4@hotmail.com

Resumen— El proyecto presenta el diseño y construcción de un preamplificador con dos distintas sonoridades, partiendo de dos diseños de preamplificadores de marcas reconocidas a nivel mundial como lo son Fender y Marshall. Empleando programas de simulación electrónica se simula los preamplificadores, estos se estudian y analizan para luego unirlos y tener un solo circuito. Posteriormente se lo lleva a una placa de pruebas para luego finalmente implementarlo en un circuito PCB y ensamblarlo en una caja. Así también se muestra mediciones y análisis parámetros técnicos como respuesta en frecuencia e impedancia. Finalmente se realiza comparaciones entre el preamplificador diseñado y uno de sus homólogos, mostrando así resultados en simulaciones y algunas entrevistas realizadas a algunos ingenieros en sonido.

Palabras clave— Diseño, construcción, preamplificador, válvulas, Fender, Marshall.

Abstract— This project deals with how a preamplifier is designed and built using valves with two different sounds, based on the two designs of the world renowned preamplifier brands of Marshall y Fender. Using electronic design and simulation to simulate preamplifiers, they are studied and analyzed to then join them in a single circuit. Next it is taken to a test board so that it can be finally installed into a PCB circuit and assembled in a box. This way the measurement and analysis of some technical parameters such as frequency response and impedance are also shown. Finally comparisons are carried out between the designed preamplifier and one of its counterparts, obtaining results with the emulation as demonstrated in meetings held with some sound engineers.

Keywords— Design, built, preamplifier, valves, Fender Marshal.

I. Introducción

Desde el año de 1946, cuando la empresa norteamericana Fender revoluciona la industria musical con la producción en masa de instrumentos y dispositivos amplificadores de audio para instrumentos musicales, empieza una carrera por encontrar nuevas posibilidades y características sonoras para que los músicos se identifiquen con estas.

En esta búsqueda de nuevas sonoridades nacen nuevas empresas una de ellas es la empresa inglesa Marshall, que 20 años después de la aparición de los primeros modelos de Fender revoluciona la industria con un nuevo sonido. Este sonido se caracterizaba por ser más agresivo que los modelos presentados por Fender sin embargo este sonido fue conseguido tras la experimentación con modelos propios de la marca de Fender.

Estas dos empresas, desde sus inicios se han caracterizado por su constante innovación en sus productos de amplificación, presentando así una infinidad de posibilidades en cuanto a lo que gustos sonoros se refiere. Además del prestigio ganado a lo largo de los años se han posicionado como líderes en el mercado mundial, ganándose la confianza de músicos profesionales como lo son: Pete Townshend (The Who), Billy Gibbons (ZZ Top), Larry "Ler" LaLonde (Primus),

para Fender y Angus y Mal Young (AC/DC), Jimi Hendrix, Zakk Wylde (Black Label Society), Steph Carpenter (Deftones), para Marshall, por mencionar algunos.

Por las razones antes expuestas es que hoy por hoy se considera como un buen punto de partida estudiar circuitos preamplificadores bases como lo son los de las marcas Fender y Marshall.

II. Diseño y Construcción

La sonoridad de un amplificador de guitarra depende de una serie de variables entre las que están: los componentes internos, el diseño, la configuración del preamplificador, la etapa de potencia, el altavoz o los altavoces elegidos y el diseño de la caja, sin tomar en cuenta el instrumento que influye en gran medida; sin embargo el presente artículo trata en particular sobre el diseño y construcción de la etapa de pre amplificación.

Un preamplificador viene definido eléctricamente por una

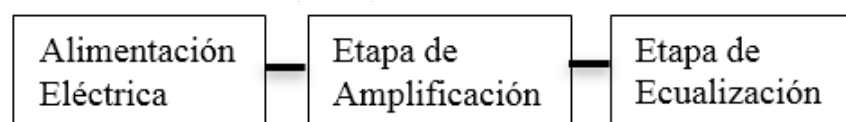


Fig. 1. Diagrama de bloques de un preamplificador.

A. Selección de circuitos preamplificadores.

El primer paso es la selección de los circuitos, para lo cual se tomó en cuenta principalmente que los diagramas eléctricos estén disponibles bibliográficamente o que se encuentren disponibles en la red para realizar el análisis respectivo. Los preamplificadores seleccionados fueron el Marshall JCM 800 y el Fender Champ 5E1.

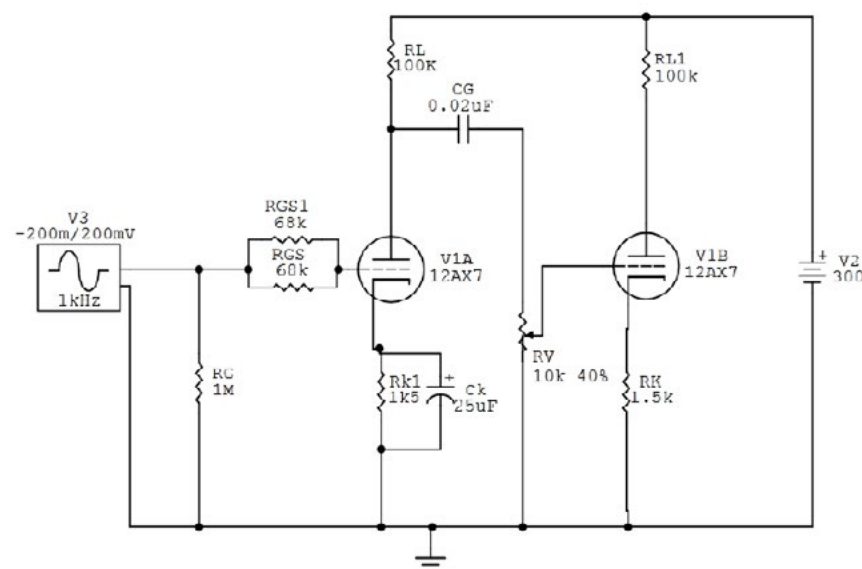


Fig. 2. Circuito Fender Champ 5E1.

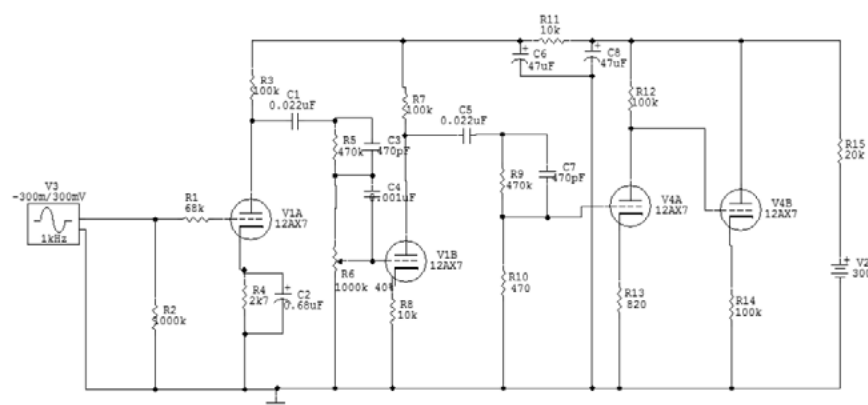


Fig. 3. Circuito Marshall JCM 800.

B. Circuito de ecualización.

El segundo paso para el diseño del preamplificador es la selección del ecualizador. Si bien un preamplificador de guitarra se encuentra definido por tres partes fundamentales (rectificación, amplificación y ecualización), la etapa de ecualización no forma parte del circuito de referencia del Fender Champ 5E1 y con este se presenta el primer reto, si debería o no ir éste dentro del circuito. Con fines experimentales se lo incluye, para así incrementar la versatilidad, ampliar los resultados y cumplir uno de los objetivos planteados que fue el de unir dos amplificadores en un solo circuito compartiendo la mayor cantidad de etapas.

Para el circuito ecualizador se toma en consideración el circuito presentado por el preamplificador Marshall JCM 800, el mismo que se mantendrá para el diseño final.

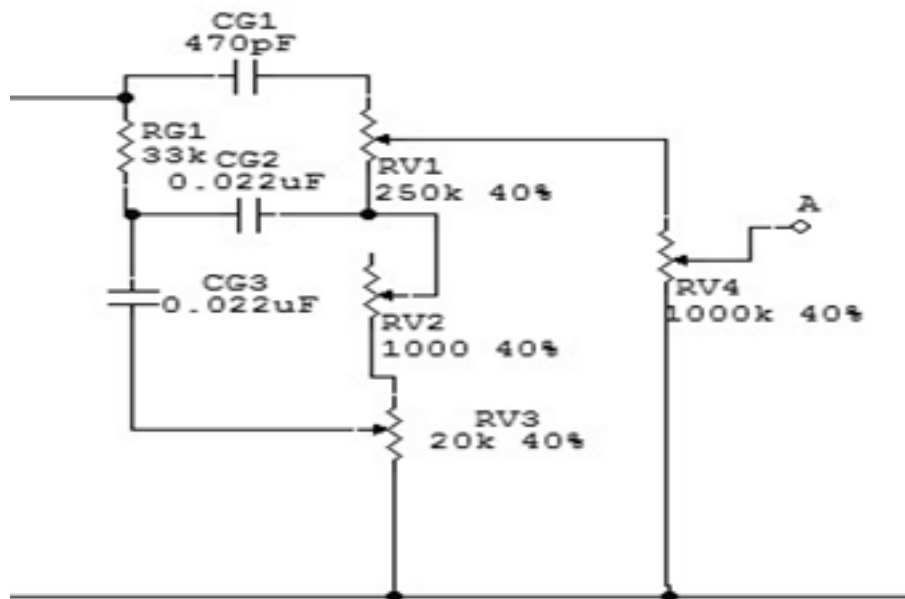


Fig. 4. Circuito ecualizador Marshall JCM 800.

C. Etapa de alimentación eléctrica.

Para la etapa de alimentación eléctrica de los circuitos amplificadores son necesarios cuatro circuitos rectificadores filtrados para obtener corrientes continuas, uno de 300 V (JCM 800), uno de 260 V (5E1) y dos de 12 V; estos últimos encargados de calentar el filamento de cada válvula (cátodo); de esto resulta evidente que un solo circuito rectificador de 12 V puede ser compartido.

Con la finalidad de simplificar una vez más el diseño se opta por la posibilidad de que los preamplificadores compartan uno de los otros circuitos rectificadores.

De forma teórica la variación del voltaje constante sobre la resistencia del cátodo (voltajes de 300 o 260 V) en un circuito amplificador con válvulas, afecta directamente a la línea de carga, la misma que indica la forma de onda amplificada resultante y que es creada a partir de los valores de los componentes electrónicos del diseño; es decir que al alterar este valor incidiremos en la forma de onda amplificada.

Mediante el análisis matemático se determina que una línea de carga con mayor voltaje genera una amplificación de la señal de entrada mayor y con menor distorsión, mientras que una línea de carga con menor voltaje genera una amplificación de la señal de entrada menor y con mayor distorsión; en otras palabras, si se aplica un voltaje de 260 V al circuito amplificador se tenderá a distorsionar la señal amplificada más que si aplicamos 300 V.

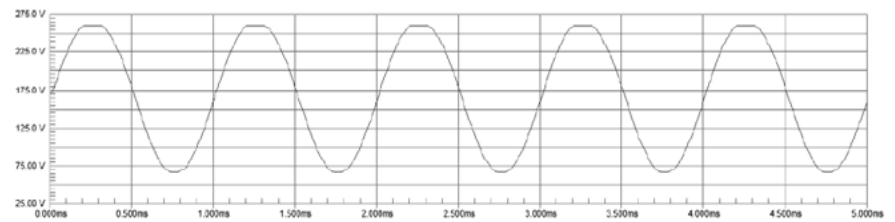


Fig. 5. Simulación preamplificador 5E1 con una fuente de voltaje de 260 V.

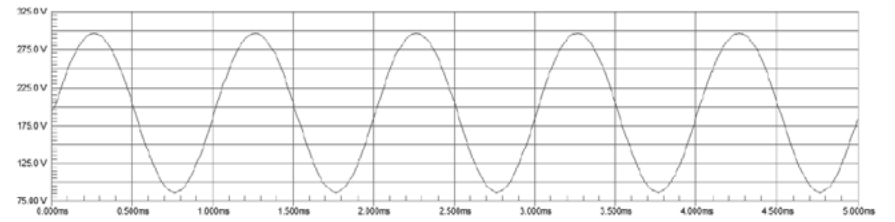


Fig. 6. Simulación preamplificador 5E1 con una fuente de voltaje de 300 V.

Para todas las mediciones realizadas se utiliza una señal de entrada de 1 kHz y una amplitud de 100 mV RMS, que es el voltaje promedio que entrega una guitarra al ser rasgada.

Conociendo los pros y contras que representa el utilizar una misma fuente para los dos amplificadores, se opta por utilizar el circuito rectificador de 300 V, siendo este el menos contraproducente.

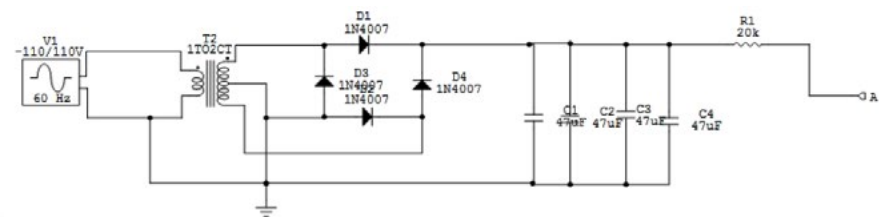


Fig. 7. Circuito rectificador de onda 300 V.

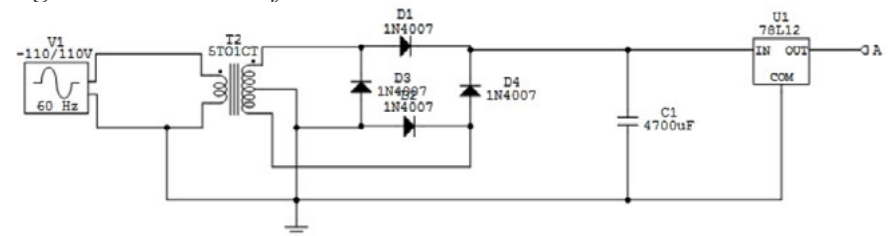


Fig. 8. Circuito rectificador de onda 12 V.

D. Simulación del preamplificador Fender Champ 5E1.

La simulación del circuito preamplificador se la realizó a través de dos programas de simulación eléctrica; CircuitMaker 2000 e ISIS de Proteus. El primero utilizado para realizar el diseño debido a que simplifica las conexiones y el segundo utilizado en la construcción.

Por simplicidad los gráficos mostrados a continuación son los de la simulación realizada en CircuitMaker 2000.

Es importante mencionar que para la simulación de este circuito se omite la etapa de rectificación y se la reemplaza por una batería de voltaje constante, que cumple la misma función.

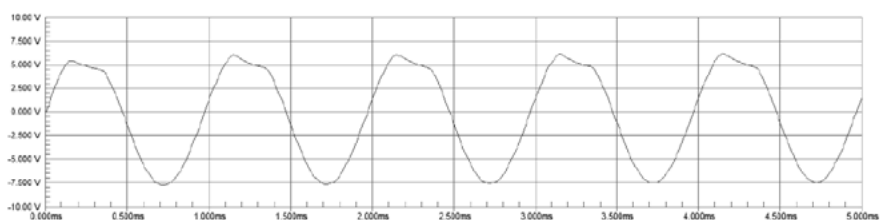


Fig. 9. Simulación del circuito Fender Champ 5E1 con etapa de ecualización.

La simulación es realizada con un voltaje de entrada cuya amplitud es de 100 mV RMS y una frecuencia de 1 kHz.

E. Simulación del preamplificador Marshall JCM 800.

Esta simulación se la realizó de forma igual que la anterior en el programa CircuitMaker 2000, con voltajes de entrada de 1 kHz y 100 mV RMS de amplitud.

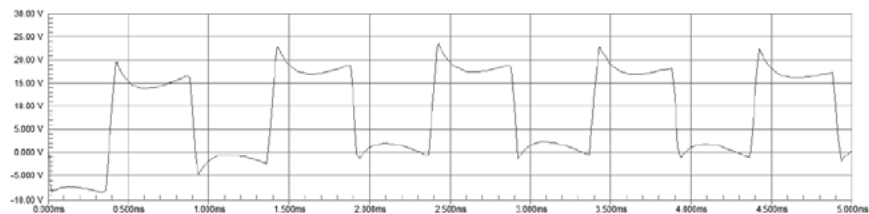


Fig. 10. Simulación del circuito Marshall JCM 800 con etapa de ecualización.

F. Diagrama de flujo de señal del circuito final.

El circuito preamplificador final se resume en la suma de todas las etapas interconectadas, los interruptores S1 y S2 rutean de manera simultánea las señales de audio a la entrada y salida, mientras que el interruptor S3 rutea los voltajes de alimentación eléctrica.

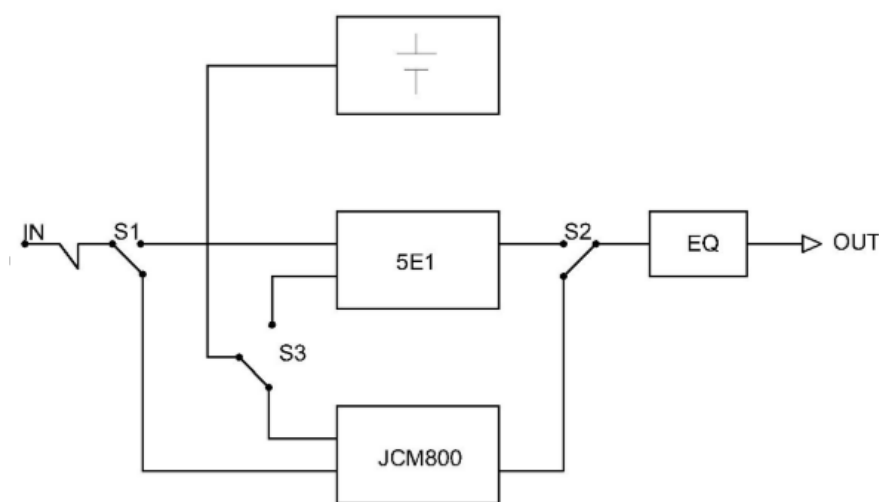


Fig. 11. Diagrama de flujo de señal del circuito preamplificador final.

G. Implementación del circuito en protoboard.

Previo a la implementación del circuito se realizó una simulación del circuito integrando todas las etapas anteriormente simuladas, de igual manera para la implementación del circuito en protoboard se realizó por etapas (rectificación, amplificación, ecualización) para finalmente unir las en un solo circuito final. La implementación en protoboard es principalmente útil para la toma de las mediciones.

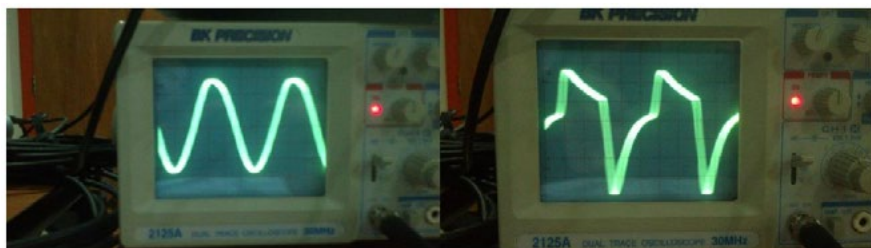


Fig. 12. Mediciones circuito Fender Champ 5E1 y Marshall JCM 800.

H. Implementación circuito PCB y ensamblaje.

La elaboración de la placa PCB es realizada por un técnico especializado, para lo cual es necesario poseer el circuito completo en un programa simulador. En el mercado ecuato-

riano se tiene como formato estándar entregar la simulación del circuito en el programa ISIS de Proteus para poder diseñar el PCB.

Dentro del programa de simulación Proteus se cuenta otro programa que trabaja en conjunto con el ISIS llamado ARES el cual permite visualizar el PCB y realizar correcciones antes de su impresión.

En el proceso de ensamblaje se realizó el montaje de componentes y se establecieron códigos de colores para los cables que conectan las válvulas, con finalidad de mantener el orden y evitar inducciones electromagnéticas.

Otro criterio importante que se tomó en cuenta durante el montaje es el uso de interruptores, con la finalidad de redirigir las señales eléctricas y simplificar el circuito e incrementar la versatilidad del preamplificador. Agrupando por un lado las señales de voltaje constante y por otro las señales de audio, facilitando también la manipulación del preamplificador.

I. Cálculo de la respuesta en frecuencia y la impedancia.

Por las dimensiones del circuito resulta bastante complejo el cálculo matemático de la respuesta en frecuencia sin embargo la evaluación experimental de la respuesta se la realiza a través de un analizador de espectro.

Por otro lado, el cálculo matemático de la impedancia se la realiza a través de la siguiente expresión matemática:

$$Z = R_L \left(\frac{V_s}{V_L} - 1 \right) \quad (1)$$

Dónde:

RL: es la resistencia de carga (1000 Ω).

VS: es el voltaje de salida.

VL: es el voltaje de entrada.

Z: es la resistencia igual a la impedancia de salida.

El valor del voltaje de entrada se lo obtiene por bandas de octava desde los 31.5 Hz a los 16 kHz, cuya amplitud es de 100 mV, mientras que el voltaje de salida es obtenido midiendo la salida del dispositivo.

III. Resultados

Impedancia preamplificadores

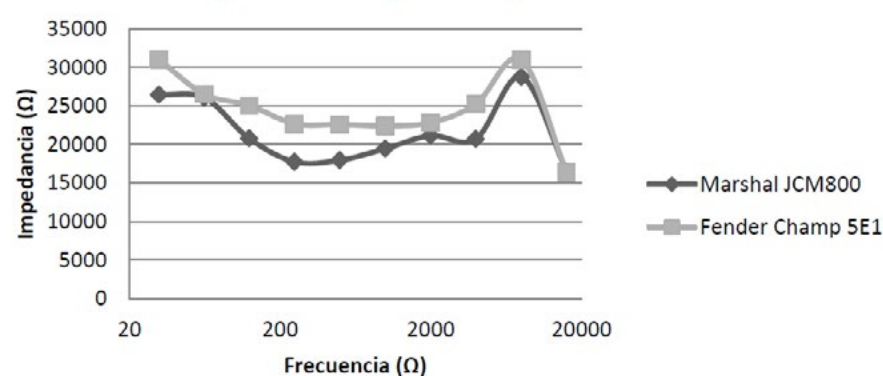


Fig. 13. Curvas de impedancia de los preamplificadores Fender Champ 5E1 y Marshall JCM 800.

A continuación se muestra la respuesta en frecuencia de los preamplificadores, analizados con ruido blanco (señal de entrada) Figura 14, y el cómo se comporta este al atravesar cada preamplificador.



Fig. 14. Respuesta en frecuencia ruido blando en el analizador de espectro.

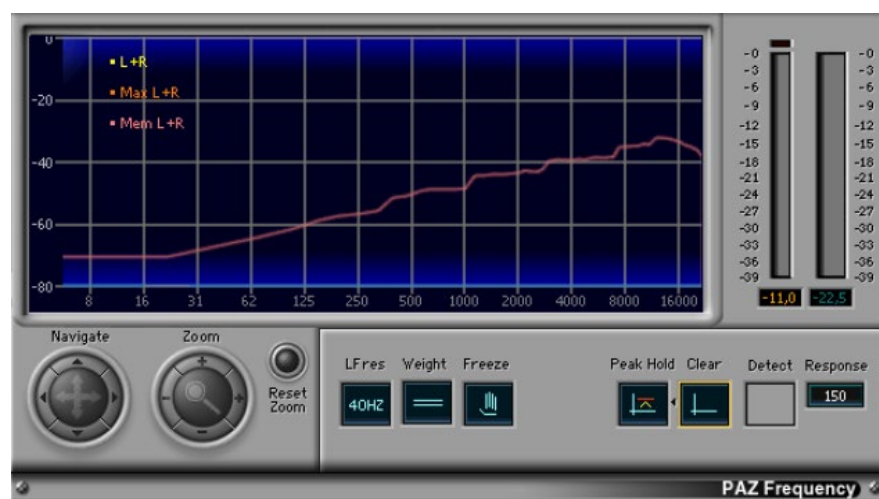


Fig. 15. Respuesta en frecuencia de preamplificador Fender Champ 5E1.



Fig. 16. Respuesta en frecuencia de preamplificador Marshall JCM 800.

La Tabla I muestra el resumen del parámetro ecualizador y como modifica la amplitud de la señal en su respectivo rango de frecuencias.

TABLA I
RESPUESTA EN FRECUENCIA DE LOS ECUALIZADORES.

	Graves 1M Ω	Medios 22k Ω	Agudos 250k Ω
Fender Champ 5E1	<60 Hz +/-2 dB	400 Hz +/-5 dB	125 +/-5 dB
Marshall JCM800	<260 Hz +/- 3 dB	800 Hz -6000 Hz +/-2 dB	1000Hz +/-4 dB

El análisis psicoacústico fue realizado a través de entrevistas a técnicos especializados, estas entrevistas fueron utilizadas

como retroalimentación para el investigador, sin embargo las opiniones, comentarios y sugerencias no son mostrados en este artículo al ser considerados subjetivos.

IV. Conclusiones

- El diseño y construcción de un solo circuito de un preamplificador a válvulas para guitarra con dos diferentes calidades tonales es completamente factible.

- Es posible establecer un solo circuito eléctrico que permita seleccionar entre dos diferentes preamplificadores, compartiendo circuitería en las etapas de rectificación y ecualización.

- La simulación mediante software no solo facilita el comprender cómo funciona la circuitería, sino también simplifica el análisis, la experimentación y la construcción.

- La implementación y el ensamblaje del circuito eléctrico es factible. Si bien estas son etapas que llevan mucho tiempo, cerca del cincuenta por ciento del tiempo de todo este proyecto fue destinado para el ensamblaje.

- La medición de parámetros técnicos del preamplificador no es del todo posible, puesto que actualmente no existe un referente que establezca las características técnicas que debe poseer un preamplificador de guitarra, así como tampoco existen metodologías únicas de medición de las distintas características. Sin embargo, la impedancia y la respuesta de frecuencia, son parámetros que pueden ser considerados, a partir de la presente investigación.

- La comparación entre los preamplificadores referenciales y contruidos fue considerada indispensable en la investigación. El mercado ecuatoriano está limitado por la escases y la demanda que presentan estos dispositivos. Por lo tanto la comparación eléctrica de los dispositivos resulta complicada, considerando las variables tiempo y dinero que implica el realizar estas mediciones. Por otro lado, la comparación psicoacústica es posible mediante grabaciones.

- Las diferencias en frecuencias medias y en graves del preamplificador son las principales características del preamplificador construido, como lo indican las percepciones de los especialistas entrevistados. Si bien no es exactamente igual al sonido del preamplificador JCM800 original, tienen sus similitudes, sin embargo debido a las apreciables diferencias, se termina hablando de un nuevo preamplificador con un nuevo sonido.

- Pese a que el preamplificador no cuenta con etapa de potencia, si es posible que este sea añadido a cualquier etapa de amplificación y funcionar en la práctica.

- Al alterar el circuito rectificador, se modifica directamente a la forma de onda resultante; otorgándole a esta mayor o menor amplitud y distorsión. En el caso del preamplificador 5E1 aumentó su amplitud y disminuyó su distorsión hasta ser prácticamente imperceptible visualmente en las simulaciones realizadas.

- La etapa de ecualización influye directamente en la forma de onda resultante. Su aporte es apreciable tanto en las simulaciones como en el osciloscopio, para el proceso que realiza el ecualizador en los diferentes preamplificadores.

- La etapa de ecualización compartida por los preamplificadores, no comparte las mismas frecuencias de corte, es decir, el realce o la atenuación de las frecuencias no son los mismos dependiendo del preamplificador que esté funcionando.

Villamar Villareal, J. J. (2011). Diseño y construcción de un preamplificador para micrófono con tecnología de tubos catódicos. Quito, Ecuador: Universidad de las Américas.

Zattola, T. (1995). Vacuum Tube Guitar And Bass Amplifier Theory. New York, USA: The Bold Strummer.

V. Referencias

Allen, D. (28 de Noviembre de 2010). Fender 5E1 Tweed Champ. Recuperado el 28 de Marzo de 2014, de Fender 5E1 Tweed Champ Project Page: <http://tknives.com/classact/tweedchamp.html/>

Álvarez, J. A. (marzo de 2012). Así Funciona. Recuperado el 5 de noviembre de 2013, de <http://www.asifunciona.com/>

Blencowe, M. (2009). Designing Valve Preamps for Guitar and Bass (Segundaed.). York, United Kingdom.

Carrión Isbert, A. (1998). Diseño acústico de espacios arquitectónicos. Barcelona, España: UPC.

Colombo, A. (6 de agosto de 2012). Universidad Nacional de la Patagonia San Juan Bosco. Recuperado el 6 de noviembre de 2013, de Facultad de Ingeniería: <http://www.ing.unp.edu.ar/electronica/asignaturas/ee016/>

Costas Soto, R. (19 de noviembre de 2008). Departamento de Tecnología Electrónica Universidad de Vigo. Recuperado el 7 de noviembre de 2013, de Convertidores AC/DC: <http://www.dte.uvigo.es/recursos/potencia/ac-dc/archivos/diodo.htm>

Dr. Tube Marshall Schematics. (01 de octubre de 2014). Recuperado el 01 de Octubre de 2014, de Dr. Tube: <http://www.drtube.com/library/schematics/69-marshallschemas#JCM800>

Fender. (1 de Octubre de 2014). Fender Musical Instruments . Recuperado el 1 de Octubre de 2014, de Fender Musical Instruments: <http://www.fender.com/es-ES/history/>

Jones, M. (2004). Building Valve Amplifiers (Primeraed.). Oxford, England: Newnes.

Kuehnel, R. (2009). Vacuum Tube Circuit Design: Guitar Amplifier Preamps (Segundaed.). Seattle, United States of America: Amp Books LLC.

Marshall Amplification. (1 de Octubre de 2014). Marshall Amplification. Recuperado el 1 de Octubre de 2014, de Marshall Amplification: <http://www.marshallamps.com/history/made-in-england>

Midge, M. &. (1 de Abril de 2008). Mhuss. Recuperado el 28 de Marzo de 2014, de Mhuss: <http://mhuss.com/MyJCM/90>

Miyara, F. (2004). Acústica y Sistemas de Sonido. Rosario, Argentina: UNR Editora.

Miyara, F. (7 de junio de 2006). FCEIA. Recuperado el 22 de septiembre de 2013, de Laboratorio de Acústica y Electroacústica: <http://www.fceia.unr.edu.ar/acustica/audio/electroac.pdf>

Pittman, A. (2004). The Tube Amp Book. China: Backbeat Books.

Rumsey, F., & McCormick, T. (2009). Introducción al Sonido y la Grabación. Madrid, España: IORTV.

AVANCES DEL AUDIO EN LATINOAMÉRICA.



Audio Engineering Society
Sección Colombia



METODOLOGÍA PARA LA MEDICIÓN DE DIRECTIVIDAD EN TIPLE COLOMBIANO

Fredy Alzate Arias / José Julián Cadavid

Fredy Alzate Arias*, José Julián Cadavid**

Facultad de Artes y Humanidades, Instituto Tecnológico Metropolitano - ITM

Medellín, Colombia

fredyalzate@itm.edu.co

jose.cadavid0335@correo.itm.edu.co

** Resumen— Los patrones de directividad de las fuentes naturales como los instrumentos de viento y cuerdas, por lo general no son tenidos en cuenta en el momento de posicionar uno o más micrófonos para realizar una grabación o para diseñar un espacio acústico de un lugar en el que se quiera ejecutar dichos instrumentos. Dichos procesos en la mayoría de los casos son realizados de manera empírica.*

En éste contexto, se evidencia la necesidad de realizar un estudio de cómo se comporta el patrón de directividad de un instrumento musical (en éste caso el tiple colombiano), en función de la frecuencia y de la posición alrededor del mismo. Dicha información puede ser utilizada tanto por productores musicales, como también para futuros desarrollos de simuladores que permitan visualizar comportamiento del patrón de directividad de un instrumento acústico o de un ensamble de los mismos, con el fin de analizar cual es la posición más recomendada para el uso de micrófonos.

Palabras clave— directividad, espectro sonoro, patrón de radiación, grabación, recintos acústicos

Abstract— The directivity patterns of natural sources such as wind instruments and strings , usually are not taken into account at the time of positioning one or more microphones for recording or to design an acoustic space of a place in which want to play these instruments. These processes in most cases are made empirically. In this context, the need for a study of how the directivity pattern of a musical instrument (in this case the Colombian tiple) , depending on the frequency and position around it behaves is evident . Such information can be used by both music producers , as well as for future development of simulators that allow viewing behavior directivity pattern of an acoustic instrument or an assembly thereof, in order to analyze what is the most recommended for the position use of microphones.

Keywords— directivity, sound spectrum, radiation pattern, recording, acoustic enclosures.

I. Introducción

La norma internacional para la realización de medidas acústicas ISO 3382-1, especifica que, para la realización de cualquier medición acústica, es necesario usar fuentes omnidireccionales. Por lo tanto, es necesario plantear una metodología para la medición de instrumentos acústicos, debido a que el sonido característico y el patrón de directividad, varía en diferentes direcciones; lo que modifica su contenido espectral en alrededor del espacio en el que se encuentra.

*

* Fredy Alzate A. labora en ITM- Medellín, Ing. Electrónico. (e-mail: fredyalzate@itm.edu.co).

** J. Julian Cadavid. Labora en ITM- Medellin, Ing. Electricista (e-mail: josecadavid0335@correo.itm.edu.co)

Actualmente el proceso de microfoneo en campo cercano de fuentes naturales, como instrumentos de viento o de cuerdas, se realiza sin un fundamento científico sobre el comportamiento físico y acústico de los instrumentos.

La directividad de un instrumento musical depende de múltiples factores, que van desde el tipo de material con el que está construido, la intensidad con el que se ejecute y en el caso de los instrumentos de cuerdas, no solo depende de la madera con la que fue construido, sino también, de el tipo de material de las cuerdas y el modo de pulsación utilizada por el ejecutante.

Hoy en día se encuentran artículos e investigaciones sobre: metodologías para la grabación de instrumentos de percusión característicos del folclor colombiano [1], medición multicanal de patrones de directividad de instrumentos de viento [2] [3], medición de directividad de instrumentos de cuerdas (guitarras y cuerdas frotadas) [4] [5], pero aún no se ha realizado un modelamiento físico y matemático de instrumentos autoctonos colombianos.

El objetivo de este artículo, es realizar una adaptación de la metodología empleada por Pätynen, Lokki y K.J.Bodon [6] [20] para la caracterización de la directividad de diversos instrumentos, pero empleando un recinto acústicamente controlado, como lo es un estudio de grabación y no una cámara anecoica; pues se busca analizar el comportamiento del patrón de directividad del tiple colombiano en un recinto que presente condiciones reales. Además de esto, la presente aproximación permite utilizar una menor cantidad de instrumentos de medida, para así adaptar el proceso a los equipos y recintos instalados en la Facultad de Artes y Humanidades del Instituto Tecnológico Metropolitano.

Lokki en su metodología propuesta para la medición de directividad de instrumentos acústicos de cuerdas y de viento, utiliza una cámara anecoica y en su interior un gran arreglo esférico de micrófonos que, junto con una interface de audio y un computador, permite hacer la captura multicanal de la energía acústica radiada por un instrumento determinado ubicado en el centro geométrico de dicho arreglo.

Para lo anterior, es necesario contar con una gran cantidad de micrófonos, cables para conexión y una estructura que permita ubicar cada micrófono en una posición específica alrededor del instrumento junto con su ejecutante.

La cantidad de señales y la resolución empleada en tiempo - amplitud para la conversión análoga-digital para las mismas, ocasionan un gran requerimiento de capacidad computacional para el procesamiento y análisis de toda la información producto de las mediciones. Por tanto, la metodología propuesta en el presente artículo, disminuye los requerimientos técnicos de manera notable, lo que permite realizar la medición y análisis de datos de una manera más sencilla y en un espacio con condiciones reales como lo es un estudio de grabación.

Es muy importante tener en cuenta los patrones de directividad de diversas fuentes acústicas, no solo para la realización de una grabación, el diseño de espacios acústicos adecuados para la ejecución de diversos instrumentos, sino también, para un adecuado refuerzo sonoro en un concierto.

Cabe anotar, que el presente artículo, es un primer acercamiento a la medición de la directividad de un tiple colombiano, pero a futuro, se espera desarrollar un software y hardware que permita simular y medir patrones de directividad de otros instrumentos musicales en tiempo real y en cualquier ambiente, además de desarrollar un dispositivo que permita parametrizar la directividad de una forma más sencilla, rápida y confiable.

II. Patrón De Radiación

Con el fin de realizar las mediciones necesarias para la parametrización de los patrones de directividad y distribución espectral de instrumentos acústicos, como lo es en este caso un tiple, es necesario realizar la mayor cantidad de datos posible alrededor del instrumento; ya que como se había mencionado anteriormente, un instrumento acústico no se comporta como una fuente sonora puntual omnidireccional.

Se tiene entonces que el patrón de radiación de una fuente, es el resultado de la solución a la prieda integral de la ecuación de Rayleigh [19] expresada de la siguiente forma:

$$P(\bar{x}, \omega) = g(\bar{r}, \omega) s(\omega) D(\bar{\varphi}, \bar{\theta}, \omega) \quad (1)$$

donde:

$\bar{x}$ es el campo de propagación del sonido radiado

ω es la frecuencia angular

$$g(\bar{r}, \omega) = \frac{e^{\frac{j\omega\bar{r}}{c}}}{\bar{r}}$$

(2)

$s(\omega)$ es la transformada de Fourier en el tiempo, de la señal de la fuente

$D(\bar{\varphi}, \bar{\theta}, \omega)$ es la función de directividad

y ω puede ser escrita como:

$$\omega = 2\pi f$$

(3)

donde f es la frecuencia.

Claramente se evidencia, como la función que describe el patrón de directividad de una fuente sonora, depende de la frecuencia y de la posición alrededor de la misma.

Es por esto que todo instrumento tiene una directividad distinta en cada frecuencia emitida.

A. Directividad

La directividad de una fuente sonora en una dirección, se define como la razón entre la presión eficaz generada por la fuente en esa dirección a la distancia r , y una presión eficaz de referencia $p_0(r)$; que suele ser la máxima que produce la fuente a esa distancia.

Existe una dependencia direccional entre la energía sonora radiada y la energía percibida por un oyente, a su vez, hay una dependencia a la frecuencia y hay una ecuación que las relaciona [16] [17].

$$D(\omega, \theta, \varphi)_r = \frac{p(\omega, \theta, \varphi)_r}{p(\omega, \theta_o, \varphi_o)_r} \quad (3)$$

Donde el numerador es la presión sonora radiada en cualquier punto de una esfera de radio r alrededor de la fuente en una posición que depende de ambos ángulos; mientras que el denominador es la presión sonora emitida captada a una distancia radial r , en ángulos de referencia

$$\varphi_o, \theta_o$$

Nuevamente, ω es la frecuencia angular tal cual se había expresado anteriormente.

Con la ecuación (3) es posible calcular cual es la pérdida de presión sonora relacionada según la frecuencia alrededor de una fuente a una distancia r .

Para el caso del presente artículo, se decidió hacer el estudio de la presión sonora en dB SPL, por tanto, la ecuación (3) quedaría expresada como:

$$D(\omega, \theta, \varphi)_r \text{ dB}_{SPL} = 20 \log \left(\frac{|p(\omega, \theta, \varphi)_r|^2}{|p(\omega, \theta_o, \varphi_o)_r|^2} \right) \quad (4)$$

Como resultado se puede obtener un diagrama polar donde la magnitud radial está dada por los decibeles de atenuación que tiene la presión sonora a medida que se cambia la posición del receptor alrededor del eje de la fuente.

La fuente direccionalmente más sencilla es una fuente esférica o mono polo, la cual radia con igual presión en todas las direcciones, de forma tal que la atenuación se da solo en función de la distancia a la fuente y esta disminuye a una razón de 6dB cada vez que la distancia se duplica. Esta es la fuente utilizada para realización de medición de parámetros acústicos de un cuarto y además, la más usada por los programas de simulación.

Sin embargo, las fuentes en las que se centra el presente artículo no pueden ser aproximadas a fuentes esféricas, debido a varios factores, tales como el tipo de madera usado para la construcción del tiple, el tipo de cuerdas empleada, y la técnica de pulsación usada por el ejecutante para excitar el instrumento.

Hay una distancia R_o a la cual, todas las fuentes pueden ser tratadas como fuentes puntuales, pues a partir de ahí la intensidad radiada disminuye respondiendo a la ley de la inversa del cuadrado de la distancia, por tanto, se puede concluir lo siguiente:

- Para distancias menores a R_o , la fuente no puede aproximarse a una fuente puntual, ésta zona recibe el nombre de campo cercano.
- Para distancias mayores o iguales a R_o , la fuente puede aproximarse a una fuente puntual, ésta zona recibe el nombre de campo lejano.

Para el desarrollo del presente trabajo, la distancia de interés fue 53 cm, con la cual se garantiza que en la intensidad sonora obtenida predomina el sonido directo de la fuente evitando saturaciones, efecto de proximidad y la captación mínima de las reflexiones producidas en las superficies del cuarto.

B. Trabajos Previos

Existen algunos trabajos previos que se han realizado a nivel científico y tecnológico sobre la parametrización y medición de directividad en los instrumentos musicales.

La primera aproximación a la recolección de una base de datos sobre la directividad de instrumentos musicales fue en la década de los setentas con Jürgen Meyer, quien realizó mediciones de la radiación promedio de varios instrumentos musicales por bandas de octava (Behler, POLLOW, Vorländer, & undefined, 2012) [9]. Más recientemente Otondo y Rindel, midieron la directividad usando un set de micrófonos ubicados alrededor del músico, utilizando 13 microfonos en el medio plano y en el plano horizontal (Polow, Behler, & Schultz, 2010) [10].

En 2010 Pätynen y Lokki (Pätynen & Lokki, 2010), usaron una ligera distribución de varios micrófonos con el fin de grabar fragmentos musicales [6] [7].

Los patrones de radiación de los instrumentos representan una parte considerable de la percepción de la acústica de un cuarto.

La directividad de varios instrumentos sinfónicos y un cantante soprano fueron investigados durante su interpretación. Para este propósito cada instrumento fue grabado por separado en una cámara anecóica con un arreglo de 22 micrófonos distribuido alrededor del músico. Los interpretes eran avisados de cambiar el tono que tocaban, mientras el instrumento emitía en dirección de su campana. El tocar con dinámica dio como conclusión que no causa cambios en la directividad según el espectro, en particular con los instrumentos de metales.

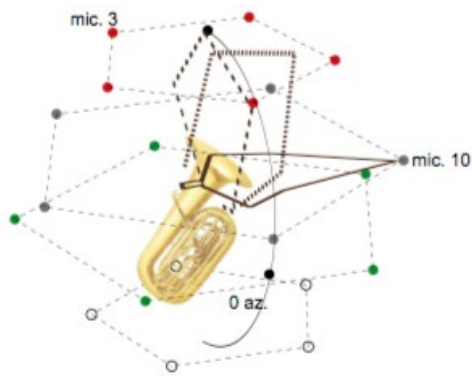


Fig. 2. Estructura para medición de directividad (Pätynen & Lokki, 2010)

Un arreglo esférico de 64 micrófonos fue construido en Graz, Austria, diseñado por Procesamiento óptico harmónico óptimo.

En el 2012 G.K Behler publicó un estudio con una metodología novedosa para la parametrización de directividad. Esta consiste en la construcción de una estructura esférica; esto con el fin de conseguir la radiación natural de instrumentos musicales, todas las direcciones deben ser medidas simultáneamente debido a la falta de repetibilidad.

Esto motivó la construcción de una esfera liviana, que fuera transportable, que contiene un arreglo de micrófonos. Se eligió como geometría para el arreglo una figura truncada de isocaedro, en el cual se distribuyeron 32 micrófonos de manera regular en toda la superficie de la estructura, la cual tiene un diámetro de 4,2 m.

La estructura consta de 90 varillas flexibles de fibra de vidrio que pueden ser unidas mediante holders de micrófonos formando una esfera.

El montaje completo fue realizado en una cámara anecóica con el fin de evitar cualquier rebote en paredes, techo y piso.

En el centro del arreglo de micrófonos se instala una silla de altura ajustable con la cual se busca ubicar al músico en el centro geométrico de la esfera.

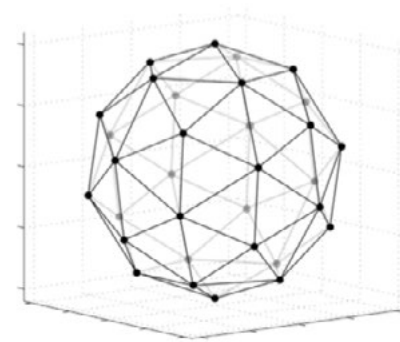


Fig. 2. Estructura Isocaedro para medición de directividad (Behler et al., 2012)

El 2015, en Colombia se realizó una investigación por C.A. Caballero y J.M. Moreno, en el que se especifica la metodología correcta para realizar la captura digital de sonido de instrumentos de percusión folclóricos con el objetivo de crear bancos de sonidos de instrumentos virtuales.

En ninguno de los casos, se cuenta con material relacionado con el estudio físico matemático de la directividad de el tiple colombiano.

C. Metodología

La metodología seleccionada para la medición y parametrización de la directividad de los instrumentos, en la presente investigación, es la utilizada por Pätynen y Lokki. En primer lugar, por la reducción de costos, ya que la elaboración de una esfera conlleva mucho más tiempo y recurso económico. Segundo, porque la metodología de Lokki puede ser implementada en un estudio de grabación con acústica variable y no necesariamente en una cámara anecóica; a la cual no se tiene acceso en el Instituto Tecnológico Metropolitano. Además, dicha metodología involucra una estructura más simple en cuanto a su montaje y mucho más flexible en cuanto la adquisición de datos con un software convencional como puede ser MatLab, C++ o una hoja de procesamiento numérico como excel.

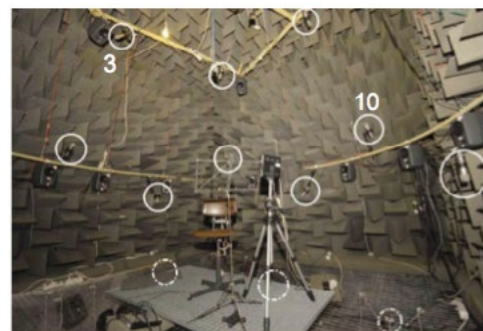


Fig. 3. Disposición de micrófonos en cámara anecóica (Pätynen & Lokki, 2010)

Para la medición del patrón de radiación de un tiple colombiano, se utilizó el equipo PAA6, el cual permite hacer medición de dB, RTA, FFT, entre otros.

Un primer acercamiento al patrón de radiación, se realizó mediante la medición con la función RTA del equipo.
Se dispuso del estudio de grabación de la facultad de Artes y Humanidades del ITM, el cual presenta un tiempo de reverberación media de 0,2s (tiempo de reverberación promedio en 500Hz y 1kHz); el cual es un valor optimo para la realización de las pruebas, debido a que hay muy poca energía reflejada.

Como una aproximación adicional a la metodología para la medición del patrón de directividad, se realizo la marcación sobre la superficie del piso, una circunferencia, cuyo radio es la distancia máxima desde el centro del instrumento. El centro del instrumento coincide con el centro de la circunferencia.
Dicha circunferencia se divide en ángulos iguales de 22,5º, en los cuales se realizan las mediciones respectivas en frecuencias características del instrumento.
Para esta primera etapa de medición fueron seleccionadas las siguientes frecuencias: 130Hz, 260Hz, 315Hz, 630Hz, 874Hz, 1kHz y 2,5kHz y a la misma altura del centro del instrumento.

Analizando los procedimientos para medir la directividad en muchos proyectos de investigación descritos en los trabajos previos,se sugiere la siguiente metodología para mejorar lo hecho hasta el momento ; una primera fase que mide de la directivdad en tiempo real y en una cámara anecoica, ver figura 4 (Norma AES 56-2008) y una segunda fase que describe el modelo físico del tiple con análisis numérico, más concretamente por diferencias finitas, además se harán medidas de la resonancia del instrumento, analizando los modos de vibración, la figura 5 muestra el procedimiento

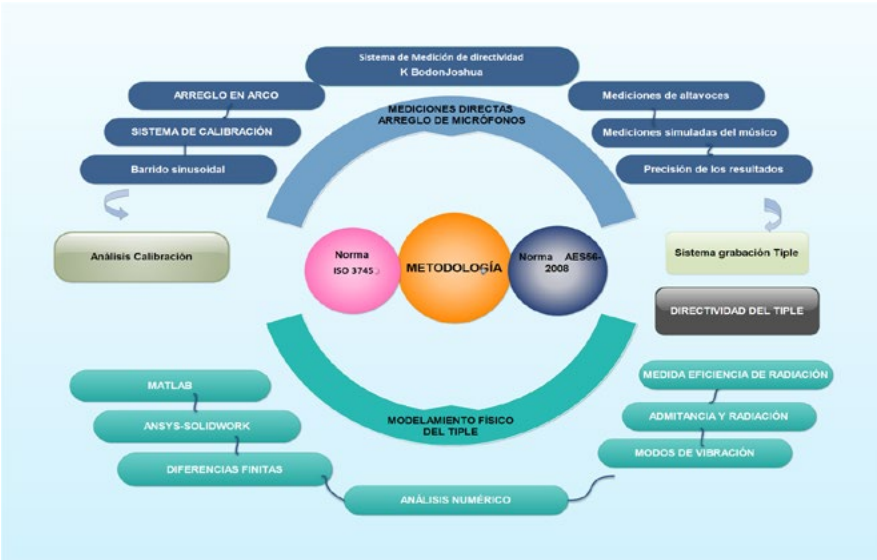


Figura 4 Metodología para medir directividad en el tiple

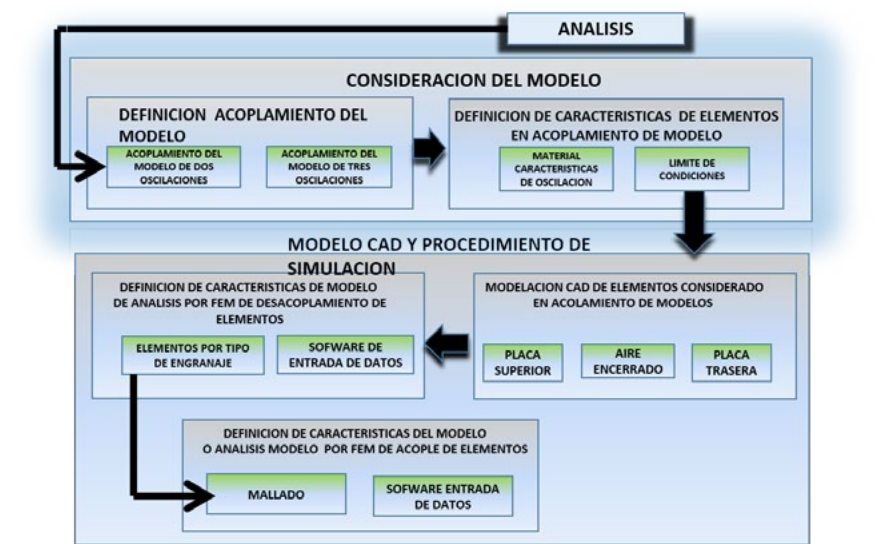


Figura 5. Simulación Análisis diferencias finitas

D. Equipos Requeridos

- 2 sonómetros PAA6
- Medidor de distancia laser
- Analizador de espectro
- Cables y conectores
- Audifonos
- 2 Tripodes

E. Calibración de Equipos

Con el fin de verificar el correcto estado de calibración de los equipos de medición PAA6, ambos fueron ubicados equidistantes al centro acústico del instrumento utilizando los trípodes antes mencionados. Con esto fue posible constatar una desviación en su medición menor al 0,32%. Debido a que cada uno de estos equipos cuenta con dos microfones para medición, realizamos el análisis estadístico de las 4 mediciones realizadas.

F. Proceso de grabación

Cada equipo de medición fue ubicado a una distancia de 0,53m tomados desde el centro del instrumento y a una altura de 70cm respecto al suelo.

TABLA I

Posición de microfones implementada (Pätynen, Pulkki, & Lokki, 2008),(Behler et al., 2012)

Angulos de Azimut	Elevación Grados
0º	0º
22,5º	0º
45º	0º
67,5º	0º
90º	0º
112,5º	0º
135º	0º
157,5º	0º
180º	0º
-22,5º	0º
-45º	0º
-67,5º	0º
-90º	0º
-112,5º	0º
-135º	0º
-157,5º	0º
-180º	0º

Se le pide al musico ejecutante que pulse la nota respectiva a cada frecuencia antes mencionada, con una intensidad constante. Cada medición es tomada dos oportunidades con los dos equipos de medición con el fin de hacer un promedio durante el análisis de datos.

En primer lugar, se realiza la medición del piso de ruido antes de hacer la medición en cada frecuencia, con el fin de sustraer algebraicamente dicho valor y obtener la medición de la fuente de interés.

Los datos adquiridos son almacenados en un archivo de texto plano separado por comas, con el fin de luego ser usados para análisis y graficos en forma bidimensional polar.

III. Resultados

Los datos son presentados de manera gráfica y detallada inicialmente con el programa Excel. En una etapa posterior de la investigación, se buscará desarrollar un hardware que permita realizar la medición a varias alturas y angulos, además de un software que mida en tiempo real los patrones de directividad y permita almacenar los datos con el objetivo de realizar simulaciones no solo de instrumentos individuales como el presente caso, sino también de arreglos de instrumentos.

Con los datos anteriores se van conformando los diferentes diagramas polares de cada frecuencia, con el fin de crear una base de datos para la mayor cantidad de frecuencias posibles, comenzando por las más bajas de un tiple. En las siguientes tablas se muestran algunos resultados obtenidos por el grupo de investigación en ingeniería de audio del ITM:

TABLA II
 Datos Obtenidos Para 130Hz de Tiple Colombiano ITM-Medellin (Elaboración propia)

Angulo	nivel promedio sonoro dB	ruido promedio de fondo dB	nivel promedio sonoro de la fuente dB
0	46,4	35,8	43,3622924
22,25	52,8	18,5	52,63094272
45	52,9	27,3	52,43176088
67,5	53,5	35,5	52,33129652
90	50,9	40,6	47,73354366
112,5	49,9	28,9	49,08917476
135	41	28,5	38,64892681
157,5	45,5	28,2	44,22566615
180	54,7	31,2	54,09917449
-22,25	48,2	34,2	46,26694209
-45	49,3	31,7	48,07213877
-67,5	50,3	30,6	49,35086814
-90	53,7	28,8	53,19129035
-112,5	51,8	29,7	51,08968468
-135	52,6	22,3	52,33051537
-157,5	48,4	27,4	47,58917476

Fig. 4. Patrón de radiación obtenido para tiple colombiano en la frecuencia de 130Hz

TABLA III
 Datos Obtenidos Para 2,5kHz de Tiple Colombiano ITM-Medellin (Elaboración propia)

Angulo	Nivel promedio sonoro dB	ruido promedio de fondo	nivel promedio sonoro de la fuente
0	73,3	33,2	73,21370814
22,25	51,3	28,4	50,65456846
45	63,1	34,1	62,78621269
67,5	57,7	27,1	57,43980257
90	72,4	22,2	72,37311647
112,5	44,1	22,5	43,3457159
135	45,3	30,5	43,55538371
157,5	48,6	31,7	47,26073736
180	50,2	36,4	48,21636472
-22,25	48,6	35,7	46,36960577
-45	38,9	36,9	25,16349351
-67,5	51,3	29,3	50,58110972

-90	46,2	16,6	45,90751341
-112,5	47,6	36,3	44,83940401
-135	45,7	28,7	44,37729703
-157,5	53,6	30,1	52,99917449

Fig. 5. Patrón de radiación obtenido para tiple colombiano en la frecuencia de 2,5kHz

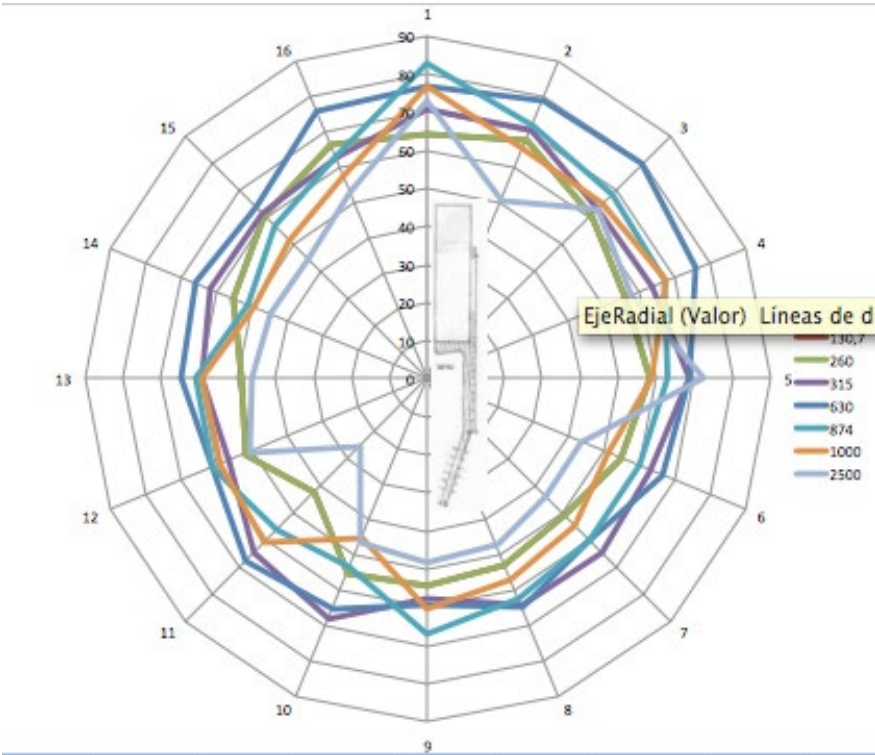


Fig. 6. Patrón de radiación obtenido para tiple colombiano en las frecuencias s especificadas

IV. Conclusiones

En el presente trabajo se puede concluir que para obtener una mayor presición en los datos, es necesario realizar un tratamiento estadístico de una gran muestra de mediciones; es decir es necesario realizar un gran número de medidas del instrumento interpretado por varios ejecutantes.

Las mediciones son dependientes de un sinnumero de variables, una de ellas es la fuerza con la que se toque el instrumento (dinámica) que puede afectar de cierta forma las medidas obtenidas; lo cual debe ser tenido en cuenta en el tratamiento de errores.

Debido a que la metodología utilizada para la presente investigación, es una aproximación a la propuesta por Pätynen y Lokki, no es posible hacer una comparación directa con los resultados encontrados por estos en los patrones de radiación de otros instrumentos. Como segunda etapa de la presente investigación, se desarrollará un hardware que permita una medición más cercana a los modelos originales. Además, la intensión inicial de la presente investigación es realizar la medición en un recinto que presente condiciones acústicas reales; ya que las ejecuciones de instrumentos musicales no se realizan en condiciones acústicas completamente controladas como el caso de una cámara anecóica, cómo lo expresado por F. Otondo [18].

Con los datos obtenidos de este trabajo, es posible llegar a hacer una comparación entre fabricantes y referencias de instrumentos musicales acerca de la efectividad acustica de los materiales con los que están contruídos, además de

realizar una evaluación crítica entre fabricantes y posible bench marking.

Este trabajo es un aporte no solo a los ingenieros de sonido para conocer la mejor manera de capturar los instrumentos, sino también a ingenieros de audio que mediante el conocimiento de los patrones de directividad de instrumentos musicales y ensambles de estos pueden diseñar lugares acusticamente mejor tratados.

Todos los datos obtenidos en este proyecto pueden ser utilizados en la consecución de un algoritmo para el desarrollo de un software que haga simulaciones de instrumentos musicales no solo virtualmente, sino también por medio de hardware mediante el uso de un dodecaedro.

V. Referencias

- [1] Carlos A. Caballero & Jamir M. Moreno (2015). Parámetros técnicos de captura en instrumentos musicales percutidos del folclor colombiano para su uso en bancos virtuales de sonidos, AES-9444 October 23, 2015. <http://www.aes.org/e-lib/browse.cfm?elib=18000>
- [2] Fernandez Comesana, D., Takeuchi, T., Morales Cervera, S., & Holland, K. R. (2012b, julio). Measuring musical instruments directivity patterns with scanning techniques. Conference presentado en 19th International Conference of Noise and Vibration. Recuperado a partir de <http://eprints.soton.ac.uk/346037/>
- [3] Fernandez Comesana, D., Takeuchi, T., Morales Cervera, S., & Holland, K. R. (2012a). Measuring musical instruments directivity patterns with scanning techniques. Recuperado a partir de http://eprints.soton.ac.uk/346037/1/ICSV_2012_Measuring_musical_instruments_directivity_patterns_with_scanning_techniques.pdf
- [4] Otondo, F., Rindel, J. H., Causse, R., Misdariis, N., & de la Cuadra, P. (2002). Directivity of musical instruments in a real performance situation. En *Proceedings of the International Symposium on Musical Acoustics (ISMA)*, Mexico City, Mexico (pp. 312–318). Recuperado a partir de http://www.otondo.net/research/man_isma_final.pdf
- [5] D. F. Comesaña, T. Takeuchi, S. M. Cervera, and K. R. Holland, “Measuring musical instruments directivity patterns with scanning techniques,” in *ResearchGate*, 2012, vol. 3.
- [6] Pätynen, J., & Lokki, T. (2010). Directivities of Symphony Orchestra Instruments. *Acta Acustica united with Acustica*, 96(1), 138-167. doi:10.3813/AAA.918265
- [7] Pätynen, J., Pulkki, V., & Lokki, T. (2008). Anechoic Recording System for Symphony Orchestra. *Acta Acustica United with Acustica*, 94(6), 856-865. doi:10.3813/AAA.918104
- [8] Bassuet, A. A., Caulkins, T., Digerness, J. K., & Knickrehm, G. A. (2013). New implementation of multi-microphone recording technique to simulate music ensembles in room acoustic auralizations. En *Proceedings of Meetings on Acoustics* (Vol. 19, p. 015135). Acoustical Society of America. Recuperado a partir de <http://scitation.aip.org/content/asa/journal/poma/19/1/10.1121/1.4801040>
- [9] Behler, G., & Pollow, M. (2009). Representation of the musical instruments directivity using dodecahedron loudspeakers. *The Journal of the Acoustical Society of America*, 125(4), 2543-2543. doi:10.1121/1.4783611
- [10] Behler, G., POLLOW, M., Vorländer, M., & undefined, others. (2012). Measurements of musical instruments with surrounding spherical arrays. *Acoustics 2012 Nantes*. Recuperado a partir de <http://hal.archives-ouvertes.fr/hal-00811213/>
- [11] Felipe Otondo, J. H. R. (2002). New Method for the Directional Representation of Musical Instruments in Auralizations. *International Computer Music Conference Proceedings*. Recuperado a partir de <http://hdl.handle.net/2027/spo.bbp2372.2002.051>
- [12] Fletcher, N. H., & Rossing, T. D. (1998). Bowed String Instruments. En *The Physics of Musical Instruments* (pp. 272-330). Springer New York. Recuperado a partir de http://link.springer.com/chapter/10.1007/978-0-387-21603-4_10
- [13] Forum Acusticum, D. A. S. (2011). Proceedings of Forum Acusticum 2011, 27 June-01 July, Aalborg, Denmark. [S. l]: European Acoustics Association.
- [14] Karjalainen, M., Huopaniemi, J., & Välimäki, V. (s. f.). of the paper. Recuperado a partir de <http://www.acoustics.hut.fi/~mak/PUB/ICA95IR.pdf>
- [15] Kristiansen, U., Støfringsdal, B. ard, Svensson, P., & Tro, J. (2001). Performance control and virtualization of acoustical sound fields related to musical instruments. En *Workshop on Current Research Directions in Computer Music*. Barcelona: Audiovisual Institute, Pompeu Fabra University. Recuperado a partir de <http://www.academia.edu/download/30493373/p46.pdf>
- [16] Pasqual, A. M. (2010). Sound directivity control in a 3-D space by a compact spherical loudspeaker array. *Universidade Estadual de Campinas*. Recuperado a partir de <http://tel.archives-ouvertes.fr/tel-00530855/>
- [17] Pollow, M., Behler, G. K., & Schultz, F. (2010). Musical instrument recording for building a directivity database. *Fortschritte der ANustiN: Tagungsband d, 36*. Recuperado a partir de http://www.researchgate.net/publication/233762472_Musical_Instrument_Recording_for_Building_a_Directivity_Database/file/9c960516683598bc5d.pdf
- [18] F. Otondo and J. H. Rindel, “Directivity of musical instruments in a real performance situation,” *ResearchGate*, Dec. 2002.
- [19] M. Pollow, *Los Patrones de directividad de la Sala de Mediciones Acústicas y Simulaciones*. Place of publication not identified, 2015.
- [20] K. J. Bodon and T. W. Leishman, “Development, evaluation, and validation of a high-resolution directivity measurement system for live musical instruments,” *The Journal of the Acoustical Society of America*, vol. 138, no. 3, pp. 1785–1785, Sep. 2015.



AVANCES DEL AUDIO EN LATINOAMÉRICA.



Audio Engineering Society
Sección Colombia



COMPARACIÓN ENTRE DIFERENTES FRECUENCIAS DE MUESTREO Y PROFUNDIDAD DE BITS EN EL DESARROLLO DE UNA PRODUCCIÓN MUSICAL

Sebastián Lopera Gómez

Sebastián Lopera Gómez*
**Facultad de Ingenierías, Universidad de
San Buenaventura Medellín**
Medellín, Colombia

Resumen—De acuerdo a las actuales posibles configuraciones de los conversores AD/DA, se planteó la pregunta si es realmente beneficioso usar parámetros superiores a los establecidos en el “libro rojo” para disco compacto de audio digital. Se realizó la misma producción musical con distintas configuraciones de frecuencia de muestreo y profundidad en bits durante los procesos de grabación, mezcla y masterización, obteniendo archivos de audio que cumplan con los parámetros establecidos para los discos de compactos de audio digital. Mediante pruebas objetivas y subjetivas se determinaron las diferencias existentes y perceptibles causadas por las diferentes configuraciones usadas en el desarrollo de las producciones musicales.

Palabras clave— audio digital, conversión AD/DA, grabación, producción de audio.

Abstract— According to current possible configurations of AD / DA converters, the question was raised whether it is really beneficial to use higher than those set in the “red book” parameters for digital audio disc. A musical production was done with different configurations of sample rate and bit depth during the process of recording, mixing and mastering, obtaining audio files that meet the parameters for digital audio compact disc. Applying objective and subjective tests the existing and perceptible differences caused by the different configurations used in the development of musical productions were determined.

Keywords— Digital audio, AD/DA conversion, recording, audio production.

I. Introducción

En 1982, las empresas Philips y Sony publicaron el “libro rojo” (Red Book), que establece las características técnicas que deben cumplir los CD-DA (disco compacto de audio digital). Este estándar especifica las propiedades que se deben cumplir, tanto físicas como de codificación del audio digital, estableciendo que el audio debe estar codificado mediante el sistema PCM (Pulse Code Modulation o Modulación por Impulsos Codificados), con una frecuencia de muestreo de 44.1 kHz y una profundidad en bits de 16. Con esta cantidad de muestras, se puede garantizar el óptimo registro y reproducción de frecuencias inferiores a 22.05 kHz, de acuerdo con el teorema de Nyquist-Shannon y con ese valor de profundidad de bits se asegura un rango dinámico de 96 dB. Aunque este estándar se ha mantenido hasta la actualidad, se desarrollan nuevos equipos y formatos que usan parámetros superiores (Rumsey, 2007). Estos nuevos avances se basan en la idea que esta cantidad de muestras por segundo y niveles de cuantificación son insuficientes, generando gran cantidad de distorsión al momento de digitalizar una señal de audio. Es por esta razón que los fabricantes hoy en día producen conversores con frecuencias de muestreo y profundidad en bits

superiores a los establecidos en el “libro rojo”, alcanzando frecuencias de muestreo de 352.8kHz, equivalente a ocho veces la frecuencia de muestreo que se usa actualmente para discos compactos de audio digital. Es de vital importancia reconocer la existencia de nuevos formatos existentes que trabajan con configuraciones diferentes a las establecidas para discos de audio digital, pero este trabajo se ha delimitado en hacer una comparación de acuerdo a los parámetros establecidos en el “libro rojo”.

Aunque la señal digitalizada tiene mayor similitud con la señal análoga al aumentar estos dos parámetros, posteriormente deben sufrir un proceso de disminución de muestras por segundo (downsampling) a 44.1 kHz y re-cuantificación de niveles a 16 bits, para así poder cumplir con los estándares del “libro rojo” y es por esto que surge la pregunta de investigación:

¿Es verdaderamente beneficioso realizar una producción musical usando archivos de audio, hardware de conversión análogo-digital y software de procesamiento con parámetros superiores a los establecidos en el “libro rojo”, o es simplemente uno de los tantos mitos que se han creado respecto al audio digital?

Al comparar frecuencias de muestreo y profundidades de bits, de manera teórica es evidente que existe una diferencia al usar diferentes parámetros logrando variar el ancho de banda del sistema y su rango dinámico. Hay estudios previos que evalúan si los individuos de una prueba diferencian entre programas musicales grabados a distintas frecuencias de muestreo, obteniendo que las personas con experiencia están en la capacidad de diferenciar entre diferentes frecuencias de muestreo (Pras & Guastavino, 2010). Un análisis cualitativo de manera verbal a los mismos sujetos indica que perciben diferencias en reproducción espacial, contenido de frecuencias altas, timbre y precisión del material evaluado.

Algunos autores defienden el uso de frecuencias de muestreo altas, como Wiesław Woszczyk que explica:

“El audio de alta definición ha sido sesgado al ancho de banda amplio y que, dada la limitada sensibilidad del oído humano a las frecuencias altas, se gana poca resolución perceptible. No hay mucha evidencia de laboratorio para contrarrestar esto, ya que la investigación psicoacústica se ha limitado en gran parte a estudiar los efectos en el rango de frecuencias entre los 20 Hz a 20 kHz.” (Woszczyk, 2003)

Además se muestra como la alta resolución temporal, espectral, espacial y dinámica en conjunto determinan la calidad de la percepción de la música y el sonido, y que la resolución temporal puede ser la más importante en la percepción del ser humano (Woszczyk, 2003).

Aunque sí existen diferencias, muchos usuarios de sistemas de audio digital se preguntan si trabajar con frecuencias de muestreo altas es más beneficioso que con frecuencias de muestreo bajas. Otros autor plantea esta pregunta y responde:

“Falso, este tema es probablemente lo que más exageran para realizar mercadeo, se aprovechan de las conclusiones simples e intuitivas ilustradas por la gran mayoría de personas, donde más es mejor.” (Aldrich, 2005)

* * Sebastián Lopera Gómez labora en la Universidad de San Buenaventura Medellín, Carrera 56 C # 51 – 110, Of. 409C. (e-mail: esp.postproduccion@usbmed.edu.co).

Dan Lavry también expone algunos mitos que ha encontrado a lo largo de su carrera, uno de ellos describe que la imagen estéreo se ve afectada al tomar un menor número de muestras, refutando esta idea asegurando que no existe la velocidad de muestreo correcta respecto a la señal. Cuando se realizan los procesos de conversión de archivos estéreo o multicanal, todos los archivos son muestreados de manera simultánea, asegurando que la relación de tiempo entre los canales se mantiene y no se ve afectada. (Lavry, The Optimal Sample Rate for Quality Audio, 2012)

Algo cierto y explicado por Dan Lavry en otro documento es que, al trabajar con archivos de audio con frecuencias de muestreo altas, se generan archivos más grandes que requieren de más espacio en disco para su almacenamiento y ralentiza la transmisión, requiriendo a su vez mayor velocidad de procesamiento computacional. Asimismo, el autor expone otro de los mitos que ha encontrado en el medio que consiste en grabar con una frecuencia de muestreo alta (192 kHz) y luego realizar un proceso de downsampling para cumplir con la frecuencia de muestreo establecida en el “libro rojo” (44.1 kHz). Frente a este mito el autor explica que lo único que se obtiene es un mayor ancho de banda, pero que los instrumentos musicales producen poca información en frecuencias alrededor de los 96 kHz, muy pocos micrófonos responden en frecuencias tan altas, los altavoces no las reproducen y por último, el oído de los seres humanos no responde a estas. (Lavry, Sampling Theory For Digital Audio, 2004).

Algunos defensores dicen que parte de la calidad que se obtiene al trabajar con 192 kHz se retiene al disminuir a una frecuencia de muestreo de 44.1 kHz ignorando que cualquier material muestreado a esa frecuencia, no puede contener información con frecuencia superior a 22,05 kHz. Dan Lavry manifiesta que al muestrear señales con frecuencias de muestreo de 192 kHz, que es tres veces mayor que la óptima, se compromete la precisión del conversor que se ve reflejado en distorsión. También afirma que el verdadero error es cuando la gente que escucha algo que les gusta, asocian mejor sonido con un muestreo más rápido, mayor precisión y mayor ancho de banda. (Lavry, Sampling Theory For Digital Audio, 2004)

La ausencia de información acerca de comparaciones entre profundidad en bits es notoria, un autor analiza esto y plantea una pregunta similar, pero esta vez analizando los niveles de cuantificación, afirmando esto como falso y que este tipo de mitos se plantean a partir de conclusiones intuitivas e información que permite obtener respuestas incorrectas. (Aldrich, 2005)

II. Parte Técnica del Artículo

A. Definición De Configuraciones A Evaluar

Como se mencionó anteriormente los parámetros para disco de audio digital descritos en el “libro rojo” establecen que los archivos de audio deben tener una frecuencia de muestreo de 44.1 kHz y una profundidad de bits igual a 16 bits. Teniendo en cuenta esto se realizaron dos comparaciones: una variando la frecuencia de muestreo y otra la profundidad en bits.

1) Variación De La Frecuencia De Muestreo

Actualmente los conversores A/D y D/A pueden trabajar con frecuencias de muestreo de 192 kHz, incluso hay algunos que permiten valores de 384 kHz. Para las comparaciones se establecen dos configuraciones posibles, una de acuerdo a lo especificado en el “libro rojo” y otra con una frecuencia de muestreo de 176.4 kHz que equivale a cuatro veces la establecida para disco de audio digital. La tabla 1 muestra las características de cada una de las configuraciones. Sabiendo que el parámetro a evaluar es la frecuencia de muestreo es importante mantener constante la profundidad de bits y únicamente cambiar lo que se desea evaluar.

TABLA 1
Características de cada configuración al variar la frecuencia de muestreo

Frecuencia de muestreo	Frecuencia máxima a muestrear	Tiempo entre muestras
44.1 kHz	22,05 kHz	22,68 μs
176.4 kHz	88,2 kHz	5,67 μs

2) VARIACIÓN DE LA PROFUNDIDAD DE BITS

A la fecha no existen conversores de audio que permitan trabajar con una profundidad de bits mayor a 24 bits. Aunque muchos de los sistemas D.A.W permiten crear archivos con niveles superiores, esto solo es útil para el procesamiento interno de las señales de audio. Teniendo en cuenta que los conversores en la actualidad sólo permiten dos configuraciones, los parámetros que se compararon se muestran en la tabla 2, junto a las características técnicas de cada una. Similar a la variación de la frecuencia de muestreo, el parámetro que no se compara se mantuvo fijo y solo se alteró la variable a evaluar que en este caso sería la profundidad de bits.

TABLA 2
Características de cada configuración al variar la profundidad de bits

Profundidad de Bits	Rango dinámico Aproximado	Niveles de Cuantización
16 bits	96 dB	65.536
24 bits	144 dB	16.777.216

Teniendo en cuenta las configuraciones previamente descritas, se realizó la captura de mínimo cuatro programas musicales con las configuraciones que se muestran en la tabla 3.

TABLA 3
Nombre de cada captura del programa musical y su configuración

Nombre de cada configuración	Frecuencia de Muestreo	Profundidad de Bits
A	44.1 kHz	24 bits
B	176.4 kHz	24 bits
C	44.1 kHz	16 bits
D	44.1 kHz	24 bits

Las capturas A y B corresponden a un conjunto de señales grabadas que permitirán evaluar los cambios de frecuencia de muestreo ya que solo se cambió la frecuencia de muestreo durante el downsampling como se ilustra en la Figura 1.

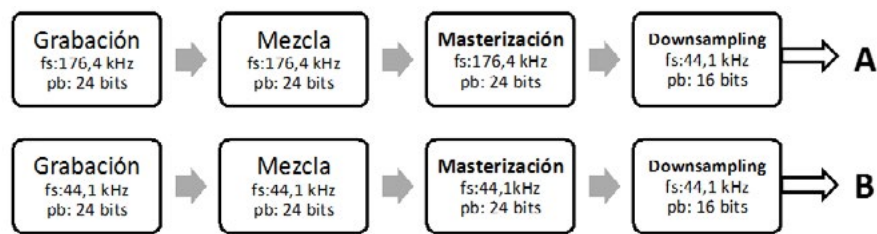


Figura 1. Etapas de la producción musical de las muestras A y B.

Mientras que las capturas C y D se usarán para evaluar los cambios de profundidad de bits, la configuración de frecuencia de muestreo (fs) y profundidad de bits (pb) durante las etapas durante la producción musical se aprecian en la Figura 2. Es por esto que, aunque las capturas A y C tienen la misma configuración el programa musical que será ejecutado es completamente diferente.

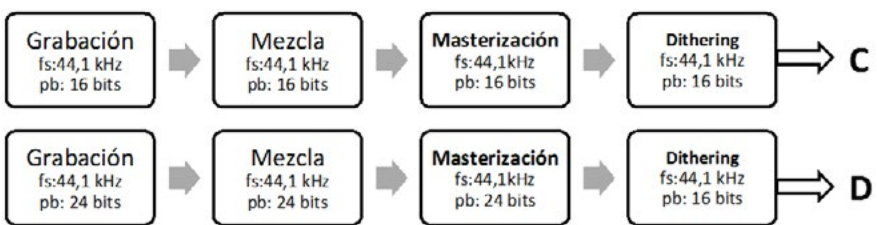


Figura 2. Etapas de producción musical de las muestras C y D

B. Parámetros A Evaluar

A continuación, se presentan algunos parámetros que permiten conocer las características propias de cada configuración.

1) Respuesta En Frecuencia

Se refiere al rango de frecuencias que el sistema es capaz de grabar y/o reproducir. Se puede encontrar representado en una imagen que describe la amplitud de la señal en cada una de las frecuencias, o como un rango de valores entre los que varía la señal, en un conjunto de frecuencias deseado.

2) Nivel De Ruido

Indica la amplitud del ruido propio del sistema causado por el diseño de los componentes electrónicos, su interacción y las conexiones de los distintos elementos. Se expresa en dB y debe indicarse si se usó alguna curva de ponderación para expresar este valor.

3) Rango Dinámico

Es la diferencia que hay entre el nivel de amplitud más alto que puede alcanzar el sistema antes de generar saturación y la amplitud del ruido de fondo. Se expresa en dB y debe indicarse si se usó alguna curva de ponderación.

4) Distorsión Armónica Total

Se produce cuando la señal que sale de un sistema no es igual a la señal que entró, esta ausencia de linealidad en los equipos de audio introduce armónicos que son múltiplos de la señal. El porcentaje de distorsión armónica total (THD%) cuantifica la distorsión armónica que entrega el sistema usando la ecuación (1) donde A_n equivale a la amplitud del armónico n y A_1 la amplitud de la frecuencia fundamental.

$$THD\% = \sqrt{\frac{\sum_n (A_n)^2}{A_1^2}} \times 100; \quad (1)$$

Luego de conocer la metodología planteada para las pruebas

objetivas, es importante definir algunas características acerca del programa musical que permitió obtener las muestras de audio que posteriormente serán presentadas en pruebas subjetivas.

C. Características del Programa Musical

Al revisar diferentes investigaciones que involucren pruebas subjetivas con programas musicales, no hay algún género que haya sido definido como estándar y suelen usar como señales de prueba interpretaciones de instrumentos individuales. Debido al planteamiento del problema, no se considera apropiado el uso de la interpretación de un solo instrumento musical, pues se busca responder si existe alguna diferencia al usar otros parámetros a los descritos en el “libro rojo” en procesos de grabación, mezcla y masterización de una producción musical, que finalmente cumplirá con los estándares definidos para disco compacto de audio mediante técnicas de *downsampling* y *dithering*.

Teniendo en cuenta las variaciones que producen en el sistema cambiar la cantidad de muestras por segundo y la profundidad de bits, se plantea que el programa musical debe tener una respuesta en frecuencia amplia, cambios de dinámica a través del tiempo y que posea instrumentos musicales a los que el oído humano este acostumbrado, haciendo que sea más fácil identificar si existen diferencias mediante pruebas subjetivas. Otro aspecto que debe considerarse es la presencia de instrumentos de percusión que posiblemente generen un reflejo estapedial.

El reflejo estapedial consiste en la tensión del músculo estapedio de manera involuntaria como consecuencia de un sonido de alta intensidad. Este reflejo es de carácter bilateral, es decir que cuando la onda sonora llega a uno de los dos oídos el reflejo estapedial ocurre en ambos oídos. El umbral de este reflejo se ubica a 85 dB por encima del umbral de audición del oyente en un rango de frecuencias que va desde los 50 Hz hasta los 4 kHz, esta contracción hace que cuando exista un aumento de la presión sonora en 10 dB para el oído interno solo sean 3 dB. (Feeney, 2014)

El flujo de señal para la captura del programa musical se aprecia en la Figura 3. Este mismo se usó para las pruebas objetivas de las distintas configuraciones.

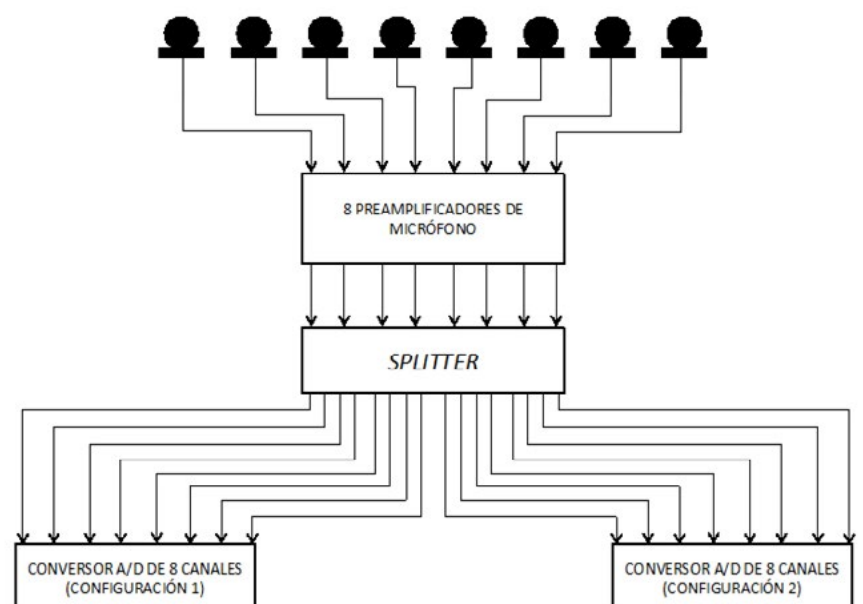


Figura 3. Diagrama de bloques del flujo de señal análogo

D. Pruebas Subjetivas

A la fecha no se ha establecido un método estándar para la realización de pruebas subjetivas, sin embargo, existe documentación sobre algunas pruebas realizadas y recomendaciones que permiten definir algunos aspectos importantes de cada prueba. Teniendo en cuenta esto se usaron como guía las recomendaciones (UIT, 1994-1997) y (UIT, 1997 - 2003), que brindan información útil para la selección de los oyentes, presentación de las muestras de audio, dispositivos de reproducción y características del lugar de escucha. Aunque estas recomendaciones no son una norma, sirven como guía para el diseño y realización de las pruebas subjetivas.

Aunque existen diferentes formas de presentar las muestras de audio a ser evaluadas se optó por realizar una comparación de dos sistemas con referencia, donde una de las dos muestras es la referencia y la otra es la muestra a comparar. Estas muestras se presentan en el siguiente orden: primero la secuencia de referencia, luego la segunda secuencia, posteriormente se repite la secuencia de referencia y por último se presenta la segunda secuencia, como se muestra en la Figura 4. La prueba subjetiva se realizó en un lugar que cumpliera las condiciones establecidas en las recomendaciones y que contará con un sistema de altavoces para la realización de la prueba. Ninguno de los sujetos tuvo conocimiento previo sobre la finalidad o justificación de la prueba, simplemente se le plantearon preguntas donde el indicará si encontraba alguna diferencia entre las muestras de audio y en que caso de ser afirmativo, indicar cual era su apreciación.



Figura 4. Reproducción de las muestras de audio en la prueba subjetiva

III. Resultados

A. Pruebas Objetivas

Para cada uno de los parámetros que se evaluó, se realizó una tabla comparativa que permite identificar de manera rápida las diferencias entre las diferentes configuraciones. El flujo de señal usado para la captura de las señales de prueba se muestra en la Figura 5. Este software de medición permite disminuir las alteraciones que produce la interfaz de adquisición de datos realizando una captura previa de sus propiedades. Al revisar estos datos, es importante tener en cuenta que todos los audios cumplen con las características exigidas por el “libro rojo” así hayan sido capturadas con configuraciones superiores, pues este trabajo se centra en estos parámetros establecidos.

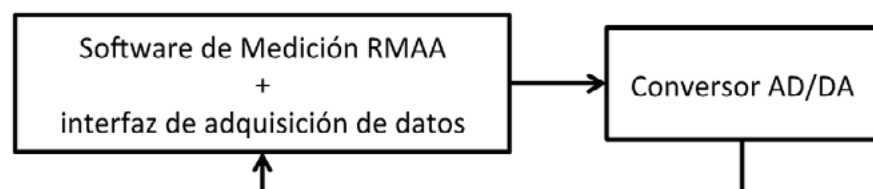


Figura 5. Flujo de señal para análisis objetivo

1) Respuesta En Frecuencia

Al revisar los datos obtenidos y presentados en la tabla 4 se observa que la respuesta en frecuencia de todas las configuraciones es muy similar y solo se aprecia una variación de

0.02 dB cuando la frecuencia de muestreo es 176.4 kHz. Los resultados de respuesta en frecuencia se presentan como un rango, sin embargo en la Figura 6 se aprecia que la respuesta en frecuencia solo

TABLA 4
Respuesta en frecuencia de las configuraciones.

Frecuencia de muestreo	Profundidad de bits	Respuesta en frecuencia	
		40 Hz – 15 kHz	20 Hz – 20 kHz
44.1 kHz	16	-0.57 dB, +1.17 dB	-8.47 dB, +1.17 dB
44.1 kHz	24	-0.57 dB, +1.17 dB	-8.47 dB, +1.17 dB
176.4 kHz	24	-0.59 dB, +1.17 dB	-8.47 dB, +1.17 dB

Aunque estos datos parecen no concordar con el teorema de Nyquist – Shannon, es importante recordar que las señales resultantes de las pruebas objetivas se les aplicaron procedimientos de dithering y re-muestreo. Adicionalmente se realizó una medición de las señales resultante previas a estos procesos cuyos resultados se revelan en la Figura 5, permitiendo afirmar que al usar frecuencias de muestreo superiores el sistema responde mejor en frecuencias mayores a 16 kHz.



Figura 6. Respuesta en frecuencia de cada configuración durante el proceso de grabación

2) Nivel De Ruido

Como se muestra en la tabla 5, la diferencia más grande de nivel de ruido se da entre las configuraciones con diferente profundidad de bits y equivale a un valor no mayor a las 0.05dB.

TABLA 5
Niveles de ruido de cada configuración

Frecuencia de muestreo	Profundidad de bits	Nivel de ruido
44.1 kHz	16	-93.8 dB
44.1 kHz	24	-93.5 dB
176.4 kHz	24	-94.2 dB

El oído humano es capaz de identificar cambios de hasta 1dB pero esto depende de muchos factores, entre ellos la frecuencia. Por esto se considera que esta diferencia de nivel de ruido es imperceptible para el ser humano. (Howard, 2009).

3) Rango Dinámico

La tabla 6 muestra los valores de rango dinámico obtenidos para cada una de las configuraciones evaluadas.

TABLA 6
Rango dinámico de cada configuración

Frecuencia de muestreo	Profundidad de bits	Rango dinámico
44.1 kHz	16	-93.9 dB
44.1 kHz	24	-94.3 dB
176.4 kHz	24	-94.4 dB

4) Distorsión Armónica Total

Los porcentajes de distorsión armónica total de cada configuración se presentan en la tabla 7. Se puede apreciar que las diferencias son pequeñas y se puede decir que el porcentaje de distorsión armónica de todas las configuraciones es inferior a 0.003.

TABLA 7
Distorsión armónica total de cada configuración

Frecuencia de muestreo	Profundidad de bits	% THD
44.1 kHz	16	0.0020
44.1 kHz	24	0.0028
176.4 kHz	24	0.0027

B. Análisis Objetivo de las Muestras de Audio

En primer lugar, se realizó un análisis objetivo sobre las muestras de audio que serían evaluadas y así determinar si existen diferencias entre las muestras que serán comparadas. Esta información será útil para determinar cuándo una respuesta es un acierto o desacierto y a partir de esto analizar los datos obtenidos. Las pruebas 1 y 2 buscan evaluar las diferencias en las etapas de una producción musical usando diferentes frecuencias de muestreo, esta variación hace que cambie la frecuencia máxima que puede ser muestreada. Considerando esto, se hizo un análisis espectral de ambas señales y así evidenciar las diferencias existentes entre las muestras que hacen parte de las pruebas 1 y 2, como se ve en las figuras 6 y 7.

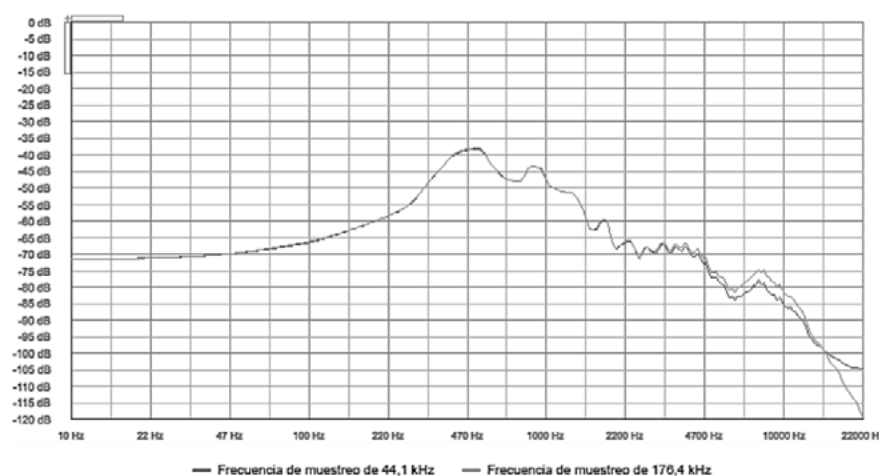


Figura 7. Análisis espectral de las muestras de audio usadas en la prueba 1

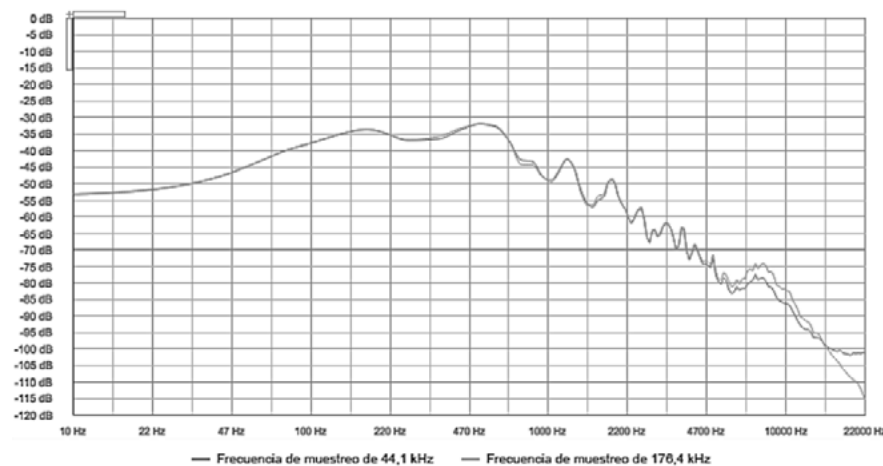


Figura 8. Análisis espectral de las muestras de audio usadas en la prueba 2

Ahora se presenta un análisis similar para las muestras que se presentaron en las pruebas 3 y 4, con la diferencia de que ya no se medirá ningún atributo relacionado con la frecuencia sino con el rango dinámico, ya que éste se ve afectado por la variación de la profundidad de bits. Aunque al graficar los niveles de amplitud pico y RMS durante 10 segundos como se muestra en las Figuras 28 y 29, se observan diferencias inferiores a 3 dB en los cambios de nivel RMS, estos cambios de nivel no están relacionados con el rango dinámico del producto final ya que ambas muestras de audio poseen el mismo rango dinámico (aproximadamente 96 dB) debido a que la profundidad de bits es igual a 16.

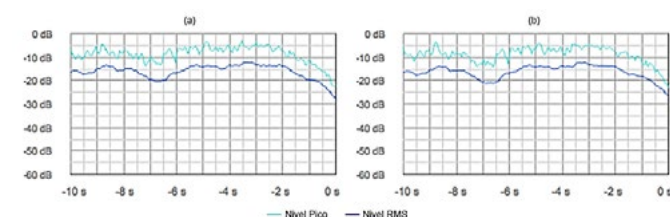


Figura 28. Nivel pico y RMS de las muestras usadas en la prueba 3

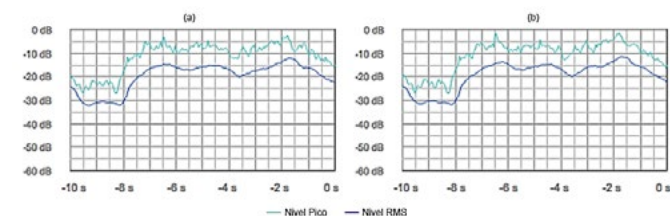


Figura 29. Nivel pico y RMS de las muestras usadas en la prueba 4

Estas pruebas se realizaron con el fin de obtener un criterio para indicar un caso de acierto o desacierto al tabular los datos, es decir que, si una persona expresó la percepción de alguna diferencia en las pruebas 1 y 2 se tomó como un acierto, de lo contrario es un desacierto. Caso contrario a las pruebas 3 y 4 donde la persona que indicó no haber apreciado alguna diferencia entre las muestras se tomó como un acierto y desacierto para la respuesta contraria.

C. Pruebas Subjetivas

A continuación, se presentan los datos obtenidos luego de procesar la información obtenida. La variable se comporta como una distribución binomial ya que solo tiene dos posibles valores, un acierto con probabilidad p y un desacierto con probabilidad $q=1-p$. Un caso concreto del teorema del límite central es la distribución binomial, que con número de muestras mayor o igual 30 se comparte como una distribución normal (Escofet, 2014). En la tabla 8 y la figura 8 con los resultados obtenidos para la prueba 1 que evaluaban las

muestras de audio donde se varió la frecuencia de muestreo. (Escofet, 2014)

TABLA 8
Personas que perciben alguna diferencia entre las muestras de la prueba 1

Número de prueba	Aciertos	Proporción	Desviación típica	Intervalo de confianza
1	28	80 %	6.76 %	66,7 % - 93,3%
2	27	77 %	7.10 %	63,2 % - 91,1%
3	18	51 %	8,45 %	34,87 % - 67,99%
4	13	37 %	8,17 %	21,13 % - 53,15%

RESULTADOS PRUEBAS SUBJETIVAS

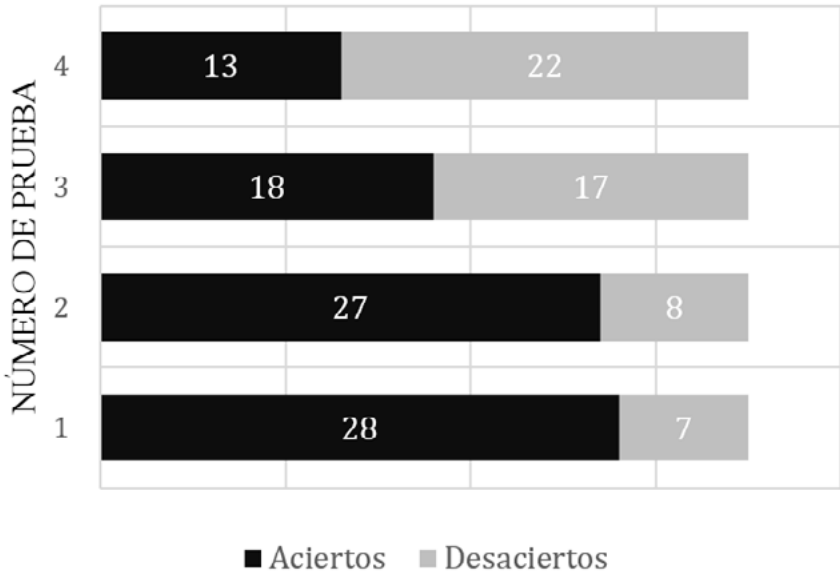


Figura 9. Cantidad de personas que perciben alguna diferencia entre las muestras de la prueba 1

Adicional a esto, se analizó si cada uno de los aciertos concordaba con las propiedades de las distintas muestras de audio de las pruebas 1 y 2 como se muestra en la TABLA 9.

TABLA 9
Personas que perciben los cambios correctos entre las muestras de audio

Número de prueba	Aciertos	Proporción	Desviación típica	Intervalo de confianza
1	17	61%	9,2 %	42,6% - 78,8%
2	27	77 %	7.10 %	63,2 % - 91,1%
3	18	51 %	8,45 %	34,87 % - 67,99%
4	13	37 %	8,17 %	21,13 % - 53,15%

IV. Conclusiones

Al caracterizar y comparar todas las configuraciones seleccionadas mediante pruebas objetivas no se evidencian grandes diferencias, asegurando que los conversores de audio no se ven beneficiados ni afectados por la variación de la frecuencia de muestreo o la profundidad de bits.

Algo diferente arrojó la realización de las pruebas subjetivas, donde el 80% de los individuos de la prueba 1 y 77% de los individuos de la prueba 2, percibieron diferencias al evaluar

las muestras de audio que fueron realizadas con diferente frecuencia de muestreo. Al realizar un análisis espectral sobre las muestras de audio de las pruebas 1 y 2 se observan diferencias en altas frecuencias. Después de aplicar un conjunto de pruebas, se concluye que estas diferencias son causadas por el procesamiento de señal mediante software tipo plug-in. Casó contrario para las pruebas 3 y 4 que evaluaban la diferencia en rango dinámico al usar dos configuraciones de profundidad de bits, donde los resultados obtenidos mediante análisis de niveles pico y RMS de ambas muestras, señalan que no existe diferencia, pero algunos individuos expresaron percibir un cambio entre las muestras. Esta percepción puede atribuirse a la dificultad, que fue comunicada por parte de los individuos encuestados de identificar cambios en el rango dinámico durante el desarrollo de las pruebas.

Finalmente se concluye el uso de diferentes frecuencias de muestreo si influye en el desarrollo de una producción musical, ya que ha sido posible evidenciar que se presentan diferencias al aplicar procesos de mezcla y masterización con software tipo plug-in, además de la cantidad de almacenamiento y recursos del sistema que son necesarios para trabajar con archivos de audio que tengan frecuencias de muestreo superiores a las establecidas en el “libro rojo”. Se deja abierto el análisis de otras configuraciones establecidas por otros sistemas de reproducción y una evaluación de diferentes configuraciones usando relojes que brinden mayor precisión, pues este es uno de los elementos más importantes en el proceso de conversión AD/DA.

V. Referencias

Aldrich, N. (2005). *Digital Audio Explained For The Audio Engineer*. Sweetwater Sound.

Escofet, C. R. (15 de Octubre de 2014). *Teorema del límite central*. Obtenido de http://www.calidad.com.mx/docs/art_64_1.pdf

Feeney, M. P. (2014). Acoustic Stapedius Reflex Measurements. En J. Katz, *Handbook of Clinical Audiology* (págs. 165-186). Wolters Kluwer Health.

Howard, D. M. (2009). *Acoustics and psychoacoustics*. Amsterdam: Focal.

Lavry, D. (2004). *Sampling Theory For Digital Audio*.

Lavry, D. (Mayo de 2012). *The Optimal Sample Rate for Quality Audio*. Pras, A., & Guastavino, C. (2010). Sampling rate discrimination: 44.1 kHz vs. 88.2 kHz . *AES 128th convention*. London.

Rumsey, F. (2007). High Resolution Audio. *Journal AES Volume 55 Issue 12*, 1161-1167.

UIT. (1994-1997). *Recomendación UIT-R BS.1116-1.Métodos Para La Evaluación Subjetiva De Pequeñas Degradaciones En Los Sistemas De Audio Incluyendo Los Sistemas De Sonido Multicanal*.

UIT. (1997 - 2003). *Recomendación UIT-R BS.1284-1. Métodos generales para la evaluación subjetiva de la calidad de sonido*.

Woszczyk, W. (2003). Physical and perceptual considerations for high-resolution audio. *AES 115th Convention*. New York.



AVANCES DEL AUDIO EN LATINOAMÉRICA.



Audio Engineering Society
Sección Colombia



MEZCLA BINAURAL MEDIANTE INTERFAZ TANGIBLE

Juan Esteban Herrera / Ricardo Andrés Moreno

Juan Esteban Herrera
Instituto Tecnológico Metropolitano
 Medellín, Colombia
 juanherrera37166@correo.itm.edu.co
 Ricardo Andrés Moreno
Instituto Tecnológico Metropolitano
 Medellín, Colombia
 ricardomoreno9013@correo.itm.edu.co

Resumen. Este artículo propone una alternativa a la forma convencional de realizar mezcla de señales de audio en procesos de postproducción musical; para lograrlo se implementa un panoramizador binaural controlado a través de una interfaz tangible de usuario (TUI), conformando un prototipo de mesa de mezcla binaural que permite la manipulación de parámetros de audio de forma práctica e intuitiva, reemplazando los tradicionales controles de tipo atenuador y perilla que han sido empleados durante décadas en las interfaces físicas de consolas de mezcla y en interfaces gráficas de estaciones de audio digital (DAW). El propósito de la interfaz TUI es permitir la interacción del usuario con objetos físicos, que son dispuestos sobre una superficie o mesa; dichos objetos constituyen representaciones tangibles de fuente(s) y receptor de audio que se usan para configurar imágenes sonoras tridimensionales que son recreadas en la producción de señales de audio binaural.

Palabras clave. Mezcla binaural, Audio 3D, Síntesis binaural, interfaz tangible.

Abstract. The present paper proposes an alternative to the conventional way of mixing audio signals in musical postproduction processes, to achieve this, a binaural panner controlled through a tangible user interface (TUI) was deployed, thus building a binaural mixing table prototype which allows practical and intuitive manipulation of audio parameters, by replacing traditional controls such as faders and knobs, which have been used for decades on the physical interfaces of mixing consoles and on the graphical interfaces of digital audio workstations (DAW). The purpose of the TUI interface is to allow user interaction with physical objects arranged on a surface or table; these objects are tangible representations of audio source(s) and receiver, used to set up 3D sound images, recreated in the production of binaural audio signals.

Keywords. Binaural audio mixing, 3D audio, Binaural synthesis, Tangible user interface.

I. Introducción

El principal propósito de usar sistemas de captura y procesamiento binaural de señales de audio, es proporcionar al oyente contenido audible con características tridimensionales, que pueden ser perceptibles gracias a la forma en que escucha el ser humano y a la interpretación que realiza el cerebro de las señales que llegan a los oídos izquierdo y derecho, este proceso se conoce como fusión binaural, y permite inferir el ángulo de precedencia de las fuentes sonoras (Xaus, 2009). El ser humano en su cotidianidad está constantemente expuesto a campos sonoros tridimensionales, por ejemplo, sonidos de la naturaleza o sonidos generados por la interacción del hombre con objetos; es por tanto coherente, en aras de realismo, que un sistema de grabación y reproducción de sonido capture y reproduzca fielmente las

condiciones del campo sonoro para poder proporcionar percepción de espacialidad, direccionalidad y distancia de las fuentes emisoras a su alrededor. Referentes musicales de los años ochenta, contemplaron la espacialidad sonora usando técnicas de grabación que permitieron establecer condiciones de espacialidad, la agrupación Pink Floyd fue una de las pioneras en experimentar con grabación binaural en su álbum “*The final cut*”, tiempo después esta actividad se popularizó con otras agrupaciones musicales entre las que se destacan producciones de artistas como: Pearl Jam “*Binaural*”, Michael Jackson “*History*”, entre otros (Fontana, Farina, & Yves, 2007); estas producciones, principalmente, utilizaban técnicas de grabación que requerían el uso de maniquíes o cabezas artificiales capaces de emular la condición de escucha humana, gracias a la implementación de micrófonos ubicados en la parte interior de sus orejas, sin embargo, el efecto binaural que genera la cabeza artificial, radica principalmente en los fenómenos acústicos causados por el choque de la onda sonora en la irregularidad de la cabeza, el pabellón y conducto auditivo, más que por micrófonos usados en la captura. Entregarle al oyente sonido 3D exige un replanteamiento total de la cadena de producción, desde la grabación de sonido y procesamiento, hasta la adaptación a formatos de audio para su posterior reproducción (Cengarle, 2012). Este artículo se enfoca en la mezcla y el procesamiento de audio en postproducción, lo cual es una labor técnica y artística que está en constante evolución, permitiendo hoy en día el uso de procesamiento digital para producir señales binaurales; estos procesos hacen uso de la síntesis binaural, sin embargo, aunque aseguran resultados de espacialidad óptimos, solo es posible percibirlos a través de sistemas de reproducción monousuario, por ejemplo a través de audífonos o de un sistema compatible de altavoces con filtros canceladores de crosstalk (Cengarle, 2012); además de un monitoreo especial, el proceso de mezcla binaural requiere herramientas de control adecuadas que permitan ajustar el nivel y ubicación de la fuente en una panorámica tridimensional. En una consola tradicional, la mezcla implica en términos eléctricos, el ajuste de controles de resistencia variable (faders o perillas), que son componentes estándar de las interfaces físicas de la mayoría de equipos y procesadores de audio, incluso en las interfaces gráficas de usuario (GUI) de sistemas DAW y de procesadores de audio virtuales, se hayan dichos controles con esquemas de operación similares a las versiones hardware. Actualmente han surgido herramientas digitales que definen otra manera de hacer las tareas de mezcla, brindando posibilidades creativas y novedosas, tal es el caso de las interfaces multitouch que hoy permiten generar nuevas sensaciones y cambiar un poco la manera tradicional de hacer mezcla de señales de audio (Carralscal & Jorda), sin embargo, para algunos usuarios, como para los autores de este artículo, las interfaces multitouch no han logrado un buen nivel de aceptación, ya que en su gran mayoría y pese a la popularidad que han alcanzado en dispositivos como pantallas, tablets y teléfonos móviles, carecen de una respuesta háptica, lo cual, las hace poco atractivas para el trabajo profesional. Por otro lado, han aparecido interfaces tangibles orientadas a realizar labores

musicales y de producción de audio, las cuales permiten al usuario interactuar directamente con datos digitales por medio de objetos físicos, éstas han resultado bien aceptadas; tal es el caso del proyecto Reactable*, un sintetizador que hace uso de una TUI para controlar generadores de sonido usando técnicas de síntesis; dado lo innovador del uso de este tipo de interfaces, se han generado nuevas formas de ver y propiciar la interacción hombre – computadora, cada vez tendiendo a desarrollar interfaces más antropomorfas, donde la máquina se acopla al pensamiento y el accionar del humano, con aplicación al campo de la música y la ingeniería de sonido (Jorda, 2005).

En este artículo se describe el diseño e implementación de un mezclador binaural, a través de una interfaz tangible de usuario, que permite realizar procesos de mezcla de audio de una manera simple e intuitiva, mediante representaciones tangibles de fuente(s) y receptor de audio para configurar imágenes sonoras tridimensionales propias de procesos de mezcla binaural.

II. Parte técnica

El desarrollo del mezclador binaural se dividió en dos etapas, A y B; la primera consistió en la elaboración de la interfaz tangible y la segunda en la elaboración de un aplicativo que se controla a través de la interfaz para la realización de mezcla binaural.

A. Diseño de la interfaz TUI

Para la elaboración de la interfaz tangible se implementó el toolkit Reactivision² y el protocolo de comunicación TUIO³, la materialización de la interfaz se tradujo en la construcción de una mesa en madera, que incorpora una cámara infrarroja para la detección de objetos y una superficie semitransparente con dimensiones (55cm x 70 cm), la proporción largo x ancho se establece en relación a la distancia focal de la cámara ubicada al interior de la mesa; se utilizó una cámara Sony PS3eye modificada en su óptica, mediante la adición de un filtro que bloquea la luz visible y un lente de foco ajustable; la cámara infrarroja se usa para leer información de posición que poseen los objetos físicos ubicados sobre la mesa (Bencina & Kaltenbrunner, 2005). Estos objetos corresponden a una serie de fichas que poseen un código visual o fiducias**, todas las fiducias son leídas de forma simultánea por la cámara infrarroja al interior de la mesa, proporcionando información de control codificada según las especificaciones del protocolo TUIO, dicha información consiste en tres elementos: un número de identificación ID, dato de posición en la superficie de la mesa expresado en coordenadas X, Y y ángulo de giro u orientación. Todos los datos son actualizados al interactuar y mover los objetos. La Fig. 1 muestra un ejemplo donde se analizan los datos correspondientes a dos objetos sonoros respecto a la posición del objeto receptor.

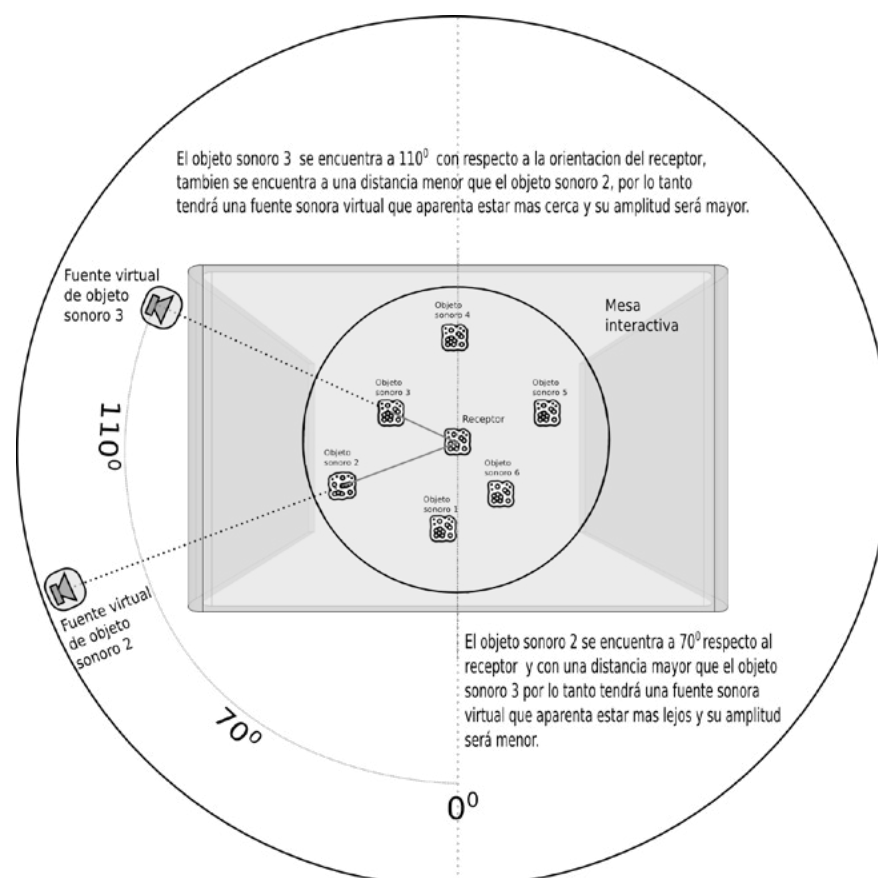


Fig 1. Ejemplo de interacción con objetos sonoros.

La mesa cumple el papel de mezclador binaural donde la distancia y el ángulo entre cada objeto instrumento y receptor, se asigna a variables que se utilizan en el aplicativo para determinar el nivel y la posición de la fuente en la panorámica binaural.

Descripción de los Objetos

La interfaz posee objetos tangibles para el control de datos mediante el empleo de fiducias, en este caso se lograron tres tipos diferentes de objetos: objeto receptor, objeto instrumento y objeto efecto.

Objeto Receptor

El objeto receptor es la representación del oyente en la mesa, este objeto es uno de los elementos principales de la interfaz, ya que con el movimiento y rotación del objeto receptor, son actualizados todos los datos de distancia y orientación de los demás objetos para mantener coherencia en la mezcla final, esto permite generar recorridos audibles entre los instrumentos que componen la mezcla. En condiciones de operación estándar se sugiere dejar fijo el objeto receptor en el centro de la mesa, en posición coherente con la orientación del oyente; la ficha del objeto receptor tiene un dibujo de cabeza que indica la orientación en la mesa.

Objeto Instrumento

La interfaz creada soporta hasta 8 objetos de instrumento, éstos son la representación de un módulo de canal en una mesa de mezcla; haciendo uso de todos los datos que entrega cada fiducia, se propone un esquema de control multiparamétrico (Hunt & Wanderley, 2000), para nivel, ángulo azimut y ángulo de elevación; donde, el nivel depende de la cercanía con el objeto receptor, el ángulo azimut corresponde al ángulo que se forma entre el eje que indica la orientación del receptor y el vector que forma el objeto instrumento,

* <http://mtg.upf.edu/project/reactable>

2 <http://reactivision.sourceforge.net/>

3 <http://www.tuio.org/>

4 Fiducia es un símbolo de referencia o código visual

tomando como origen la posición del receptor. Por último, el ángulo de elevación se establece mapeando el ángulo de rotación del objeto instrumento; donde al girar hacia la derecha aumenta su elevación hasta 45 ° y al girar hacia la izquierda disminuye su elevación hasta -45 ° respecto al receptor.

Objeto Efecto

Este objeto se usa para el control de efectos sonoros de tiempo y modulación; cada objeto efecto actúa como un canal auxiliar que recibe la señal de los objetos tipo instrumento, donde el nivel de envío depende de la distancia entre objeto instrumento y objeto efecto; por otro lado, el ángulo de rotación se usa para controlar parámetros según el efecto que representa. En la tabla 1 se describe la asignación sugerida e implementada para diferentes tipos de efectos:

TABLA 1
MAPEO DE PARÁMETROS EN EFECTOS

Tipo de efecto	Control de ángulo de rotación
Reverb	Decay
Delay	Subdivisión de tiempo
Flanger, phaser	Feedback

B. Programación

Se utilizó la plataforma de programación Pure Data (PD)^{***}, ésta permite crear entornos de interacción para aplicaciones multimedia a través de objetos y extensiones; se programó un patch^{****} que sirve de puente de comunicación entre la interfaz y la plataforma DAW que contiene los canales de audio a mezclar. Las señales de audio se envían a PD desde el DAW usando la herramienta Jack^{*****}, ésta actúa como un servidor de audio de baja latencia que permite compartir audio entre aplicaciones; el cliente, en este caso, el DAW de preferencia del usuario, puede ejecutar procesos de audio independientes y compartir las señales con el servidor a través de puertos de audio virtuales. Por otro lado, PD recibe los datos de la interfaz tangible para calcular las distancias y ángulos entre los objetos de instrumento, efecto y receptor. En la figura 2 se observa un diagrama de bloques general del sistema completo.

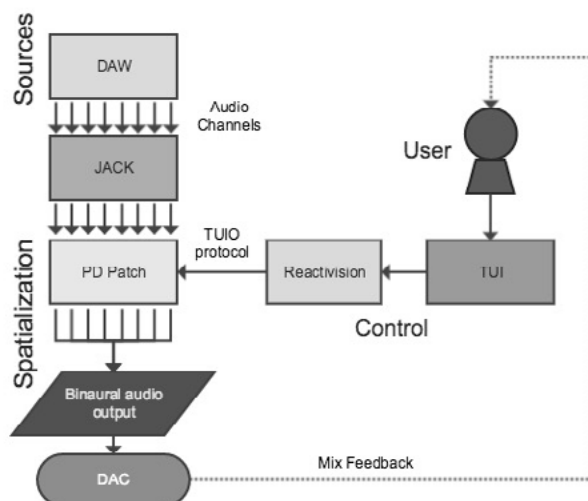


Fig. 2. Diagrama de bloques general

*** <http://puredata.info/>

**** Se denomina patch al ejecutable programado en Pure Data.

***** <http://jackaudio.org/>

Síntesis binaural

Para lograr el procesamiento de audio y convertir las señales monofónicas a señales binaurales, se implementó el external para PD, CW_Binaural~ (Doukhan & Sedes, 2009), este objeto realiza un proceso de filtrado digital en tiempo real, usando respuestas al impulso relativas a la cabeza (HRIR); para este proyecto se utilizaron las bases de datos del Ircam, estas bases de datos corresponden a mediciones de HRIR realizadas a varias personas con diferentes características físicas en una cámara anecoica (Ircam, 2010); también se experimentó con una base de datos propia construida a partir de mediciones de una cabeza artificial 01dB Mk2, tomando como referencia el protocolo de medición propuesto por el Ircam. Dado que las bases de datos no contienen las mediciones para todas las posiciones espaciales posibles y estas posiciones además varían con el movimiento de cada fuente a espacializar, se requirió usar una instancia del objeto CW_Binaural~ por cada canal del aplicativo; dicho objeto actúa como un interpolador de las funciones de transferencia relativas a la cabeza (HRTF), con el fin de aproximar y obtener en todo momento los filtros necesarios que caracterizan la posición deseada para cada una de las fuentes a procesar.

Como se describió en la etapa A, el ángulo azimut de cada fuente representada por cada uno de los objetos tipo instrumento, se obtiene a partir de las coordenadas de posición de los objetos instrumento y receptor, y se calcula según la expresión mostrada en (1).

$$\theta = \text{atan2}((or.y - oi.y)/(or.x - oi.x)) \quad (1)$$

Donde, es el ángulo azimut, or.x y or.y son las coordenadas del objeto receptor y oi.x y oi.y son las coordenadas del objeto instrumento. La función atan2 es una modificación de atan implementada en la mayoría de lenguajes de programación, incluido PD, se utiliza para entregar el ángulo de un vector entre 0 y 2π respecto al eje de abscisas; luego, a este ángulo, se le suma el ángulo correspondiente a la rotación o eje de orientación del receptor para obtener correctamente la ubicación o imagen virtual del instrumento. El ángulo de elevación se encuentra limitado entre -45° a +45° y se controla con el giro de cada objeto instrumento; al situar un nuevo objeto instrumento sobre la mesa, el aplicativo asigna por defecto un ángulo de elevación 0°, esto quiere decir que el instrumento se encuentra a la altura de los oídos, a partir de este punto la rotación del objeto se programa para funcionar como una perilla infinita, donde el incremento o decremento en el ángulo de elevación es controlado por la dirección de rotación de cada objeto instrumento hacia derecha o izquierda respectivamente, con la posibilidad de ser visualizado en el aplicativo PD.

El valor de amplitud del instrumento corresponde a la magnitud del vector, cuyo origen es la posición del receptor, y el destino, la posición del instrumento; dicha magnitud se calcula como se muestra en (2)

$$mag = \sqrt{((or.y - oi.y)^2 + (or.x - oi.x)^2)} \quad (2)$$

El valor de magnitud, al igual que el ángulo, se calcula con valores expresados en número de píxeles dados por la imagen

que capta la cámara, este valor luego se escala para operar en un rango entre 0 a 1, que corresponde al nivel del canal L y R de la señal binaural obtenida para los ángulos azimut y elevación dados. En la Fig. 3, se muestra el patch principal del mezclador binaural, con una interfaz gráfica que sirve para monitorear los niveles de entrada de las señales a procesar, los niveles de salida de la señal binaural, los ángulos de azimut y elevación, y la posición de cada instrumento en la superficie de la mesa.

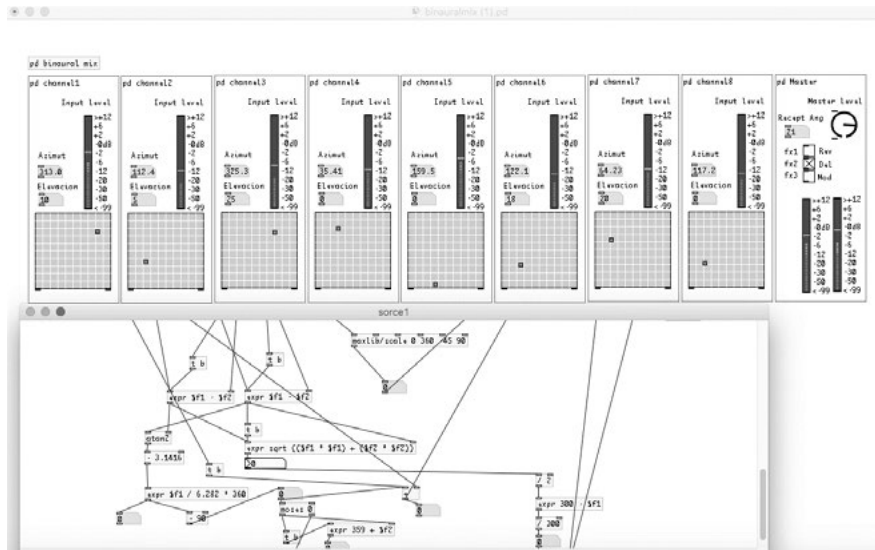


Fig. 3. Patch principal del mezclador binaural creado en PD

Reproducción binaural

La reproducción de audio binaural de esta interfaz se hace posible a través del uso de audífonos, este método de reproducción de audio ofrece una buena impresión espacial (Barlett & Barlett, 1999), garantizando que las señales procesadas en la síntesis binaural con las HRIR del oído izquierdo y derecho lleguen en correspondencia a los oídos del receptor, evitando problemas de diafonía (Cengarle, 2012). Sin embargo, la precisión en la identificación del ángulo de procedencia de una fuente cuando se escucha una señal binaural, depende en gran medida de la antropometría de la cabeza artificial usada, obteniendo mejores resultados cuando ésta se aproxima a la antropometría del usuario. En el caso del patch programado se brinda la opción de cambiar la base de datos de HRIR cargada en `cw_binaural~` buscando aproximar las respuestas a las características del usuario mediante un test de escucha previo en pro de ofrecer una experiencia audible más realista.

III. Resultados

Se obtuvo un prototipo de mesa que incorpora una interfaz tangible y un aplicativo para ser usado en procesos de mezcla binaural de señales de audio. La interacción con la interfaz permite transmitir datos de posición para realizar síntesis binaural de ocho canales de instrumento simultáneos, más un canal master con 3 envíos auxiliares para efectos.

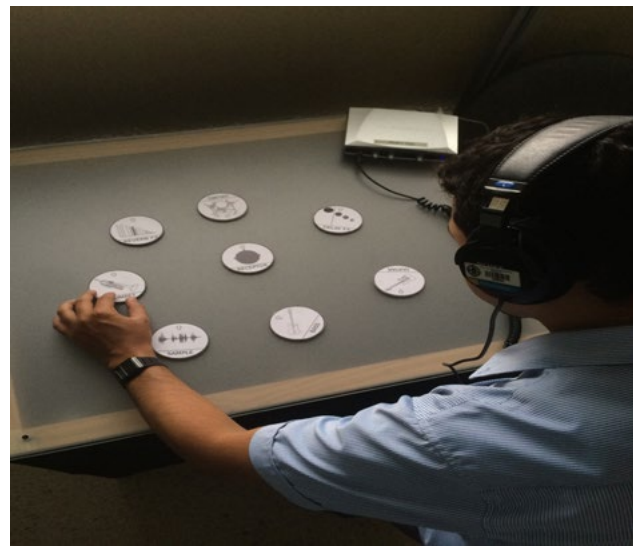


Fig. 4. Prototipo de mesa para mezcla binaural

Se obtuvo un prototipo de bajo costo de producción gracias a que en todas las etapas de diseño y construcción se utilizó software, librerías, protocolos y extensiones de libre distribución. Luego de un proceso de experimentación con la interfaz y el aplicativo de síntesis binaural, en sistemas de reproducción con audífonos; y en el sistema de altavoces con tecnología Opsodis^{*****}; se obtienen imágenes sonoras tridimensionales que recrean la escena física configurada sobre la mesa. El aplicativo incorpora un factor de atenuación por distancia entre objetos que puede ser modificado; sin embargo su comportamiento es lineal; como trabajo futuro se plantea la implementación de curvas de atenuación modificables y programación de un algoritmo que simule efecto doppler.

IV. Conclusiones

Este proyecto contribuyó a proporcionar una herramienta no convencional para facilitar los procesos de mezcla binaural como respuesta al notable incremento de utilidades o herramientas software en formato plugin o standalone para procesamiento binaural y a la necesidad de usar interfaces de control que se ajusten a los procesos de una panorámica tridimensional.

Las pruebas de interacción con el prototipo mostraron que el tiempo de actualización de las posiciones de objetos, está sujeto a las características técnicas de la cámara PS3eye, alcanzando en el mejor de los casos una actualización de 45 fps; esto se traduce en la escucha de saltos en posición de los instrumentos cuando se realizan movimientos o giros bruscos sobre la superficie de la mesa, esta situación es de especial cuidado cuando se desea grabar el audio correspondiente a un movimiento del objeto receptor; para mejorar esta condición se hace necesario implementar a futuro una cámara de mejores especificaciones técnicas.

Luego de testear el prototipo con un sistema de monitoreo a través de audífonos, se evidencian problemas comunes en la mezcla como el incremento en frecuencias bajas por efecto de oclusión y lateralización de fuentes en las señales binaurales, por lo tanto se utilizó un sistema de monitoreo alterno basado en altavoces, para este caso, el reproductor Opsodis, con éste se encontró que aunque se evitan los problemas anteriormente mencionados con audífonos, se requiere de

***** <http://resource.isvr.soton.ac.uk/FDAG/VAP/pdf/OPSODIS.pdf>

unas condiciones de escucha controladas para garantizar en un sweetspot pequeño una percepción binaural.

En este artículo se presentó una descripción de cómo construir un prototipo de mesa de mezcla binaural de bajo costo; que puede reproducirse o considerarse como referente para diseñar y construir un prototipo propio ajustado a las necesidades del usuario.

V. Referencias

Barlett, B., & Barlett, J. (1999). On-location recording techniques. Focal Press.

Bencina, R., & Kaltenbrunner, M. (2005). The Design and Evolution of Fiducials for the reacTIVision System. In Proc. 3rd International Conference on Generative Systems in the Electronic Arts. Melbourne.

Carralscal, J., & Jorda, S. (s.f.). Multitouch interface for audio mixing.

Cengarle, G. (2012). 3D audio technologies: applications to sound capture, post-production and listener perception.

Doukhan, D., & Sedes, A. (2009). CW_binaural~: A binaural synthesis external for Pure Data. In Proc. 3rd Puredata Int. Conv. (Pd-Con09). Sao Paulo.

Fontana, S., Farina, A., & Yves, G. (2007). Binaural for popular music: A case of study. Proceedings of the 13th International Conference on Auditory Display, (pág. 6). Montréal.

Hunt, A., & Wanderley, M. (2000). Towards a model for instrumental mapping in expert musical interaction. In Proceedings of the 2000 International Computer Music Conference, (págs. pp. 209-212).

Ircam, R. A. (2010). <http://recherche.ircam.fr>. Obtenido de http://recherche.ircam.fr/equipements/salles/listen/system_protocol.html

Izhaki, R. (2008). Mixing audio: concepts, practices and tools. Focal Press.

Jorda, S. (2005). Digital Lutherie: Crafting musical computers for new musics performance and improvisation. Tesis, Universitat Pompeu Fabra, Tecnologia, Barcelona.

Kaltenbrunner, M., & Bencina, R. (2007). reacTIVision: a computer-vision framework for table-based tangible interaction. In Proceedings of the 1st international conference on Tangible and embedded interaction (págs. pp. 69-74). ACM press.

Klemmer, S. (2004). Tangible User Interface Input: Tools and Techniques. University of California, Berkeley.

Moggridge, B. (2007). Designing Interactions . London: MIT Press.

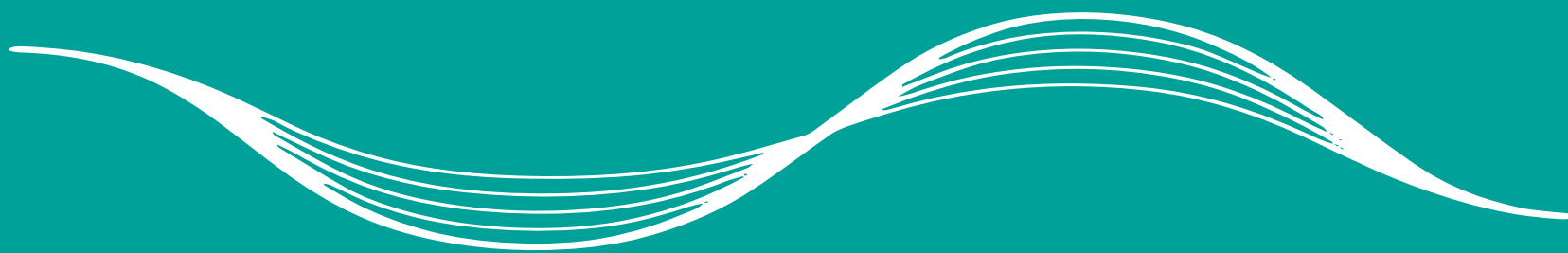
Moylan, W. (2002). The art of recording: understanding and crafting the mix.

Xaus, A. (2009). La percepción acústica egocéntrica y principios de sistemas espaciales. Barcelona.

AVANCES DEL AUDIO EN LATINOAMÉRICA.



Audio Engineering Society
Sección Colombia



DESARROLLO DE ALGORITMO DE INTERPRETACIÓN DE RESPUESTAS AL IMPULSO EN TRES DIMENSIONES

Federico Nahuel Cacavelos / Augusto Bonelli Toro / Alejandro Bidondo

Federico Nahuel Cacavelos*, Augusto Bonelli Toro**,
Alejandro Bidondo***

Universidad Nacional de Tres de Febrero,
Buenos Aires, Argentina

fnahuelc@gmail.com
abonellitoro@gmail.com
abidondo@untref.edu.ar

Resumen—En el presente trabajo se presenta el desarrollo de un algoritmo de interpretación de respuestas al impulso en tres dimensiones. Esta herramienta posibilita la visualización de vectores de intensidad en un sentido espacial, siendo útil para distintas ramas de la acústica arquitectónica, la producción de audio y la industria. Esta visualización permite distinguir de forma espacial y temporal, el sonido directo de las reflexiones en una respuesta al impulso (RI) grabada con un arreglo de micrófonos. Diversas técnicas de procesamiento digital de señales fueron abordadas para la evaluación de los datos. Finalmente, se constató el funcionamiento del algoritmo con una serie de mediciones.

Palabras clave— Respuesta al impulso en 3 dimensiones, Erizo, Espacialización, Soundfield, Ambisonics.

Abstract— In this paper, the development of an interpretation algorithm for 3D impulse responses is presented. This new tool allows to identify the spatial provenance for each time window of an impulse response recorded in aA-Formatmicrophone. In this way, it is possible to discern between the direct sound and the different reflections temporarily and spatially. To understand the basics of this method, a brief introduction to all concepts involved is presented. Alternatively, different types of procedures are used for data evaluation. Finally, system applications are shown in controlled acoustical environments to evaluate the precision and the potential of this tool.

Keywords— Impulse response 3D, Hedgehog, Spacialization, Soundfield, Ambisonics.

I. Introducción

La captación de respuestas al impulso ha sido ampliamente estudiada con anterioridad en diversas áreas de la acústica. El uso más conocido está dado en la industria de la producción de audio comercial. La convolución de este tipo de respuestas al impulso con señales anecoicas posibilitan recrear ambientes sonoros envolventes. En el presente trabajo se presenta un uso alternativo a este tipo de respuestas, con aplicaciones en la caracterización de los parámetros acústicos y diseño de recintos.

Este trabajo presenta la metodología y consideraciones que se realizaron en el desarrollo de un algoritmo de interpre-

tación de respuestas al impulso en tres dimensiones. Para esto se utilizó la plataforma MATLAB como entorno de programación. Esta plataforma permite la visualización de vectores de intensidad en tres dimensiones. De esta manera, se puede distinguir de forma espacial y temporal, el sonido directo de las reflexiones en una respuesta al impulso grabada con un arreglo de micrófonos. Para esto, diversas técnicas de procesamiento digital de señales fueron abordadas para la evaluación de los datos. Finalmente, se constató el funcionamiento del algoritmo con una serie de mediciones.

Aplicaciones de la respuesta al impulso en tres dimensiones

Los espacios acústicos son comúnmente evaluados utilizando técnicas de medición de respuestas al impulso omnidireccionales. Estas técnicas permiten obtener caracterizaciones en términos de nivel y tiempo, pero no proporcionan información sobre la dirección. En este sentido la ventaja de la comprensión y caracterización de los parámetros acústicos a través de la medición de respuestas al impulso en tres dimensiones es la de abarcar también una caracterización en términos de dirección. La obtención de la dirección permite entre otras, relacionar las reflexiones sonoras a las características físicas de la sala, observar la distribución direccional de la energía temprana y tardía, así como identificar las superficies que generan reflexiones indeseadas. Por otro lado, la comparación de las direcciones de los componentes de respuestas al impulso de distintas salas puede aportar a la identificación de diferencias y similitudes de patrones acústicos.

Si bien existen en la actualidad parámetros acústicos que describan características espaciales de la respuesta al impulso, es posible desarrollar nuevos parámetros al contemplar la dirección como una variable independiente.

Formato A y Formato B

El formato A de micrófonos permite la captación de cuatro señales de audio. Para esto es necesario utilizar un arreglo de micrófonos con cuatro cápsulas sub-cardioides ubicadas en las caras de un tetraedro.

Desde un aspecto teórico las cápsulas deberían coincidir en un solo punto, pero debido a limitaciones físicas obvias, esto resulta imposible. Esta condición produce una distorsión de fase en longitudes de onda menores a cuatro veces la distancia entre las cápsulas. Por este motivo es necesario implementar filtros que corrijan este problema y que representen la especialización sonora como si las cápsulas estuvieran virtualmente colocadas en el centro del tetraedro.

El formato B consiste en cuatro señales denominadas W, X, Y y Z. La señal W representa el componente de presión de la señal captada en forma omnidireccional, mientras que X, Y y Z representan los vectores de velocidad relacionados con las direcciones horizontales (X e Y) y la dirección vertical (Z). El micrófono formato B se obtiene utilizando tres micrófonos figura de 8 y un micrófono omnidireccional. Otra manera de obtener dicho formato, es realizando una combinación entre las señales en formato A y la posterior aplicación de un filtro de compensación de fase. [1, 2].

*N. Cacavelos estudiante en Ingeniería de Sonido en Universidad Nacional de Tres de Febrero, Valentín Gómez 4828/38, Buenos Aires, Argentina y Universidad de San Buenaventura, Carrera 56C N° 51-110, Medellín, Colombia.(fnahuelc@gmail.com.ar).

** A. Bonelli estudiante en Ingeniería de Sonido en la Universidad Nacional de Tres de Febrero, Valentín Gómez 4828/38, Buenos Aires, Argentina (abonellitoro@gmail.com).

*** A. Bidondo coordinador de carrera de Ingeniería de Sonido en la Universidad Nacional de Tres de Febrero, Valentín Gómez 4828/38, Buenos Aires, Argentina (abidondo@untref.edu.ar).

Vectores intensidad a partir de señales Formato B

Como afirman Farina y Tronchin en su investigación de la Universidad de Parma (2013), los vectores intensidad en una dirección (correspondientes a una o más fuentes) se pueden obtener mediante el producto de la presión (es decir, la señal W del micrófono) y el vector velocidad deseado (X , Y o Z). Posteriormente, con los valores de la intensidad en cada uno de sus ejes, se puede calcular el módulo de la misma. Para obtener los ángulos de desplazamiento (azimuth y elevación) de cada uno de los puntos del vector intensidad se pueden utilizar las ecuaciones trigonométricas expuestas en el trabajo anteriormente mencionado [3].

Log Sine Sweep y algoritmo de Kirkeby

La técnica de medición de respuestas al impulso con Log Sine Sweep (LSS) consiste utilizar como señal de excitación de un recinto a una señal del tipo seno, la cual varía exponencialmente entre dos frecuencias. La ventaja radical que adquiere esta técnica respecto a un barrido lineal es que al tratarse de un espaciado logarítmico, presentará una mayor duración para los tonos de baja frecuencia. De esta manera se pueden lograr mayores relaciones señal ruido aun en baja frecuencia. En contraposición a un barrido lineal, este último mantiene la misma duración para todo el rango de frecuencias [4, 5], siendo así necesario aplicar mayores tiempos de duración de barrido.

La aplicación del filtro inverso de la fuente también produce una mejora en la calidad de la respuesta al impulso medida. El uso de esta herramienta permite reducir las posibles alteraciones que pueda producir la alinealidad de la fuente utilizada. Para esto, es preciso contar con la respuesta al impulso de la fuente utilizada medida en condiciones anecoicas. A fines prácticos, es posible lograr una buena evaluación de la misma aun en condiciones acústicas no ideales. En la siguiente sección se desarrollarán algunas de las herramientas plausibles de utilizar. De forma complementaria se utiliza el algoritmo de Kirkeby para la aplicación del filtro inverso a las señales medidas. La aplicación de este filtro mejora en forma el contenido en frecuencia, eliminando la dispersión temporal producida por la alinealidad de la fuente.[5]

II. Procedimiento

Adquisición de respuestas al impulso

Para la obtención de las respuestas al impulso se utilizó la técnica desarrollada por Farina con el LSS en conjunto con el algoritmo de filtro inverso de Kirkeby, el cual reduce las alteraciones en la señal producidas por las características de la fuente

Dada la dificultad para obtener la respuesta de la fuente omnidireccional en condiciones anecoicas se implementaron una serie de métodos alternativos capaces de obtener resultados aceptables. La colocación del micrófono omnidireccional fue hecha lejos de una superficie reflectante de manera de obtener el sonido directo de la fuente lo suficientemente espaciado de las reflexiones tempranas. El criterio para elegir dicha distancia fue que sobrepase al menos una longitud de onda de la frecuencia más baja de análisis. En los casos

donde no se conseguía la distancia suficiente, otra técnica utilizada estuvo dada por colocar el micrófono lindante a la superficie reflejante. Este método modifica el patrón de captación del micrófono a semiesférico e imposibilita la adquisición de reflexiones producidas por la superficie.

Es preciso aclarar que la aplicación del filtro inverso de la respuesta de la fuente no realiza modificaciones respecto a las diferencias de respuesta entre las cápsulas. Es por esto, que futuras mediciones deben ser tomadas en cuenta a fin de realizar una caracterización de las capsulas utilizadas.

Mecanismo de corrección de ganancias

Uno de los aspectos importantes a tener en cuenta refiere a la configuración de ganancias de la interface. La inequidad de relaciones de ganancia entre los distintos canales puede provocar la obtención de resultados erróneos.

Con este fin se desarrolló un algoritmo de corrección de ganancias. En el mismo se puede iniciar la medición con cualquier criterio de ajuste de ganancias, preferentemente el que logre la mayor relación señal ruido sin distorsión. Luego, se introduce una señal de ruido blanco con la misma duración (puede ser generada con la misma interface) en cada uno de los canales utilizados. Las diferencias relativas entre las sumas energéticas de cada una de las señales adquiridas representarán las relaciones de ganancia entre canales. De esta manera se debe aplicar correcciones de amplitud a fin de obtener la misma suma energética en todos los canales.

Limitaciones prácticas y filtros

Como se mencionó anteriormente, hay una restricción en frecuencia debido a la distancia entre las cápsulas. Para compensar la distorsión de fase en alta frecuencia es necesario emplear un filtro pasa – bajos con una frecuencia de corte acorde al micrófono utilizado. En este caso, se utilizó un filtro FIR de fase lineal, de orden 400 y frecuencia de corte en 5 kHz, como se puede apreciar en la figura 1.

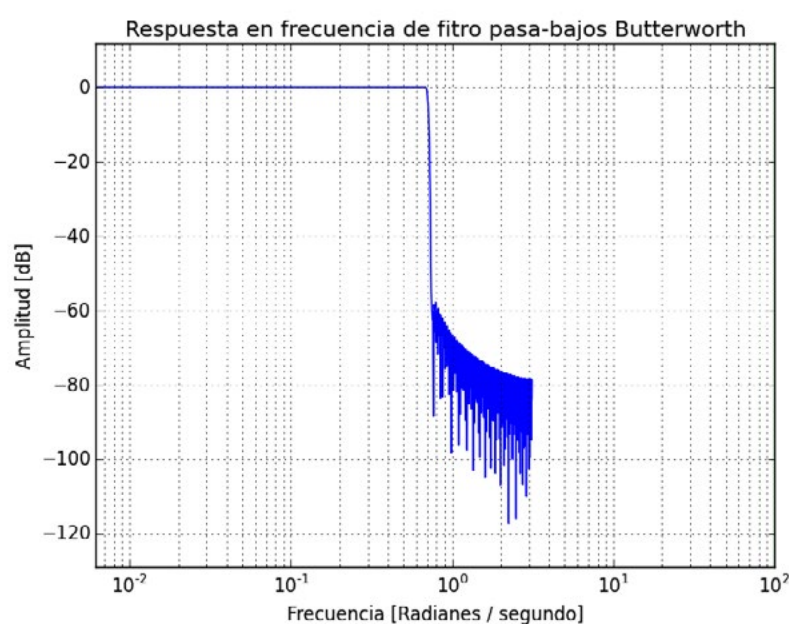


Fig. 1. Filtro FIR de fase lineal para evitar distorsión de fase

La necesidad de implementar un filtro de fase lineal radica en la posible alteración de la interpretación de resultados al introducir un corrimiento de fase que no es propio de la señal medida.

Presentación de gráficas

Las gráficas se representan en forma de vectores con un mismo origen y distintas direcciones. Cada uno de los vectores representa a cada una de las ventanas temporales. El largo del vector representa su magnitud, la cual está dada por la suma energética de los niveles de intensidad que contiene. Dichas magnitudes se presentan normalizadas en base al sonido directo. El color del vector simboliza la distancia temporal al sonido directo (de rojo al azul). Con el fin de facilitar el entendimiento de los resultados, los gráficos se presentan en los tres planos ortogonales que incluyen el punto de origen.

Tamaño temporal de los vectores graficados

El problema más discutido en el desarrollo del procedimiento de cálculo fue la elección de las ventanas temporales que representan cada uno de los vectores graficados. La más pequeña variación en el alcance de muestras contenidas en cada ventana influencia la dirección y magnitud de los vectores. Es por esto que distintos métodos de selección fueron desarrollados con el fin de comparar y contrastar los resultados encontrados.

La problemática radica en que el número de muestras que se adquieren de una ventana temporal impacta sobre la resolución frecuencial obtenida y por consiguiente de la menor frecuencia de análisis. Por consiguiente, una ventana temporal con menor número de muestras implica una limitación en baja frecuencia, desplazando a valores superiores la frecuencia inferior de análisis. Por otro lado, el tomar una ventana temporal extensa se debe considerar que la misma puede contener información proveniente de más de una reflexión, dando lugar a interferencias constructivas y destructivas, por lo que el análisis de los resultados debe ser tomado con criterio.

Las diferencias en la sensibilidad de las cápsulas a frecuencias bajas se consideran un factor crítico que puede introducir errores en el análisis de los datos.

En la figura 2 se muestran dos graficas obtenidas para tamaños de ventana distintos.

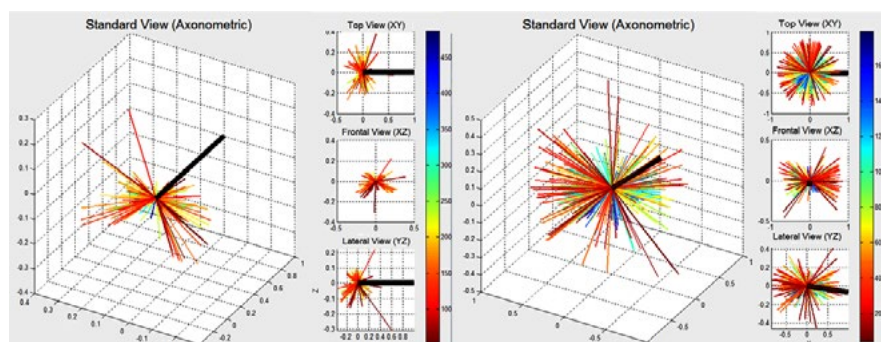


Fig. 2. Resultados obtenidos para distintos tamaños temporales (Izquierda: 1 ms; Derecha: 10ms)

Criterio de selección del sonido directo

Distintos métodos fueron estudiados para la extracción del sonido directo de respuestas al impulso y de sus reflexiones [7, 8]. La técnica utilizada en este trabajo incorpora un algoritmo de extracción de picos que permite al usuario seleccionarlos manualmente, si así lo desea. De esta manera se analizan los picos del nivel de presión sonora de la señal

que distancien en una proporción específica del nivel sonoro máximo. De los puntos encontrados se adopta el primero en orden temporal como centro del sonido directo. Una ventana temporal es aplicada de forma de extraer el sonido directo proveniente de la fuente. Esto puede observarse en la figura 3.

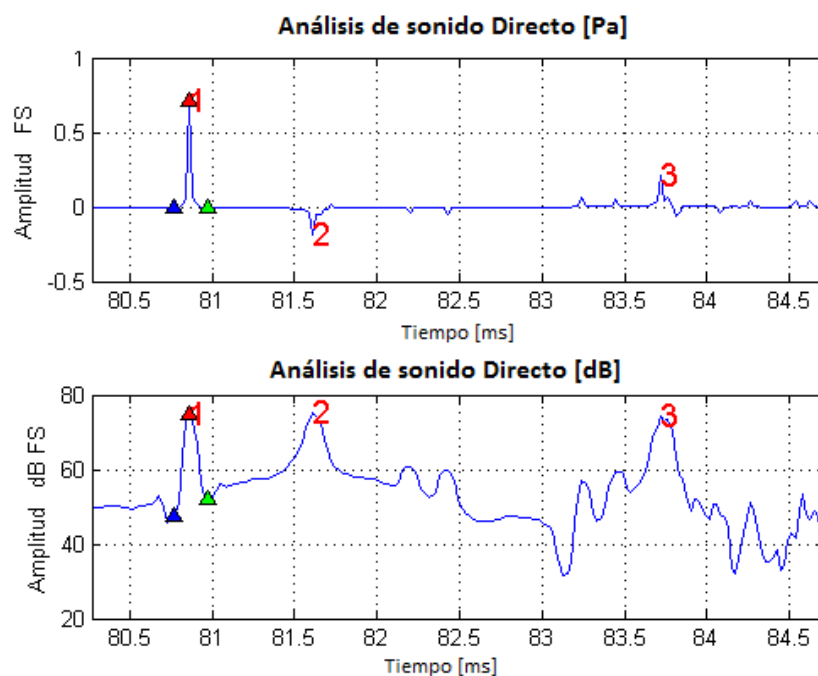


Fig. 3. Extracción del sonido directo

En caso de contar con la respuesta al impulso de la fuente en condición anecoica o de que el sonido directo y las reflexiones se encuentren lo suficientemente separadas temporalmente (una longitud de onda de la frecuencia más baja), es posible realizar un análisis adicional que permite obtener las posiciones específicas de cada una de las reflexiones aun existiendo superposición. El reflectograma (Figura 4) hace uso de la función de correlación cruzada evaluando la ventana encontrada con el vector completo de la respuesta al impulso.

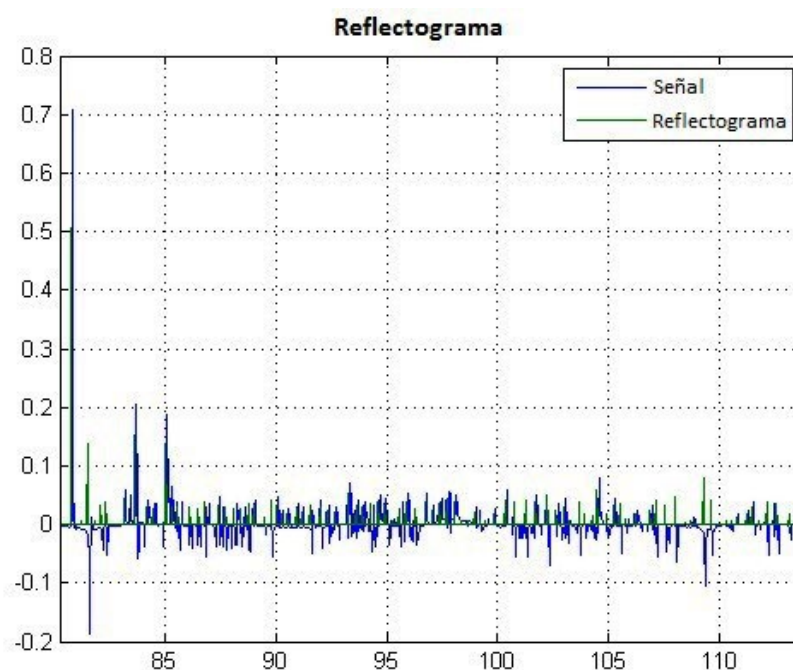


Fig. 4. Reflectograma de la señal

Criterio de selección de ventanas representativas

Otro de los factores importantes radica en la elección de las ventanas más representativas a graficar. La numerosa cantidad de ventanas que resultan del análisis producen en consecuencia una densidad de vectores que imposibilita la correcta visualización de los datos. Por este motivo se opta por

extraer los que tuviesen las mayores magnitudes de nivel de intensidad. Para esta selección también se tiene en cuenta la proximidad de las ventanas, evitando la superposición de las mismas. Este criterio de evitar o no la superposición se encuentran actualmente en discusión. Al considerar superposición de ventanas es preciso aclarar que cada muestra de la respuesta al impulso tendrá asociada una ventana temporal, por lo cual sin importar el tamaño de ventana que se utilice siempre habrá tantas ventanas como muestras del archivo de audio. En la figura 5 se muestra el resultado para los tres casos descritos anteriormente.

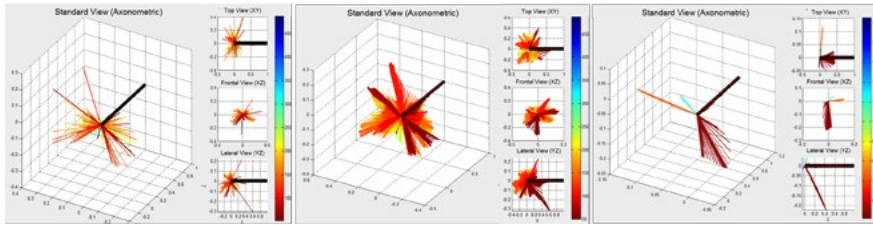


Fig. 5. Distintos criterios de selección de ventanas (Izquierda: Se consideran todas las ventanas; Centro: Se consideran las 256 de mayor nivel; Derecha: Se consideran las 256 de mayor nivel sin superposición)

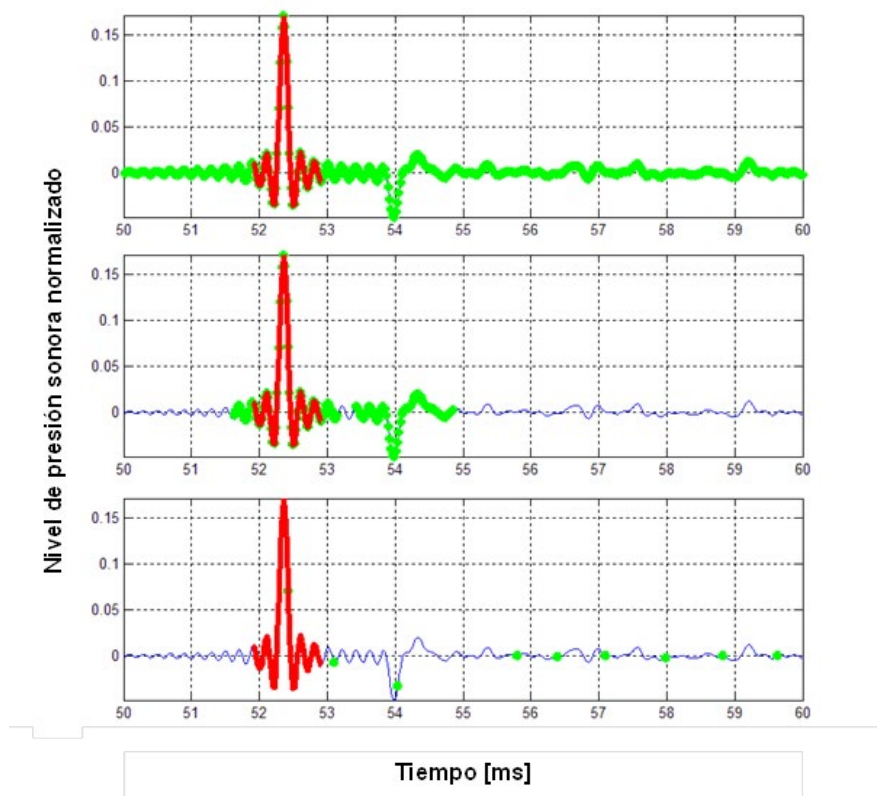


Fig. 6. Distintos criterios de selección de ventanas (Arriba: Se consideran todas las ventanas; Centro: Se consideran las 256 de mayor nivel; Debajo: Se consideran las 256 de mayor nivel sin superposición)

III. Implementación y resultados

Con el fin de analizar los métodos empleados, se utilizaron mediciones conocidas. A continuación se muestra un ejemplo para St. Margaret's Church de York hecha por Aglaia Foteinouy Simon Shelleyen (2011) con un micrófono Soundfield SPS422B y una fuente Genelec S30D. Se utilizó la técnica de LSS de 20 a 22Khz con 15s de duración y 96Khz de frecuencia de muestreo. En la figura 7 se pueden apreciar la planta con las posiciones de medición. También se incluye en la misma grafica los resultados del análisis.

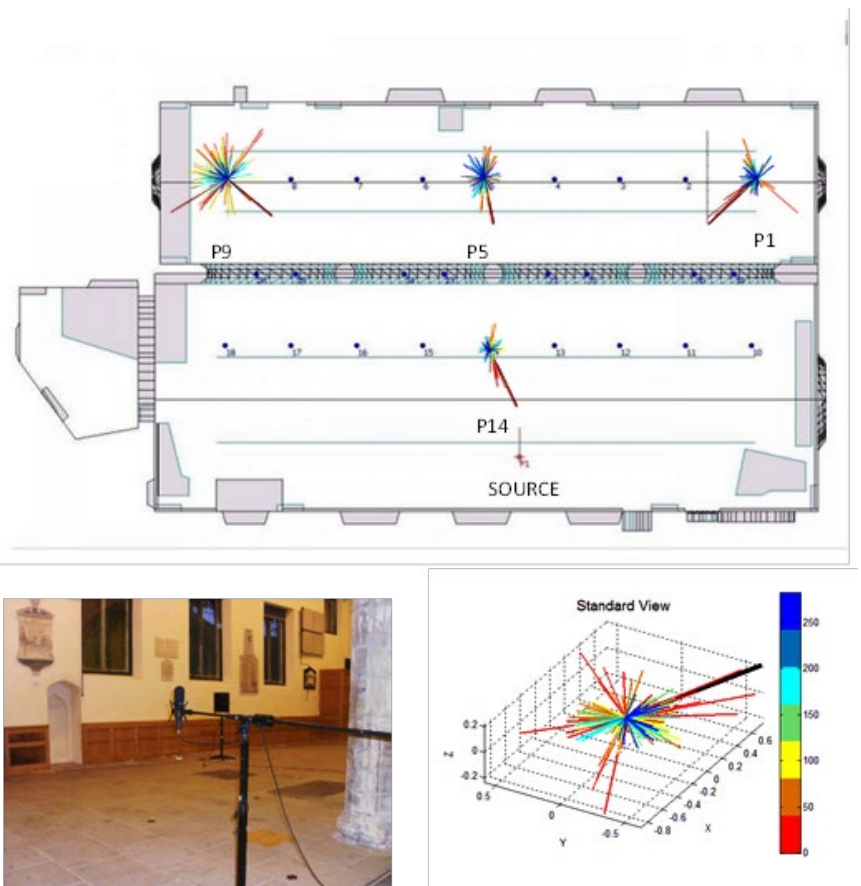


Fig. 7. Puntos de medición y resultados de St. Margaret's Church (arriba); Foto de la medición (abajo-izquierda); Gráfico de la respuesta al impulso 3D para la medición (abajo-derecha)

Esta medición se realizó en distintos puntos ubicados de manera equidistante a lo largo de la habitación. Esta dispersión geométrica es favorable para evaluar la variación en la dirección del sonido directo y sus reflexiones. Se puede observar que los resultados no fueron orientados manualmente hacia la fuente sino que la orientación es producto del algoritmo. Es posible observar como los resultados encontrados en las posiciones cercanas a las esquinas de la sala adquieren componentes laterales comparables en magnitud al sonido directo dada la proximidad a las superficies reflejantes. Situación que no se ve reflejada en las posiciones P5 y P14. Esto advierte que los resultados obtenidos son acordes con la acústica de la sala. Del mismo modo se observa que en la posición P14, el sonido directo adquiere un mayor protagonismo en los vectores graficados.

De forma complementaria se evaluó el sistema completo realizando una medición en una situación real y conocida. Para esto se utilizó una sala preparada acústicamente con material Fonac clase 1 50mm, un micrófono Soundfield SPS200 y una fuente KRK Rokit8. Como se menciona anteriormente, es recomendable realizar una medición de la fuente en condiciones anecoicas, para esto se usaron las técnicas descritas anteriormente con una ventana de 8ms (dadas la geometría de la sala y la mínima distancia a superficies reflejantes), alcanzando una frecuencia inferior de análisis de 121Hz. Usando la técnica de LSS (50s de duración entre 80 y 15Khz con una frecuencia de muestreo de 44100) e implementando la técnica de Kirkeby8, se llevaron a cabo las mediciones de respuesta al impulso. Posteriormente se hizo la conversión de formato A al formato B por medio de la herramienta proporcionada por la firma Soundfield.

En la figura 8 se muestra un esquema de planta con algunos objetos presentes en el momento de la medición. También se incluye en la misma grafica los resultados del análisis

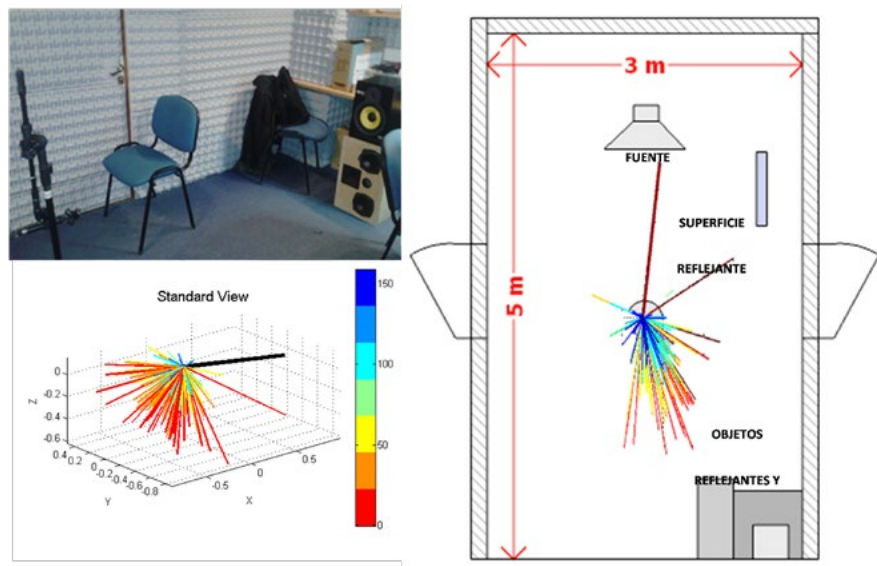


Fig. 8. Planta de puntos de medición y resultados de medición realizada en una sala tratada acústicamente

Es posible de observar en los resultados que la dirección del vector correspondiente al sonido directo mantiene una correspondencia con la ubicación de la fuente. No obstante, ésta se orienta levemente hacia la derecha (donde hay gran componente energético de reflexiones). Esto puede darse debido al solapamiento de reflexiones en la misma ventana temporal, ya que la medición fue realizada en una sala donde la diferencia temporal entre los caminos recorridos es pequeña. También se puede advertir sobre la energía proveniente de las superficies reflejantes y difusoras.

IV. Conclusiones

Un algoritmo para la obtención de respuestas al impulso en tres dimensiones fue desarrollado e implementado. Los resultados obtenidos prueban su correspondencia con las distintas situaciones acústicas evaluadas, evidenciando la validez de la teoría utilizada. De este modo, se puede advertir que el método entraña una herramienta útil y práctica para el análisis de cualquier tipo de entorno acústico. Sus alcances y aplicaciones deben ser estudiados en el futuro a fin de construir las bases para la descripción de problemáticas acústicas. Se advierte que al igual que las respuestas al impulso, otros parámetros pueden ser también analizados mediante la interpretación de vectores de intensidad, por lo que aplicaciones en diversas áreas de la acústica pueden verse beneficiadas con el desarrollo de esta herramienta.

Es necesario señalar que el estudio debe continuarse por la caracterización de precisión en magnitud y dirección de los vectores obtenidos relacionándolos a la situación acústica planteada.

V. Referencias

- [1] Ortolani, F (2015). Introduction to Ambisonics- A tutorials for beginners in 3D audio, 13-16.
- [2] Farina A, (2015). A-format to B-format conversion. (2015). Pc-farina.eng.unipr.it.
- [3] Farina A., Tronchin L. (2013). 3D Sound Characterisation in Theatres Employing Microphone Arrays.
- [4] Farina A. (2000). Simultaneous measurement of impulse res-

ponse and distortion with a swept-sine technique.

[5] Farina A. (2007). Advancements in impulse response measurements by sine sweeps.

[6] Kirkeby O. Nelson P. (1999). Digital Filter Design for Inversion Problems in Sound Reproduction.

[7] Sturm B., Defrance G. (2010). Detection and estimation of arrivals in room impulse responses by greedy sparse approximation.

[8] Usher J. (2010). An improved method to determine the onset timings of reflections in an acoustic impulse response.

AVANCES DEL AUDIO EN LATINOAMÉRICA.



Audio Engineering Society
Sección Colombia



AUDIOJUEGO CON IMPLEMENTACIÓN DE AUDIO BINAURAL

Andrés Carmona / Diego Sierra / Ricardo Moreno

Andrés Carmona*, Diego Sierra** y Ricardo Moreno***
Facultad de Ingenierías, Universidad de San Buenaventura
Medellín, Colombia

*andres.carmona@tau.usbmed.edu.co

**diego.sierra@tau.usbmed.edu.co

***ricardo.moreno@usbmed.edu.co

Resumen. El desarrollo de los videojuegos ha estado enfocado principalmente a la industria del entretenimiento. Por otra parte, el desarrollo de audiojuegos ha centrado su atención en la creación de herramientas didácticas de aprendizaje para población con discapacidad visual. Solo hasta años recientes se ha presentado un auge de este tipo de juegos dirigidos a la industria del entretenimiento. Un audiojuego es un juego electrónico en el que, a diferencia de un videojuego, la interacción se logra con base en la percepción y respuesta a una serie de estímulos sonoros. En este proyecto se describe el desarrollo de un audiojuego implementando técnicas de audio binaural interactivo, que puede ser jugado tanto por personas videntes como con discapacidad visual.

Palabras clave: audio binaural, audio interactivo, audio inmersivo, audiojuego, audio 3D.

Abstract. Video games has been focused mostly to the entertainment industry. Meanwhile the audio games has been directed as didactic element for visually impaired population, only until recent years there has been a boom in this type of games to entertainment. An audio game is similar to a video game, but unlike all the weight of this entertainment is loaded into the sound. This project describes the development of an audio game implementing techniques binaural audio and interactive, that can be play by both sighted and visually impaired players.

Keywords: binaural audio, interactive audio, audio game, sound design, game play.

I. Introducción

Los primeros registros de audiojuegos datan de 1974, con juegos como Touch me, lanzado por Atari. Este consistía en tocar una serie de botones que se iluminaban y producían sonidos de acuerdo a una secuencia generada por el juego en forma automática. El reto era memorizar la secuencia para reproducirla en el mismo orden, incrementando su dificultad con cada ejecución (Kurtz, 2004). Aunque contaba con un componente visual, tanto los jugadores videntes como con discapacidad visual podían disfrutarlo de igual forma. Años más tarde saldrían juegos similares, como Simon o Merlin (Smith, 2011).

Los audiojuegos iniciaron como juegos para personas con discapacidad visual, adaptando juegos comunes a través del sonido, como Blackjack y Batalla Naval (López, 2008). Luego comenzaron a surgir audiojuegos dirigidos a todo público. Uno de los primeros fue Drive, lanzado en el 2002 (Velleman, Tol, Huibets, & Verwey, 2004). Es un juego de carreras sin interfaz visual, basado únicamente en sonidos.

* A. F. Carmona estudia en la Universidad San Buenaventura Medellín (e-mail: andres.carmona@tau.usbmed.edu.co).

** D. Sierra estudia en la Universidad San Buenaventura Medellín (e-mail: diego.sierra@tau.usbmed.edu.co).

*** R. A. Moreno labora en la Universidad San Buenaventura Medellín (e-mail: ricardo.moreno@usbmed.edu.co).

Actualmente se encuentran en el mercado distintos tipos de audiojuegos, como Papa Sangre, lanzado en diciembre del 2010 y desarrollado por el estudio Something Else. Este juego cuenta con implementación de audio binaural a través de su propio motor de audio patentado: Papa Engine. Otros audiojuegos de *Something Else* son: *The Nightjar*, lanzado en el 2011, Papa Sangre II, en el 2013, y Audio Defence: Zombie Arena, en el 2014. En el 2013 la desarrolladora Epicycle presentó Blindside, audiojuego desarrollado en Unity, para iOS, Mac y PC.

La implementación de audio en los videojuegos ha tomado elementos propios del diseño sonoro en la industria cinematográfica, adaptándolos a las dinámicas y requerimientos de un juego, y combinando planos visuales con planos sonoros que –a diferencia del cine tradicional– son planos animados y sonoros programados en secuencias no lineales o interactivas. Prescindir de la imagen en un juego virtual establece retos técnicos y de diseño sonoro para generar la jugabilidad; sin embargo, pueden explorarse aspectos como la espacialidad sonora, para crear elementos de interacción y efectos sonoros que puedan ser aprovechados en pro de la historia del juego. Michel Chion sugiere que escuchar sonidos sin ver o conocer su causa puede crear un impacto emocional enorme, pues esto hace que el sonido se revele en todas sus dimensiones, haciendo que el oyente dirija su atención a rasgos normalmente ocultos del sonido o que pasan desapercibidos (Chion, 1994). Eliminar la imagen pone al espectador o jugador ante una condición de escucha consciente.

La aparición de tecnologías de producción y posproducción permiten crear productos de audio con sonidos que pueden ser percibidos en un determinado punto del espacio 3D, como los sistemas surround y los métodos de grabación binaural. Las técnicas de espacialización sonora en videojuegos han evolucionado conforme a estos formatos y sistemas de reproducción del sonido (Breebaart & Faller, 2007). Francis Rumsey muestra un resumen general de los esquemas de programación, la relación de los conceptos de síntesis para el procesamiento de audio, el diseño de los ambientes sonoros y la implementación de técnicas de mezcla envolvente para ubicación de fuentes sonoras virtuales en sistemas de reproducción 5.1 y 7.1 (Rumsey, 2013). Si bien los formatos envolventes han sido usados comercialmente en la industria de los videojuegos (Miller, 2006), son pocos los referentes que se tienen acerca del uso de técnicas binaurales. En el presente proyecto se describe el uso de técnicas de captura binaural de sonido usando cabezas artificiales, y la implementación de herramientas para el diseño y programación de eventos sonoros dinámicos e interactivos que permitieron la creación de un audiojuego titulado El Bosque. A continuación se describen los procesos realizados en las etapas de preproducción, producción y posproducción del audiojuego.

II. Parte técnica del artículo

A. Preproducción

En esta primera etapa se concibieron todos los detalles relativos al juego, plasmados en un documento de diseño (Schell, 2008) que contiene desde la historia hasta el concepto del juego, y un desglose de cada elemento de audio utilizado.

1. Concepción del juego

En el audiojuego El Bosque, el jugador se representa en primera persona por un personaje que despierta en medio de un bosque oscuro. El bosque es un lugar lleno de misterio, en cuyo terreno hay una serie de trampas. Al acercarse a ellas se emiten sonidos espacializados que indican su ubicación. El jugador debe prestar atención a los eventos sonoros que emiten las trampas, para poder cambiar el trayecto y esquivarlas. El objetivo del jugador será salir del bosque guiado únicamente por el sonido de una campana que proviene del pueblo más cercano. Deberá salir de allí en un tiempo límite, esquivando las trampas sin detenerse; de lo contrario, en poco tiempo escuchará la presencia y persecución de un enemigo desconocido.

2. Diseño

Se planteó un juego para plataforma PC, tipo arcade, por niveles de dificultad, enfocado a un rango de edad entre trece y veintisiete años. Tiene menús hablados, para que pueda ser jugado por personas con discapacidad visual. En la Fig. 1 se presenta la estructura general del juego, con cada una de las escenas y las distintas opciones a elegir en cada menú.

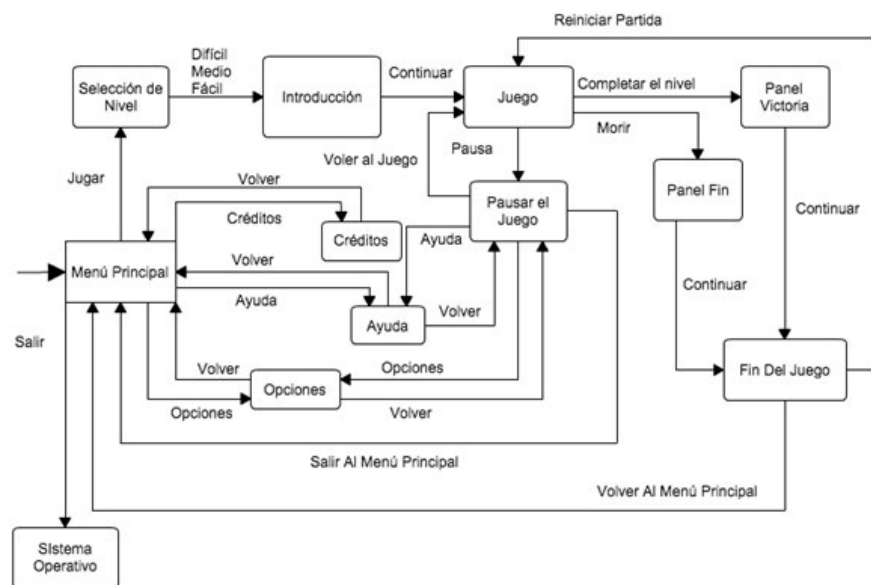


Fig. 1. Diagrama de flujo de pantallas en el juego
Producción

B. Producción

1. Programación

El proyecto fue desarrollado en la plataforma Unity, programando scripts en C# para cada una de los elementos que componen el juego; como por ejemplo los menús, personajes, movimientos, trampas, inteligencia artificial del enemigo, eventos interactivos, etc. Las interacciones del jugador con su entorno –según las entradas y los fenómenos indirectamente controlados por él– son traducidos en información sonora.

1.1. Elaboración de scripts para el desarrollo de menús

Aunque El Bosque es un juego que se basa únicamente en sonidos, cuenta con una interfaz gráfica básica para los menús: principal, pausa y fin del juego. Los menús también actúan como menús hablados, posibilitando el acceso a personas con discapacidad visual e indicando la tecla que deben seleccionar de acuerdo a la opción deseada en cada menú.

1.2. Edición de script primera persona

Se implementó un asset predeterminado de Unity para los movimientos y las propiedades físicas del personaje. Toda la implementación de los audios binaurales se desarrolla en esta fase. Se agregaron cinco fuentes, que son activadas por medio de variables booleanas, contadores, comparadores y las funciones OnTriggerEnter (cuando el personaje choca con un interruptor, el cual activa un evento en el juego) y OnCollisionEnter (cuando choca con un obstáculo), que emite un sonido aleatorio, dependiendo de la fuente activada.

1.3. Enemigo, trampas, contador y fin del juego

Son los sucesos por los cuales se da por terminado el juego.

1.3.1. Enemigo

El enemigo que persigue al jugador durante la partida se compone de una inteligencia artificial que consta de dos estados: ambulante y persecución. Cuando el jugador está estático, el enemigo lo persigue; cuando el jugador se mueve, el enemigo también lo hace, pero este avanza de manera aleatoria por el escenario. Si el jugador es alcanzado por el enemigo, se dará por terminada la partida.

1.3.2. Trampas

Son obstáculos que, al hacer contacto con ellos, se activan y generan la terminación de la partida. Como se muestra en la Fig. 2, una trampa se constituye de varias zonas: la sección A permite que al entrar y salir se reinicie la trampa; la sección B genera una alerta audible, indicándole al jugador que está cerca de una trampa y la ubicación de la misma; esta depende del ángulo con que entra el jugador a la trampa. La sección C da por terminada la partida.

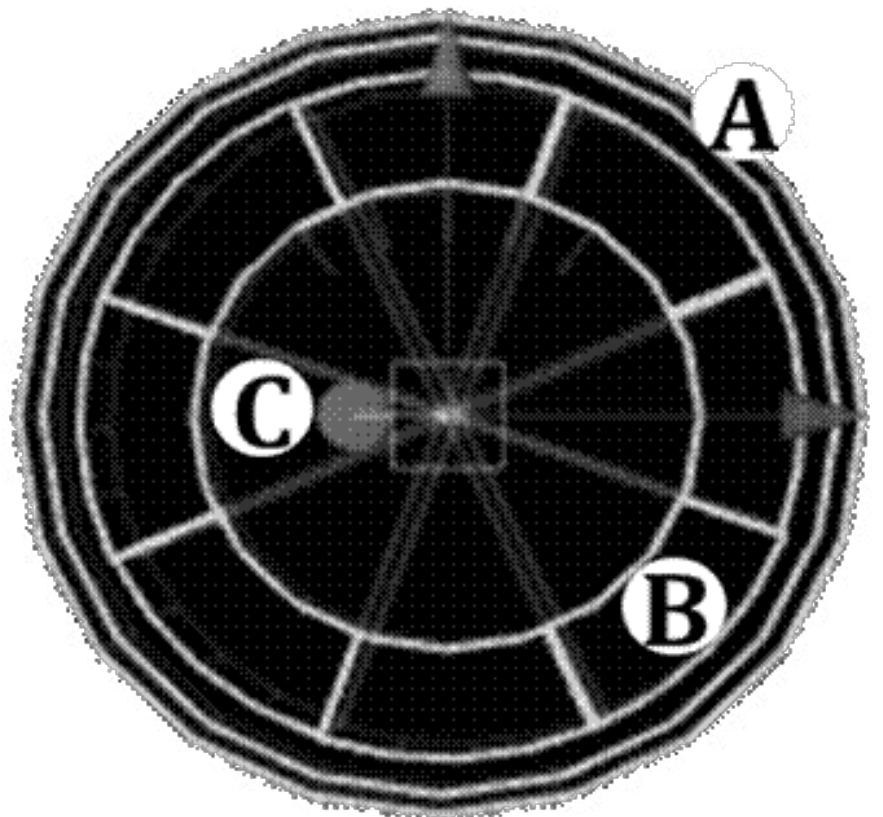


Fig. 2. Secciones de una trampa

Con relación al número de trampas, esta varía según el nivel de dificultad seleccionado. Hay tres niveles: fácil, medio y difícil; con ocho, catorce y veinticuatro trampas respectivamente.

1.3.3. Contador

La partida tiene un tiempo límite de diez minutos. Si el conteo llega a cero, el jugador perderá.

1.3.4. Fin del juego

Para ganar la partida el jugador tendrá que llegar al pueblo, guiado por el sonido de una campana, y sin haber activado una sola trampa durante todo el juego, no ser atrapado por el enemigo y llegar por debajo del tiempo límite.

2. Producción de audio

2.1. Grabación

Se decidió limitar en lo posible el uso de librerías externas, por lo que se creó un banco sonoro exclusivo para el juego. Por ende, se realizaron diferentes grabaciones en distintos lugares. Todas se realizaron con una frecuencia de muestreo de 44100 Hz y 16 Bits, de acuerdo al estándar de audio profesional propuesto en el libro rojo (Philips and Sony Corporation, 1980).

2.1.1. Grabación de efectos

La grabación de efectos se realizó en el estudio A de la Universidad de San Buenaventura, Medellín. En la Fig. 3 se muestran los dos modelos de cabeza artificial usados, uno de ellos con torso.

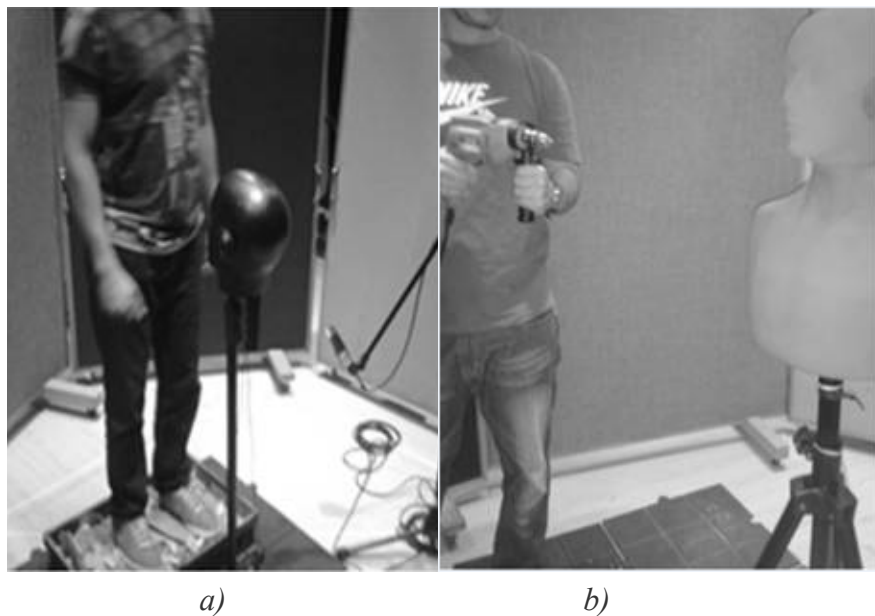


Fig. 3. Grabación de efectos: a) cabeza binaural, b) cabeza binaural 01 dB con torso

En la Fig. 4 se muestran los ángulos tenidos en cuenta para la grabación. Esto con el fin de tener una cobertura de 360° en el plano horizontal (Masiero, 2012).

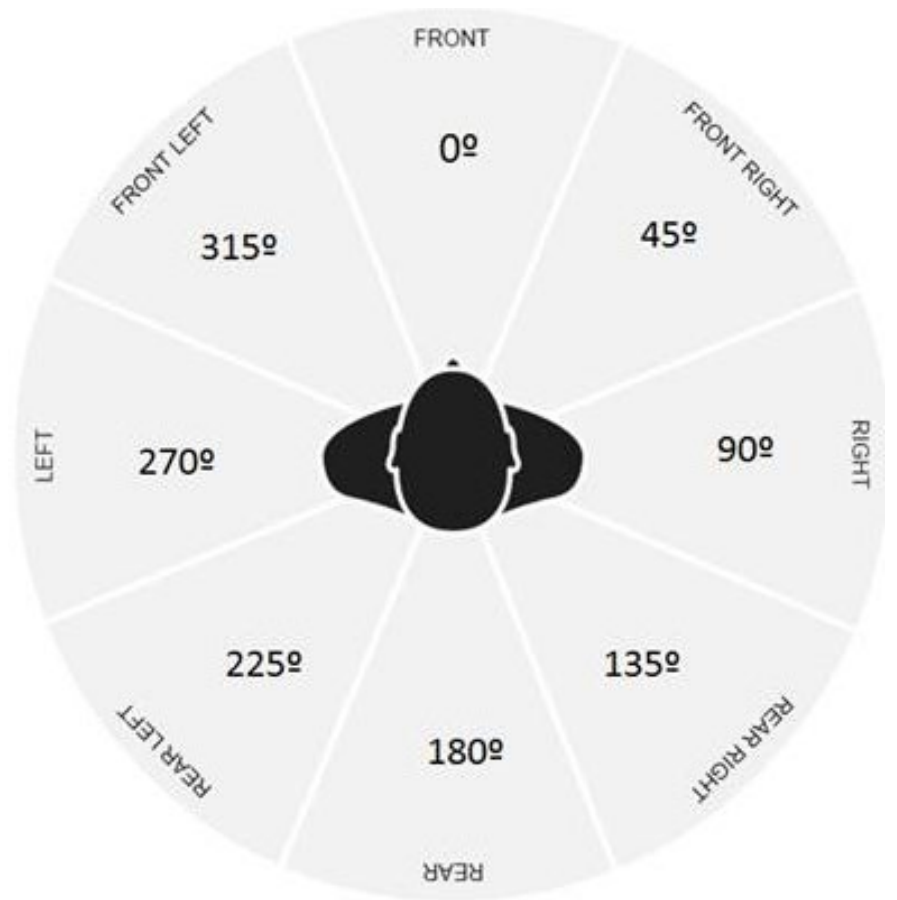


Fig. 4. Distribución esquemática en el plano horizontal, utilizada para la grabación con las cabezas binaurales (Lalwaina, 2015)

2.1.2. Banco sonoro

Para la grabación del banco sonoro requerido por el juego, se realizó un scouting* de locaciones, entre las cuales se seleccionó el Parque Zoológico Santa Fe, de la ciudad de Medellín. Las capturas se hicieron con el prototipo de cabeza binaural (Fig. 5a), el micrófono superdireccional Sennheiser ME 66 (Fig. 5b) y un par de grabadoras Zoom H4. En este lugar se logró grabar una gran variedad de animales: mamíferos (ovejas y monos), aves (guacamayas, grullas, patos, loros, guacharacas, sinsontes, flamencos), y varias cascadas y fuentes de agua en las que reposan algunos de ellos.

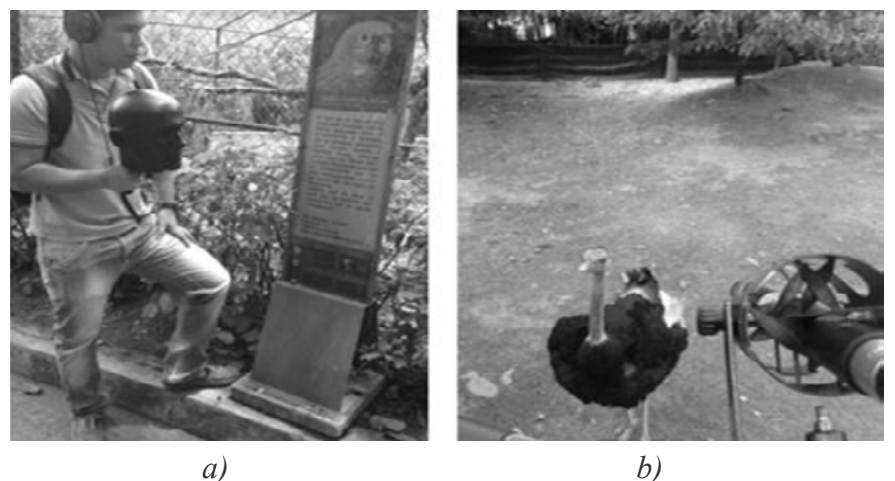


Fig. 5. Capturas en el zoológico: a) prototipo de cabeza binaural, b) micrófono superdireccional Sennheiser ME 66.

2.1.3. Ambientes

En la Fig. 6 se muestra la grabación de ambientes realizada en el museo El Castillo, ubicado en la ciudad de Medellín, lugar rodeado de jardines y zonas arboladas. El registro se

* Proceso de análisis para selección de locaciones y escenarios para la grabación.

realizó en las horas de la madrugada, para evitar el ruido producido por el tráfico vehicular cercano a este lugar.

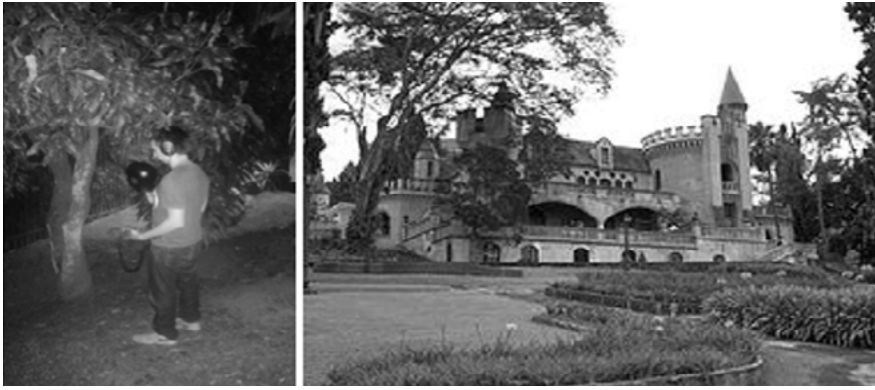


Fig. 6. Grabación de ambientes

2.1.4. Diálogos

Los menús hablados se grabaron en estudio con el micrófono de condensador Neumann Kms 105.

2.2. Música

La música utilizada para el juego fue realizada en Ableton Live. Esta solo se reproduce durante el inicio del audiojuego y no durante la partida, para no sobreponerse a los ambientes y efectos binaurales que transcurren durante el juego.

3. Implementación de audio

Para el juego se emplearon los motores de audio de Unity y de Fmod. Ambos constan de parámetros que se encargan de reproducir y captar los sonidos.

Los sonidos son emitidos por audio sources, que luego son captados por un audio listener, el cual está adjunto a la cámara principal y se encarga de mostrar las diferentes escenas del juego.

Según la necesidad, en algunos casos se asignó al audio source un audio mixer, que permite aplicar una atenuación de volumen, insertar efectos que procesen la señal de audio, y enviar y devolver (send & return) la señal de un bus a otro. La Fig. 7 muestra cómo es el enrutamiento de señales de audio en Unity.

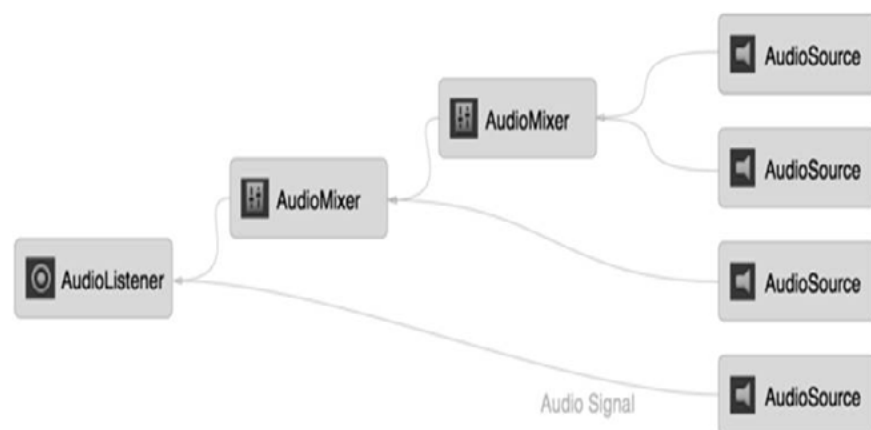


Fig. 7. Enrutamiento de señales de audio en Unity

El audio source cuenta con una propiedad llamada spatial blend, que establece qué tanto del motor 3D tiene efecto en este, asignando 0 para que el sonido no sea procesado (usado para los sonidos grabados binauralmente) y 1 para espacializarlos (usado para los sonidos que no pudieron ser

registrados de forma binaural).

Los eventos de Unity son emitidos mediante pulsadores, los cuales reproducen sonidos binaurales grabados en cabezas artificiales.

El motor de audio de Fmod trabaja mediante eventos, a los cuales se les puede asignar diferentes propiedades. Estos eventos se trabajaron en tres formas distintas: la primera, programación por ejecución de eventos y parámetros; la segunda, mediante la activación de sonidos instantáneos con la función PlayOneShot; y la tercera, mediante sonidos estacionarios desde el event_emitter proporcionado por Fmod. La recreación del bosque se hace por medio de scatterer sounds. Estos son unos contenedores donde se pueden agregar varios sonidos y configurar su comportamiento a través de nodos, como random (escoge sonidos aleatoriamente), modulator (altera el pitch de los sonidos aleatoriamente), attenuator (atenúa el volumen de los sonidos a partir de un determinado umbral que se asigne), entre otros. Por ejemplo, con cinco cortos sonidos de pájaros, mediante la asignación de nodos, es posible crear varias horas de cantos de muchos pájaros, emulando una situación específica requerida por el juego. En la Fig. 8 se hace un esquema del proceso.

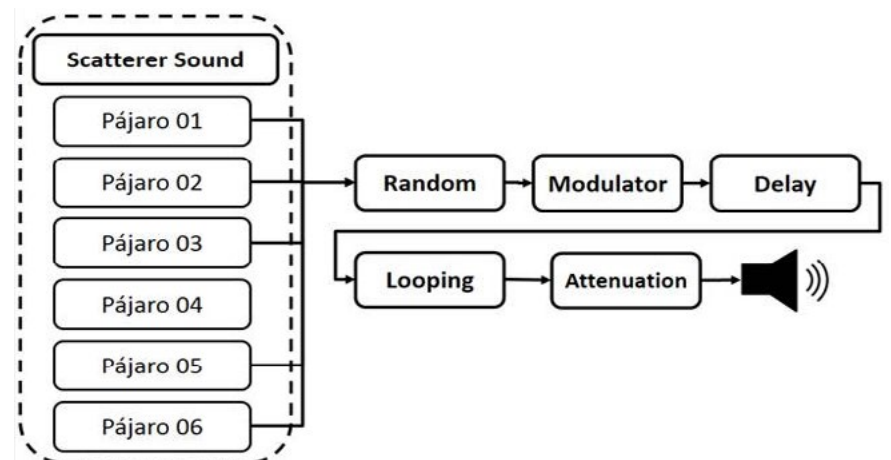


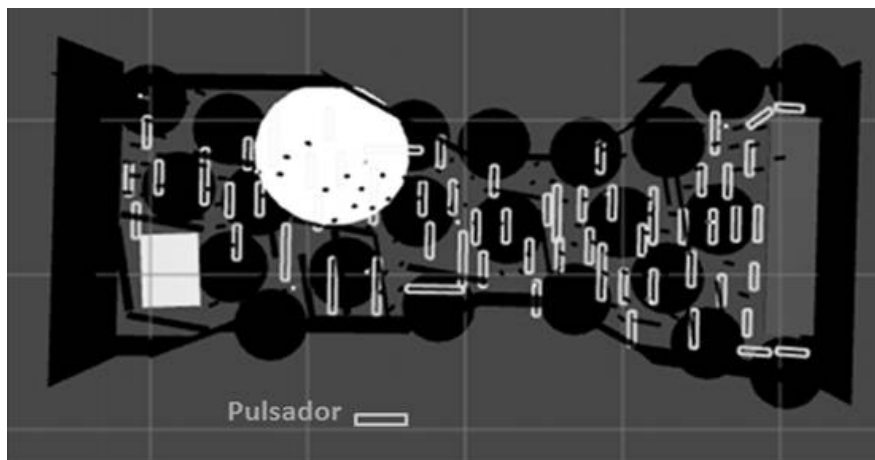
Fig. 8. Esquema de recreación del canto de pájaros a través de un scatterer sound

C. Posproducción

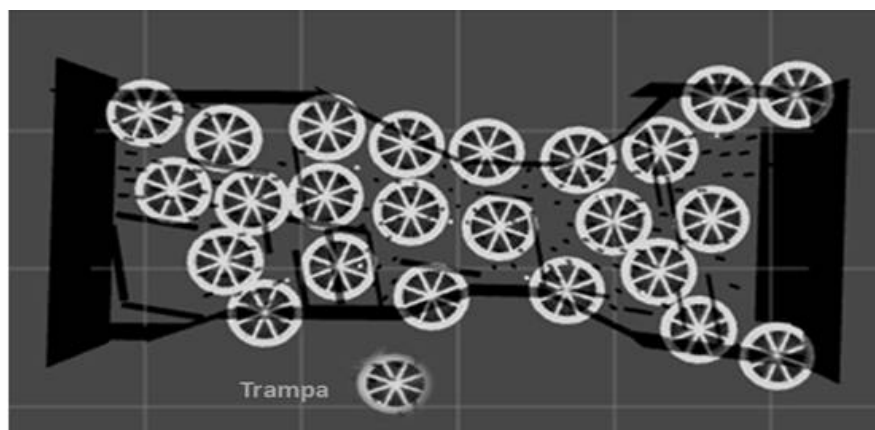
El Digital Audio Workstation (DAW) utilizado para la edición de sonido del juego fue Reaper. Los ambientes grabados se limpiaron filtrando en algunos casos frecuencias bajas y altas. Las grabaciones en el zoológico tuvieron mayor trabajo de edición, por la gran cantidad de ruido de fondo contenido. El proceso de edición fue cuidadoso, para no afectar las diferencias interaurales de tiempo y nivel de las señales binaurales, preservando sus condiciones de espacialidad.

Pruebas

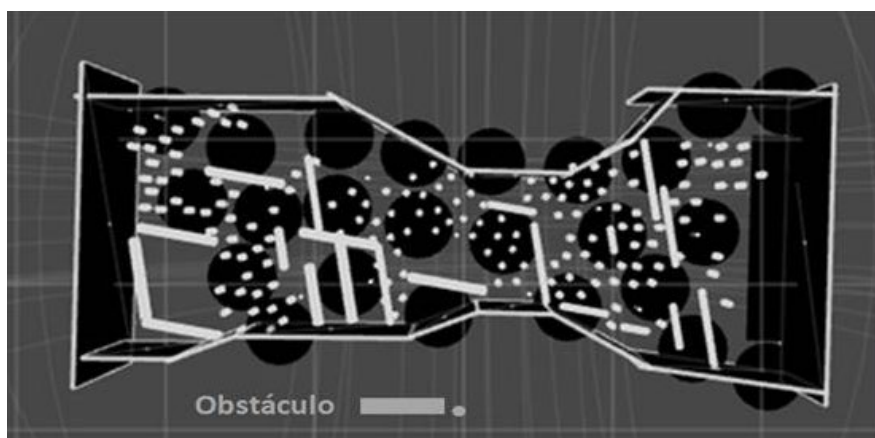
En la Fig. 9 se muestra un mapa de todos los elementos que posee el escenario del juego, distribuidos en lo que fue el terreno del bosque. En esta etapa el audiojuego adquiere funcionalidad, para testear y verificar el estado de cada uno de los eventos interactivos programados.



a)



b)



c)

Fig. 9. Elementos que componen el escenario del juego en Unity: a) pulsadores, b) trampas y c) obstáculos

III. Resultados

El procesamiento de audio se analizó en cuatro casos, tomando como referencia el porcentaje de audio manager, como se muestra en la Fig. 10. En primer lugar las escenas del menú que se componen por la interacción de botones, el menú hablado y la música de fondo; el procesamiento de la CPU en este caso fue de 0.53 ms, donde el audio consumió el 5.3% del total de procesamiento. A continuación se evaluó el uso de sonidos binaurales mientras se corre juego, donde el procesamiento de la CPU fue de 1.41 ms y el audio consumió el 2.0% del total de procesamiento. En la tercera parte se evaluaron los sonidos que hacían uso del motor 3D de Unity, el procesamiento de la CPU fue de 2,30 ms y el audio consumió 2.7%. Finalmente se evaluaron los sonidos que implementan el motor de audio de Fmod, para este caso el procesamiento de la CPU momento fue de 2,52 ms y el audio consumió el 3,8% del total de procesamiento.

Hierarchy	CPU:0.53ms	GPU:0.00ms	Frame Debugger				
Overview	Total	Self	Calls	GC Alloc	Time ms	Self ms	
▶ Camera.Render	30.6%	4.2%	1	0 B	0.16	0.02	
▶ GUI.Repaint	19.1%	13.5%	1	0 B	0.10	0.07	
Overhead	16.3%	16.3%	1	0 B	0.08	0.08	
▶ BehaviourUpdate	12.4%	2.9%	1	0 B	0.06	0.01	
▶ AudioManager.Upda	5.3%	4.0%	1	0 B	0.02	0.02	
▶ EnlightenRuntimeMa	5.2%	3.3%	1	0 B	0.02	0.01	
Profiler.FinalizeAndS	5.0%	5.0%	1	0 B	0.02	0.02	
▶ MonoBehaviour.OnM	3.5%	0.3%	1	0 B	0.01	0.00	
Canvas.SendWillRenc	1.1%	1.1%	1	0 B	0.00	0.00	
Cleanup Unused Cac	0.1%	0.1%	1	0 B	0.00	0.00	
▶ Graphics.PresentAnc	0.1%	0.0%	1	0 B	0.00	0.00	
▶ Loading.UpdatePrelc	0.1%	0.1%	1	0 B	0.00	0.00	

Fig. 10. Panel de procesamiento: modo desarrollador de Unity mientras se ejecuta el menú principal

La implementación de grabaciones binaurales, en efecto, por no tener que ser procesadas, permite aligerar el uso de la CPU. El procesamiento disminuye con respecto al uso del motor 3D de Unity y Fmod, dado que estos motores computan las posiciones de la fuente y del oyente, procesan parámetros como atenuación por distancia, efecto dopler, reverberación, filtros, propagación, entre otros. Estas herramientas ofrecen muchas posibilidades para el desarrollo de juegos y son eficaces acústicamente, sin embargo los hacen computacionalmente más costosos.

En definitiva se obtuvo un audiojuego que implementa audio binaural. Como consecuencia del efecto crosstalk en los altavoces estéreo convencionales, el cual interfiere con la reproducción binaural (Kotorynski, 1990), es necesario el uso de auriculares para jugarlo. El producto final se presenta como un ejecutable para plataformas Macintosh y Windows. El peso final del juego es de 171 MB, se evitó usar formatos de audio comprimidos para garantizar una calidad de audio máxima; igualmente, al no ser un formato comprimido es rápido de reproducir (Trinit, 2009), a costa de ocupar bastante espacio, pero esto se compensa al no haber almacenamiento ni procesamiento de gráficos.

El banco sonoro resultante está desglosado en doscientos veintiún archivos de audio en formato wav, a una resolución de 44100 Hz y 16 Bits. En su mayoría son audios grabados binauralmente, resultantes de las distintas capturas hechas tanto en el estudio de grabación como en exteriores.

IV. Conclusiones

Este tipo de software puede ser implementado en el aprendizaje de niños con discapacidad visual. En futuros proyectos se podría investigar y elaborar softwares orientados en el desarrollo de estructuras cognitivas de representación espacial, a través de la interacción con ambientes virtuales basados en sonidos.

La implementación de sonido binaural fue una herramienta idónea para el desarrollo del juego, por su capacidad de generar una mayor inmersión y un sentido más realista de la espacialización. Adicionalmente, esto se potencia con el uso de técnicas interactivas de programación de audio y con las intenciones del diseño sonoro, presentándose como alternativa en una industria donde el sonido en los videojuegos ha alcanzado un estado de madurez que implica que esté a la paridad de desarrollos gráficos. Para esto es necesaria

la implementación de técnicas de espacialización sonora en aras del realismo.

El siguiente paso para el proyecto es hacer pruebas relacionadas con la evaluación de ciertos parámetros del juego, como el tiempo límite establecido para cada nivel de dificultad o si el método para la detección de las trampas es funcional de acuerdo a la interactividad y espacialidad planteadas como objetivo del juego. También sería interesante evaluar entre la dos métodos usados (grabación binaural y motores 3D) cual presenta mayor nivel de realismo e inmersión. A nivel personal y como una primera impresión, la grabación binaural se presenta más realista, pero tiene un inconveniente que radica en la función HRTF, que cambia con cada jugador, de manera que las sensaciones que se pueden percibir, con una grabación que introduce una HRTF distinta a la del individuo, pueden ser confusas. Estas pruebas deberán hacerse tanto con personas videntes como con las que tienen discapacidad visual. Además, para la evolución del juego se debería llevar a plataformas móviles (Android, iOS), posibilitar otros idiomas, agregar características adicionales (como nuevos niveles, escenarios, enemigos y progresión de habilidades en el personaje) e implementar la combinación de sonidos capturados binauralmente con sonidos mono, que a través de herramientas de síntesis binaural puedan ser espacializados en tiempo real.

V. Referencias

Breebaart, J.; & C. Faller (2007): Spatial audio processing: MPEG Surround and other applications. John Wiley & Sons.

Chion, M. (1994): Audio-Vision: sound on screen. New York: Columbia University Press, 14th ed.

Kotorynski, K. (1990): Digital binaural/stereo conversion and cross-talk cancelling. Presented at the 89th convention, 1990 September 21-25, Los Angeles: Audio Engineering Society.

Kurtz, B. (2004): Encyclopedia of arcade video games. Atglen, PA: Schiffer Publishing.

Lalwaina, M. (2015): Surrounded by sound: how 3D audio hacks your brain. Obtenido de: <http://www.theverge.com/2015/2/12/8021733/3d-audio-3dio-binaural-immersive-vr-sound-times-square-new-york>

López, C. (2008): «Breve historia de los mundos virtuales». En AA. VV.: Arte y arquitectura digital, net.art y universos virtuales. Conferencias presentadas en las Jornadas Internacionales celebradas los días 5 y 12 de marzo del 2008 en la Facultad de Geografía e Historia de la Univ. de Barcelona. Barcelona: Univ. de Barcelona, pp. 135-42.

Masiero, B. S. (2012): Individualized binaural technology: measurement, equalization and perceptual evaluation. PhD tesis. Berlín: Logos Verlag.

Miller, R. E. (2006): 5.1 surround and 3D (full sphere with height) reproduction for interactive gaming and training simulation. Audio Engineering Society.

Philips and Sony Corporation. (1980): System description compact disc digital audio (Red book).

Rumsey, F. (2013): «Game audio: transforming simulation and in-

teractivity». Journal of Adolescent & Adult Literacy, vol. 61, n.º 7/8, julio-agosto, pp. 609-12.

Schell, J. (2008): The art of game design: a book of lenses. Elsevier.

Smith, M. (2011): Simon who? The story behind a playground favorite.

Trinit. (2009): Efectos de sonido. En Programación de videojuegos. Zaragoza: Asociación de Informáticos de Zaragoza.

Velleman, E.; Tol, R. v.; Huibets, S., & H. Verwey (2004): «3D shooting games, multimodal games, sound games and more working examples of the future of games for the blind». En K. Miesenberger, J. Klaus, W. Zagler, & D. Burger (eds.): Computers helping people with special needs. Berlín: Springer-Verlag, pp. 257-63.



AVANCES DEL AUDIO EN LATINOAMÉRICA.



Audio Engineering Society
Sección Colombia



MÉTODOS ALTERNATIVOS DE ESPACIALIZACIÓN COMPUESTA [PARTE 1]

Juan David Sierra

Juan David Sierra
Facultad de Artes, Pontificia Universidad Javeriana
 Bogotá, Colombia
 jdsierral@icloud.com

Resumen—El presente proyecto procura hacer un análisis de diferentes algoritmos de espacialización para desarrollar múltiples métodos alternativos de generar posiciones aparentes de fenómenos sonoros. Sin embargo, más allá de todo el desarrollo investigativo, este proyecto representa un esfuerzo por lograr una conjunción clara entre diferentes ramas del conocimiento como la matemática, la física, el análisis de señales y sistemas y los lenguajes de programación en pro de un objetivo artístico. En últimas el lector encontrará la descripción detallada detrás del desarrollo de una suite de algoritmos en forma de plug-ins que sirven a un solo propósito. Generar una explosión de posibilidades en mezcla a la hora de la manipulación de la imagen estéreo.

Palabras clave: Audio 3D, Binaural, Espacialización, Plug-ins, Software.

Abstract—The present project intends to do an analysis of different spatialization algorithms in order to develop multiple alternative methods to simulate apparent positions of sound phenomena. Nevertheless, beyond the research development, this project represents an effort to clearly merge different types of knowledge such as mathematics, physics, signal and system analysis, programming languages, etc... pointing towards an artistic objective. At last the reader will encounter a detailed description behind the development of a suite of algorithms presented as plugins for a single purpose: They pursue a burst of possibilities in mixing engineer regarding the manipulation of stereo imaging.

Keywords— 3D Audio, Binaural, Spatialization, Plug-ins, Software.

I. Introducción

En general el consumo actual de la música se realiza de una manera despreocupada y poco concienzuda y en el mejor de los casos, si el escucha tiene plena conciencia de la música y el arte que de ella se deriva, es poco frecuente que tenga una preocupación equiparable o si quiera comparable por obtener un buen sonido dentro de su experiencia auditiva.

Ello se evidencia en las claras falencias en el posicionamiento de sistemas de reproducción, en la falta de cuidado con la percepción espacial dentro de un registro sonoro, en la escasa atención que se le presta a la calidad de los componentes y equipos de reproducción en general y en últimas, a la falta de importancia y credibilidad sobre los sistemas de reproducción que ofrecen nuevas y mejores tecnologías para enriquecer el proceso auditivo como lo son los formatos multicanal. De hecho, si bien la reproducción en estéreo se ha ganado una gran parte de los audio-escuchas, existen nuevas tendencias comerciales y culturales que han llevado a un incremento en el uso de sistemas monofónicos por parte de los usuarios promedio. Por ejemplo, sistemas portátiles que electrónicamente funcionan en estéreo pero la posición y funcionamiento acústico de los parlantes elimina totalmente los desarrollos en espacialización.

Debido a esto, es importante incentivar el desarrollo de nuevas y mejores tecnologías que acerquen a los consumidores de fonogramas a mejores y novedosas experiencias auditivas y a la vez procurar la educación del público general para que logre valorar dichos procesos. De esta manera buscar que

tengan un acercamiento mucho más consciente a aquello que es el sonido y el papel que juega dentro de la música y en últimas dentro del arte.

Adicionalmente, los procesos técnicos asociados a la producción sonora y al procesamiento de señales en pro del arte han tenido grandes desarrollos en diversas ramas; los ecualizadores son cada vez mejores y de mayor variedad, los compresores son cada vez más transparentes y con mayor diversidad colorística, y así mismo procesos utilizados en la espacialización han sufrido desarrollos inimaginables como sucede con las reverberaciones tanto algorítmicas como de convolución. Sin embargo, hay un proceso que se ha mantenido casi estático desde el surgimiento del estéreo. Éste es el proceso de panoramización o paneo el cual si bien es efectivo a la hora de localizar una señal en una imagen estéreo limitada, es innegable que es un método bastante primitivo si se compara con todos los desarrollos en materia de espacialización asertiva y natural dentro de la percepción humana.

En otras palabras, el paneo funciona, pero diferentes estudios psicoacústicos evidencian que la diferencia de nivel no es un parámetro suficiente para hacer creer al cerebro que una fuente sonora está ubicada en una posición en específico. Diferentes marcadores psicoacústicos y la conjunción de los mismos son esenciales para una espacialización realista.

Sin embargo, no sólo no es común en el mercado un proceso así, sino que no es solicitado por los ingenieros de mezcla en general. De hecho, esto abre las posibilidades a una investigación que puede satisfacer un mercado con una clara necesidad técnica de la cual aún no se percata y a la vez enriquecer las experiencias auditivas de los audio-escuchas para forjar mejores procesos de mezcla.

II. Metodología

La principal propuesta de emulación será un algoritmo compuesto en el que se encuentren anidadas las múltiples herramientas de espacialización en un solo procesador. Sin embargo; para desarrollar este algoritmo, se debe separar el problema en partes mas pequeñas con el fin de resolver, poco a poco, y de manera analítica marcadores psicoacústicos diferentes (tales como Diferencias Interaurales de Tiempo (ITD), Diferencias Interaurales de Intensidad (IID), cambios espectrales, reflexiones tempranas y reverberación) e ir moldeando estos análisis de acuerdo a las necesidades de un ingeniero de mezcla.

Con el fin de emular estos marcadores, debe hacerse un procesamiento básico de señal y para ello, será necesario desarrollar algoritmos que permitan determinar todas las variables a partir de un número selecto de parámetros. Esto resultará en una ecuación o un conjunto de ecuaciones que serán posteriormente codificadas en un lenguaje de programación tal como C++ y así generar un plug-in.

A. Paneo (IID)

El primer diseño algorítmico requerido es el de un simulador simple de IID que será básicamente un control de paneo. Sin embargo, el paneo requiere cierta comprensión sobre el comportamiento de ondas sonoras reproducidas a través de parlantes. Debido a la expansión tridimensional de las ondas en un cuarto, no existe una superposición perfecta de

ondas en su interior; a pesar de que la amplitud obtenida a través de la reproducción simultánea de un programa común a través de varios parlantes depende en gran medida de la acústica de la sala, en general una ganancia de 3dB (en lugar de 6dB esperada de una superposición perfecta) es comúnmente aceptada; lo anterior se conoce como la ley del paneo. Con el fin de simular el comportamiento deseado, se asume que el eje independiente representa la posición p de paneo entre 0 y 1 y por consiguiente la atenuación a medio camino (0.5) debería tener un valor de -3dB, es decir, un factor de 0.747. Para lograr esto, se definen las siguientes funciones exponenciales que determinan la atenuación:

$$y_l(t) = x_l(t) \cdot p^{0.5} \quad (1)$$

$$y_r(t) = x_r(t) \cdot (1 - p)^{0.5} \quad (2)$$

Sin embargo, si se requiere otra definición para la ley del paneo, se requiere una generalización. Asuma que d es la profundidad del paneo utilizado y s es el exponente de la función exponencial requerida para generar la atenuación de la profundidad del paneo de la señal. En el caso de una simple ley del paneo de 3dB, el siguiente algoritmo describe dicha relación. $-3 = 20 \cdot \log(0.5^{0.5})$ y cuando se evalúa para una ley del paneo de 6dB la expresión resulta ser $-3 = 20 \cdot \log(0.5^{0.5})$. Por tanto, una expresión general para definir este exponente s sería:

$$-d = 20 \log(0.5^s) \quad (3)$$

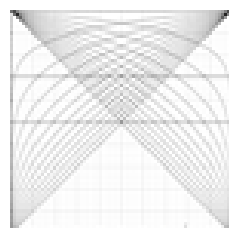
$$\log_{0.5}(10^{-d/20}) = s \quad (4)$$

$$\frac{\log 10^{-d/20}}{\log 0.5} = s \quad (5)$$

$$\frac{-d}{20(-0.3)} \approx s \quad (6)$$

$$d/6 \approx s \quad (7)$$

Por lo tanto, para lograr un paneo adaptativo, la función que define la ganancia de cada canal sería:



$$y_l(t) = x_l(t) \cdot p^{d/6} \quad (8)$$

$$y_r(t) = x_r(t) \cdot (1 - p)^{d/6} \quad (9)$$

Fig. 1. Posibles Ganancias para distintos ajustes de la ley del paneo.

B. Delay (ITD)

Un algoritmo de retraso sencillo guarda una señal en un buffer y la lee en un tiempo posterior para volver a reproducir la señal. Si se asume una señal de entrada $x(t)$ y una señal de salida $y(t)$ lo anterior se expresaría como:

$$y(t) = x(t - t_d) \quad (10)$$

Sin embargo un análisis simple muestra los inconvenientes en hacer esto. Si se asume una distancia interaural promedio de 17cm, el retraso mas largo que se requiere para simular ITDs será aquel generado por una fuente sonora colineal a los oídos. En esta posición la fuente sonora llegará primero al oído izquierdo y tiempo después al oído derecho. Si la diferencia en tiempo esta dada por una separación espacial de 17cm, la misma será de medio milisegundo. Dicha diferencia de tiempo con una frecuencia de muestreo de 44.1kHz, implica cerca de 22 muestras. Por otro lado, si la fuente sonora

se encuentra en frente del escucha, la diferencia en tiempo será nula. No obstante, habría solo 22 posiciones discretas para una fuente sonora virtual con posiciones generadas a partir de retrasos. Si se quiere un conjunto mas amplio de posiciones virtuales posibles, entonces se requeriría una mayor frecuencia de muestreo o un método alternativo que permita la reproducción simulada de retrasos mas cortos que una muestra.

Dadas las complicaciones en cuanto a la capacidad computacional, se implementaron 3 métodos de interpolación. Se utilizó interpolación lineal, cuadrática y cúbica para definir cada muestra en relación al conjunto correspondiente de muestras junto a esta. Por ejemplo, en la interpolación cuadrática se usaron 3 muestras para definir cada una de las muestras interpoladas ubicadas a una fracción de muestra de su propia muestra base. A través del análisis matemático se definió un sistema de ecuaciones donde cada muestra n' se definió a través de una ecuación cuadrática con coeficientes definidos en relación con las amplitudes de su propio conjunto de muestras relacionadas.

$$y(n') = \frac{\gamma - 2\beta + \alpha}{3} (\phi)^2 + \frac{5\beta - 4\alpha - \gamma}{3} (\phi) + \alpha \quad (11)$$

Donde α , β y γ son las amplitudes respectivas de las muestras n , $n+1$ y $n+2$ para obtener la n' muestra ubicada a una fracción ϕ de un periodo de muestreo posterior a la posición de muestreo estándar. Con este método es posible definir cualquier retraso requerido, y mas preciso que un retraso de una muestra, con el fin de obtener una simulación ITD mas precisa a través de un retraso L-R correlacionado en el procesador propuesto.

C. R3

El propósito de este procesador será la eco-localización virtual de una fuente sonora en un punto específico en un espacio tridimensional a través de la simulación de ITDs e IIDs discutidos anteriormente. Sin embargo, su relevancia reside en la parametrización de la realidad de manera tal que las diferencias de nivel y tiempo sean calculadas a partir de la definición de una única posición en un espacio tridimensional. De acuerdo a esto, suponga una posición $P(x,y,z)$ para ubicar una fuente puntual en relación con un escucha ubicado en el origen en R3 de un sistema cartesiano. Se asume que la posición de los oídos izquierdo y derecho son $(0,-r,0)$ y $(0,r,0)$ respectivamente. Para evaluar las distancias (d_l y d_r) desde la fuente sonora al oído izquierdo y derecho se definen las siguientes funciones respectivamente:

$$d_l = P_{(x,y,z)} O_{(0,-r,0)} = \sqrt{x^2 + (y+r)^2 + z^2} \quad (12)$$

$$d_r = P_{(x,y,z)} O_{(0,r,0)} = \sqrt{x^2 + (y-r)^2 + z^2} \quad (13)$$

A través de estas distancias es posible calcular el tiempo de arribo escogiendo una velocidad constante para la propagación del sonido de 340 m/s

$$v_s = d/t = 340 \text{ m/s} \quad (14)$$

$$t = d/v_s \quad (15)$$

$$t_l = \frac{\sqrt{x^2 + (y+r)^2 + z^2}}{340 \text{ m/s}} \quad (16)$$

$$t_r = \frac{\sqrt{x^2 + (y-r)^2 + z^2}}{340 \text{ m/s}} \quad (17)$$

Esto determinará el tiempo de retraso a cada oído. Adicionalmente, esta distancia determinará el nivel de la señal después de recorrer dicha distancia. Asumiendo un patrón de dispersión esférico, es posible calcular, a través de la intensidad acústica inicial, la intensidad en un punto p a una distancia d . Esto se logra dividiendo la intensidad inicial por el área de la esfera de radio d . Dado que la intensidad acústica puede expresarse como sigue, una simplificación de toda la relación sería la siguiente.

$$I_p = \frac{I_o}{4\pi r^2} \quad (18)$$

$$\frac{p_p^2}{z} = \frac{p_o^2}{4\pi z r^2} \quad (19)$$

$$G = \frac{p_p}{p_o} = \frac{1}{2r\sqrt{\pi}} \quad (20)$$

En donde G es la ganancia entre los puntos P y O , I_p e I_o son las intensidades acústicas en el punto de interés y en el origen del sonido respectivamente, p_p y p_o son las presiones acústicas en el punto de interés y en el origen respectivamente, z es la impedancia acústica y r es el radio de la esfera generada por el patrón de dispersión asumido dentro del fenómeno.

Finalmente, sustituyendo el radio por la distancia obtenida anteriormente, es posible calcular las ganancias relativas de la fuente sonora a cada oído

$$G_l = \frac{1}{2\sqrt{\pi(x^2+(y+r)^2+z^2)}} \quad (21)$$

$$G_r = \frac{1}{2\sqrt{\pi(x^2+(y-r)^2+z^2)}} \quad (22)$$

Así pues, mezclado el resultado para el cálculo de las distancias y el cálculo de las ganancias obtenemos lo siguiente, asumiendo una señal de entrada $x(t)$ y una de salida $y(t)$:

$$y(t) = G \cdot x(t - t_d) \quad (23)$$

$$y_l(t) = \frac{x_l\left(t - \frac{\sqrt{x^2+(y+r)^2+z^2}}{340 \text{ m/s}}\right)}{2\sqrt{\pi(x^2+(y+r)^2+z^2)}} \quad (24)$$

$$y_r(t) = \frac{x_l\left(t - \frac{\sqrt{x^2+(y-r)^2+z^2}}{340 \text{ m/s}}\right)}{2\sqrt{\pi(x^2+(y-r)^2+z^2)}} \quad (24)$$

Son entonces estas las ecuaciones por medio de las cuales se define la salida del procesador R3.

D. Binaural

El procesador binarual, como se espera, utilizará una librería de HRTFs (Head Related Transfer Functions) para calcular, a través de convoluciones, las afecciones espectrales de reproducción sonora localizada en relación a un escucha único. A través de la librería HRIR (Head Related Impulse Responses) cada combinación de azimuth y elevación con una profundidad constante, lo cual da una posición única, es asignada a un par de respuestas impulso experimentalmente obtenidas de manera tal que la convolución se calcula con dicha respuesta impulso para simular dicha posición a partir de mar-

cadores psicoacústicos, en particular aquellos relacionados con las afecciones espectrales. Esto se representa como sigue asumiendo una señal de entrada $x(t)$ y una de salida $y(t)$

$$y(t) = x(t) * h(t) \quad (26)$$

$$y(t) = \int_0^t x(u)h(u-t)du \quad (27)$$

En donde $h(t)$ es la respuesta impulso y $*$ denota la convolución. No obstante, debido a implicaciones computacionales, la convolución de señales mayores a 128 muestras se calcula comúnmente a través del producto de las transformadas de Fourier tanto de la señal como de la respuesta impulso denotado como sigue:

$$\hat{y}(t) = \mathcal{F}^{-1}[Y(n)] \quad (28)$$

$$y(t) = \mathcal{F}^{-1}[X(n) \cdot H(n)] \quad (29)$$

$$y(t) = x(t) * h(t) \quad (30)$$

donde $\mathcal{F}^{-1}$ denota la transformada inversa de Fourier y $X(n)$, $H(n)$ y $Y(n)$ denota las transformadas de Fourier de $x(t)$, $h(t)$ y $y(t)$ respectivamente. Finalmente una última variable debe ser incluida para representar la profundidad dentro del sistema de manera tal que actúe como una ganancia para todo el procesamiento. Debido a esto, se asume un sistema coordinado esférico donde (ρ, θ, φ) representarán la profundidad, el azimuth y la elevación respectivamente y $x(t)$ y $y(t)$ serán la señal de entrada y de salida respectivamente

$$y_{l(\rho,\theta,\varphi)}(t) = \rho \mathcal{F}^{-1}[X(n) \cdot H_{l(\theta,\varphi)}(n)] \quad (31)$$

$$y_{r(\rho,\theta,\varphi)}(t) = \rho \mathcal{F}^{-1}[X(n) \cdot H_{r(\theta,\varphi)}(n)] \quad (32)$$

E. R3-Binaural

Finalmente el último procesador será la conjunción entre R3 y Binaural y por tanto será, de hecho, la salida de R3 convolucionada con la respuesta impulso obtenida de la librería HRIR usada en Binaural asumiendo una señal de entrada $x(t)$ y una de salida $y(t)$ como sigue:

$$y_{l(x,y,z,\theta,\varphi)}(t) = \frac{x_l\left(t - \frac{\sqrt{x^2+(y+r)^2+z^2}}{340 \frac{m}{s}}\right) * h_{l(\theta,\varphi)}(t)}{2\sqrt{\pi(x^2+(y+r)^2+z^2)}} \quad (34)$$

$$y_{l(x,y,z,\theta,\varphi)}(t) = \frac{x_l\left(t - \frac{\sqrt{x^2+(y-r)^2+z^2}}{340 \frac{m}{s}}\right) * h_{l(\theta,\varphi)}(t)}{2\sqrt{\pi(x^2+(y-r)^2+z^2)}} \quad (35)$$

Sin embargo, esta conjunción no funciona perfectamente debido a que R3 se basa en un sistema coordinado rectangular mientras que Binaural se basa en uno esférico; por tanto una transformación de coordenadas se necesita para mezclar ambas definiciones. Para esto las siguientes equivalencias se proponen

$$x = \rho \sin\theta \cos\varphi \quad (36)$$

$$y = \rho \sin\theta \sin\varphi \quad (37)$$

$$z = \rho \cos\theta \quad (38)$$

Las cuales a través de algunas simplificaciones dan la salida de R3-Binaural como sigue:

$$y_{l(\rho,\theta,\varphi)}(t) = \frac{x_l\left(t - \frac{\sqrt{\rho^2 + r^2 + 2\rho r \sin\theta \sin\varphi}}{340 \frac{m}{s}}\right) * h_{l(\theta,\varphi)}(t)}{2\sqrt{\pi(\rho^2 + r^2 + 2\rho r \sin\theta \sin\varphi)}} \quad (39)$$

$$y_{r(\rho,\theta,\varphi)}(t) = \frac{x_l\left(t - \frac{\sqrt{\rho^2 + r^2 - 2\rho r \sin\theta \sin\varphi}}{340 \frac{m}{s}}\right) * h_{r(\theta,\varphi)}(t)}{2\sqrt{\pi(\rho^2 + r^2 - 2\rho r \sin\theta \sin\varphi)}} \quad (40)$$

Esto da todos los algoritmos necesarios para construir los plug-ins deseados; sin embargo, los algoritmos por sí mismos definen la matemática detrás de los procesos y no cómo podría un computador calcularlos. Por esta razón, es necesario el desarrollo de software para escribir código viable para ejecutar dichos algoritmos.

III. Programación e Implementación

Para la implementación de este algoritmo en forma de plug-in, una librería de código llamada JUCE para C++ se requiere para programar toda la estructura que, más que procesar audio, permitirá la conformación de un plug-in VST o AU. Esta también incluirá la GUI (Graphic User Interface) y todo el código necesario para el manejo de la información dentro del plug-in. Sin embargo, para este propósito, solamente el código correspondiente al procesamiento de audio será expuesto. En especial, la función `processBlock` será llamada periódicamente para entregar buffers de audio al procesador y entregar el buffer procesado cada vez que el host del plug-in lo solicite. De esta manera, ésta será la función principal para analizar para entender el código correspondiente a cada procesador.

A. Paneo

Para el paneo, un código sencillo es implementado; sin embargo, la complejidad residirá en cómo definir las ganancias respectivas. Para hacer esto, una función de paneo se define para determinar la ganancia en relación con un valor de 0 a 1 asignado a la variable `pan` y diferentes comportamientos se definen dependiendo de si la señal calculada es la del canal izquierdo o el derecho

```
float panner (float pan, int chan) {
    if (chan == 0) {
        pan = pow(1-pan, 0.5);
    } else if (chan==1) {
        pan = pow(pan, 0.5);
    }
    return pan;
}
```

Posteriormente un valor para *Pan* (ganancia izquierda) y para *pAn* (ganancia derecha) se definirán a través de la función anterior

```
pan = panner(panPotValue, 0);
pAn = panner(panPotValue, 1);
```

Finalmente, todo el sistema es multiplicado por la respectiva señal para los canales izquierdo y derecho dentro del `processBlock`. Como se ve, los buffers de entrada y salida (`inL`, `inR`) y (`outL` y `outR`) respectivamente, también se definen en este punto a través de pointers.

```
void PannerAudioProcessor::processBlock
(AudioSampleBuffer& buffer) {
    const float* inL = buffer.getReadPointer(0);
```

```
    const float* inR = buffer.getReadPointer(1);
    float* outL = buffer.getWritePointer(0);
    float* outR = buffer.getWritePointer(1);
    int bufferSize = buffer.getNumSamples();
    for (int i = 0; i < bufferSize; i++) {
        outL[i] = Li * pan;
        outR[i] = Ri * pAn;
    }
```

B. Delay

Para la implementación del delay, se requiere un buffer secundario para almacenar la señal retrasada; sin embargo, este segundo buffer se reproduce en bucle a medida que más y más muestras son procesadas. Adicionalmente, las variables globales *dlyReadPos* y *dlyWritePos* son agregadas para almacenar la posición de lectura y escritura respectivamente dentro del buffer del delay. No obstante, estos no son realmente utilizados dentro del *processBlock*, en cambio se utilizan las variables locales `drp` and `dwp` y al final se almacenan sus valores en las variables globales para mantener los valores hasta la próxima llamada al *processBlock*.

```
void FracDlyAudioProcessor::processBlock
(AudioSampleBuffer& buffer) {
    int drp, dwp;
    int bufSize = buffer.getNumSamples();
    int dlySize = dlyBuf.getNumSamples();
    const float* inL = buffer.getReadPointer(0);
    const float* inR = buffer.getReadPointer(1);
    float* outL = buffer.getWritePointer(0);
    float* outR = buffer.getWritePointer(1);
    const float* dlyDataL = dlyBuf.getReadPointer(0);
    const float* dlyDataR = dlyBuf.getReadPointer(1);
    float* dlyWriteL = dlyBuf.getWritePointer(0);
    float* dlyWriteR = dlyBuf.getWritePointer(1);
    drp = dlyReadPos;
    dwp = dlyWritePos;
```

Finalmente el procesamiento real de los buffers es implementado a través del almacenamiento de la señal de entrada en un buffer temporal circular. Nótese como `Li` y `Ri` son variables locales usadas para definir el valor interpolado y por tanto corresponden directamente a la salida del procesador.

```
for (int i = 0; i < bufSize; i++) {
    dlyWriteL[dwp] = inL[i];
    dlyWriteR[dwp] = inR[i];
    float Li, Ri;
```

La función de interpolación usa posteriormente un conjunto de muestras, una parte fraccional del periodo de muestreo y un tipo de interpolación para calcular el valor interpolado de cada muestra a través del algoritmo descrito anteriormente

```
Li = interpol(dlyDataL[drp],
dlyDataL[(drp + 1) % dlyBufSize],
dlyDataL[(drp + 2) % dlyBufSize],
dlyDataL[(drp + 3) % dlyBufSize], frac, inter);
Ri = interpol(dlyDataR[drp],
dlyDataR[(drp + 1) % dlyBufSize],
dlyDataR[(drp + 2) % dlyBufSize],
dlyDataR[(drp + 3) % dlyBufSize], frac, inter);
outL[i] = Li; outR[i] = Ri;
```

Nótese también como `drp` y `dwp` recorren de manera circular el buffer temporal redefiniendo sus índices a 0 en cuanto llegan al límite del buffer.

```
if (++drp >= dlyBufSize){drp = 0;}
if (++dwp >= dlyBufSize){dwp = 0;}
dlyReadPos = drp;
dlyWritePos = dwp;}
```

Es importante notar que el tiempo del delay está realmente definido a través de la distancia relativa inicial en muestras entre los índices drp y dwp.

C. R3

Para R3 el código es básicamente la conjunción entre el delay y el panning explicado anteriormente. Por tanto, la forma explícita y concisa de definir el buffer de salida es muy sencilla

```
outL[i] = Li * gainL; outR[i] = Ri * gainR;
```

Sin embargo, es importante notar que la relevancia de R3 reside en la parametrización de la cantidad de delay y la proporción entre las ganancias de cada canal. Todas estas variables están definidas en otra función llamada en cuanto cualquier parámetro dentro del procesador cambie llamada setParameter. En este fragmento de código en particular, el parámetro modificado es el azimuth azimuParam, no obstante el código se comporta de manera equivalente si cambiase la elevación. Aquí la distancia entre las variables dL y dR se define a través de la definición de distancia extraída de la transformación entre coordenadas cartesianas a esféricas

```
void R3AudioProcessor::setParameter
(int index, float newValue) {
switch (index) {
case azimuParam : azimu = newValue;
dL = sqrt(pow(depth, 2) -
2 * r * depth * sin(eleva * M_PI / 180) *
cos (azimu * M_PI / 180) +
pow(r, 2));
dR = sqrt(pow(depth, 2) +
2 * r * depth * sin(eleva * M_PI / 180) *
cos (azimu * M_PI / 180) +
pow(r, 2));
```

Posteriormente esta definición será utilizada para calcular las ganancias de acuerdo a los algoritmos anteriormente desarrollados

```
gainL = 1/(2 * sqrt(M_PI) * dL);
gainR = 1/(2 * sqrt(M_PI) * dR);
```

Y finalmente, la longitud de los retrasos son también definidas a través de los mismos algoritmos. Note como la variable global c se utiliza como la velocidad de propagación del sonido (340m/s)

```
dlyL = dL/c;
dlyR = dR/c;
```

Con estas distancias el verdadero retraso es utilizado definido como una diferencia entre muestras y una parte fraccional definida como sigue:

```
dlyReadPosL = (dlyWritePosL - (int)(dlyL
* getSampleRate()) + dlyBufSize) % dlyBufSize;
dlyReadPosR = (dlyWritePosR - (int)(dlyR
* getSampleRate()) + dlyBufSize) % dlyBufSize;
```

```
fracL = fmod(dlyL * getSampleRate(), 1);
fracR = fmod(dlyR * getSampleRate(), 1);
```

Nótese como esto define todos los valores requeridos para las funciones de interpolación y por tanto los retrasos así como las ganancias respectivas de cada canal.

D. Binaural

Para el Binaural se debe crear un buffer circular temporal para almacenar el exceso de muestras y agregar las mismas al próximo processBlock como sigue:

```
for (int i = 0; i < bufSize; i++) {
outL[i] = (outLp[i] + proRdL[i]) / depth;
outR[i] = (outRp[i] + proRdR[i]) / depth;
proWtL[i] = outLp[i + bufSize];
proWtR[i] = outRp[i + bufSize];}
```

Nótese como la primera parte de la función for define la salida inmediata pero la segunda parte escribe el remanente en otro buffer llamado proBuf cuyo pointer de escritura es proWt y cuyo pointer de lectura es usado en la primera parte como proRd. Toda la definición de salida es multiplicada por un factor de profundidad que da la impresión de distancia debido a que los HRIR sólo nos darán la locación esférica a un radio común.

Finalmente el resto del processBlock se muestra a continuación; nótese como las funciones de convolución requieren de las entradas izquierda y derecha (inL, inR), la forma compleja de las respuestas impulso (Transformadas de Fourier de las respuestas impulso) previamente calculadas en la función setParameter, y escribe su salida en los arrays locales outLp y outRp usados posteriormente para definir las salidas del procesador.

```
void BinauralAudioProcessor::processBlock
(AudioSampleBuffer& buffer) {
const float *inL = buffer.getReadPointer(0);
const float *inR = buffer.getReadPointer(1);
float *outL = buffer.getWritePointer(0);
float *outR = buffer.getWritePointer(1);
int bufSize = buffer.getNumSamples();
const float *proRdL = proBuf.getReadPointer(0);
const float *proRdR = proBuf.getReadPointer(1);
float *proWtL = proBuf.getWritePointer(0);
float *proWtR = proBuf.getWritePointer(1);
nFFT = 2 * bufSize;
float *outLp = new float [nFFT];
float *outRp = new float [nFFT];
convolution (inL, irLc, outLp, nFFT, nFil, nSig);
convolution (inR, irRc, outRp, nFFT, nFil, nSig);
```

La función de convolución es implementada a través del producto de las Transformadas de Fourier de la entrada y la respuesta impulso tras haberla rellenado con ceros para lograr una longitud equivalente al tamaño del buffer. Todas las Transformadas de Fourier fueron hechas con una librería C++ de LIBROW para procesamiento relacionado con transformadas directas como inversas de Fourier

E. R3-Binaural

Finalmente el plug-in R3-Binaural viene de una conjunción muy simple de todo el código anterior. Por tanto, como fue mencionado anteriormente, la salida será la salida de Binaural con las ganancias y los delays generados por R3


```

for (int i = 0; i < bufSize; i++) {
    outLp[i] = (outLp[i] + proRdL[i]);
    outRp[i] = (outRp[i] + proRdR[i]);
    proWtL[i] = outLp[i + bufSize];
    proWtR[i] = outRp[i + bufSize]; }
for (int i = 0; i < bufSize; i++) {
    dly2L[dwpL] = outLp[i];
    dly2R[dwpL] = outRp[i];
    outL[i] = dlyL[drpL] * gainL;
    outR[i] = dlyR[drpR] * gainR;}}

```

Nótese como las definiciones de outLp y outRp usan el mismo buffer para redefinirlo por lo que no se requieren más buffers adicionales entre la salida de Binaural y R3

IV. Resultados

En general, la intención es la de desarrollar herramientas comprensivas para que un ingeniero de mezcla pueda sacar provecho de la imagen estéreo y de los procesos de espacialización multicanal durante el proceso de mezcla. Por consiguiente la mejor manera de mostrar los resultados es por medio del resultado individual de cada procesador.

A. Paneo

Aunque un paneo simple es sólo una correlación de ganancias de tal manera que el nivel no se afecte de manera general, entender las relaciones apropiadas entre las ganancias de los canales izquierdo y derecho es crucial. Por esto, la cuantificación de la ley del paneo es un proceso fundamental a la hora de desarrollar algoritmos que utilicen un paneo simple. Así, este procesador tendrá como fin ultimo, ser un plano base para el resto de los plug-ins.

B. Delay

A pesar de que un retraso simple no es un proceso complejo en el procesamiento digital de señales, surgen varias complicaciones a partir del manejo de la información en buffers. Adicionalmente, aunque usualmente no se requiere la interpolación en un retraso simple dado que por lo general la precisión de una muestra es suficiente, es importante tener en cuenta que dicho desarrollo se hizo específicamente para generar retrasos de mayor exactitud que pudieran ser usados como herramientas de espacialización.

C. R3

Aunque los ITDs y los IIDs son usados regularmente durante procesos de mezcla, usualmente no existe una correlación entre los mismos y una posición particular en un espacio tridimensional. Esto será la función principal de este procesador, presentar los ITDs y los IIDs de tal manera que el usuario no requiera decidir que diferencia de tiempo y nivel aplicar, sino solo en donde quisiera escuchar la fuente sonora. Sin embargo como R3 solo utiliza ITDs y IIDs, las posibilidades son limitadas debido a que hace falta información de otros marcadores psicoacústicos.

D. Binaural

Este procesador es un algoritmo completamente diferente de los anteriores, debido a que depende por completo de datos experimentales. Sin embargo esta información será particularmente importante para simular la influencia espectral

en la percepción de la reproducción localizada de la fuente sonora. Esta es una herramienta particularmente engañosa dado que esta pensada para reproducción holofónica por lo que a pesar de ser muy precisa en cuanto a la ecolocalización, puede ser algo perjudicial para el espectro debido al uso de filtros agresivos por la convolución con HRIRs. Por esto, debe ser usada cuidadosamente durante el proceso de mezcla.

E. R3-Binaural

Este procesador incluye todos los otros algoritmos por lo que es la herramienta mas completa para la espacialización; sin embargo, la herramienta mas completa no siempre es la mas apropiada para determinadas tareas. Esta es la razón por la cual se tienen todos los subcomponentes disponibles. En ultimas R3-Binaural combina todo el desarrollo analítico de R3 en relación a los ITDs y IIDs, así como los logros experimentales del Binaural a través de sus HRTFs.

V. Conclusiones

En ultimas, queda una suite de herramientas enteramente dedicadas a la espacialización además de un plug-in de audio completamente dedicado a la espacialización multicanal increíbles posibilidades creativas en el ámbito de la percepción espacial. Todo esto se da a través de la implementación analítica de definiciones básicas de IIDs y la posibilidad de incluir ITDs y eventualmente quizás la posibilidad de incluir afectaciones espectrales a través de HRTFs. Con suerte, estas implementaciones podrían mejorarse en el futuro y quizás este tipo de procesamiento pueda impactar directamente en las prácticas de mezcla comunes para mejorar una parte de la mezcla que no ha sido tan desarrollada por ingenieros de mezcla actuales.

VI. Referencias

- [1] Ballou, G. (1997). Handbook for Sound Engineers - The New AudioCyclopedia (Second Edition ed.).
- [2] Beranek, L. L. (1996a). Acoustics.
- [3] Carrión, A. (1998). Diseño acústico de espacios arquitectónicos.
- [4] William Briggs, L. C., Bernard Gillett. (2013). Calculus for Scientists and Engineers.
- [5] William E. Boyce, R. C. D. (2000). Ecuaciones diferenciales y problemas con valores en la frontera.).



AVANCES DEL AUDIO EN LATINOAMÉRICA.



Audio Engineering Society
Sección Colombia



MÉTODOS ALTERNATIVOS DE ESPACIALIZACIÓN COMPUESTA [PARTE 2]

Juan David Sierra

Juan David Sierra

Facultad de Artes, Pontificia Universidad Javeriana

Bogotá, Colombia

jdsierral@icloud.com

Resumen—El presente proyecto procura hacer un análisis de diferentes algoritmos de espacialización para desarrollar múltiples métodos alternativos de generar posiciones aparentes de fenómenos sonoros. Sin embargo, más allá de todo el desarrollo investigativo, este proyecto representa un esfuerzo por lograr una conjunción clara entre diferentes ramas del conocimiento como la matemática, la física, el análisis de señales y sistemas y los lenguajes de programación en pro de un objetivo artístico. En últimas el lector encontrará la descripción detallada detrás del desarrollo de una suite de algoritmos en forma de plug-ins que sirven a un solo propósito. Generar una explosión de posibilidades en mezcla a la hora de la manipulación de la imagen surround.

Palabras clave: Audio 3D, Surround, Espacialización, Plug-ins, Software.

Abstract—The present project intends to do an analysis of different spatialization algorithms in order to develop multiple alternative methods to simulate apparent positions of sound phenomena. Nevertheless, beyond the research development, this project represents an effort to clearly merge different types of knowledge such as mathematics, physics, signal and system analysis, programming languages, etc... pointing towards an artistic objective. At last the reader will encounter a detailed description behind the development of a suite of algorithms presented as plugins for a single purpose: They pursue a burst of possibilities in mixing engineer regarding the manipulation of surround imaging.

Keywords— 3D Audio, Surround, Spatialization, Plug-ins, Software.

I. Introducción

En general el consumo actual de la música se realiza de una manera despreocupada y poco concienzuda y en el mejor de los casos, si el escucha tiene plena conciencia de la música y el arte que de ella se deriva, es poco frecuente que tenga una preocupación equiparable o si quiera comparable por obtener un buen sonido dentro de su experiencia auditiva. Ello se evidencia en las claras falencias en el posicionamiento de sistemas de reproducción, en la falta de cuidado con la percepción espacial dentro de un registro sonoro, en la escasa atención que se le presta a la calidad de los componentes y equipos de reproducción en general y en últimas, a la falta de importancia y credibilidad sobre los sistemas de reproducción que ofrecen nuevas y mejores tecnologías para enriquecer el proceso auditivo como lo son los formatos multicanal. De hecho, si bien la reproducción en estéreo se ha ganado una gran parte de los audio-escuchas, existen nuevas tendencias comerciales y culturales que han llevado a un incremento en el uso de sistemas monofónicos por parte de los usuarios promedio. Por ejemplo, sistemas portátiles que electrónicamente funcionan en estéreo pero la posición y funcionamiento acústico de los parlantes elimina totalmente los desarrollos en espacialización.

Debido a esto, es importante incentivar el desarrollo de nuevas y mejores tecnologías que acerquen a los consumidores de fonogramas a mejores y novedosas experiencias auditivas y a la vez procurar la educación del público general para que logre valorar dichos procesos. De esta manera buscar que tengan un acercamiento mucho más consciente a aquello que es el sonido y el papel que juega dentro de la música y

en últimas dentro del arte.

Adicionalmente, los procesos técnicos asociados a la producción sonora y al procesamiento de señales en pro del arte han tenido grandes desarrollos en diversas ramas; los ecualizadores son cada vez mejores y de mayor variedad, los compresores son cada vez más transparentes y con mayor diversidad colorística, y así mismo procesos utilizados en la espacialización han sufrido desarrollos inimaginables como sucede con las reverberaciones tanto algorítmicas como de convolución. Sin embargo, el paneo es bastante primitivo aún para un sistema surround. Su implementación a través de tres paneos estéreos en conjunción con un nivel independiente para el canal central muestra que su primitividad es innegable si se compara con el conocimiento disponible en relación con la asertividad de la percepción espacial del hombre.

En otras palabras, el paneo funciona, pero es evidente que la diferencia de nivel en la reproducción es, simplemente, insuficiente para hacer que el cerebro piense que una fuente sonora esta, en efecto, ubicada en una posición específica. Diferentes marcadores psicoacústicos y la precisa conjunción entre estos son esenciales para una aproximación realística. Sin embargo, la rareza no es lo único peculiar de estos procesadores; más importante aún, es el hecho de que estos procesadores no están siendo demandados por ingenieros de mezcla en general. De hecho, esto representa una oportunidad para investigar una necesidad técnica aún desconocida por las personas y entre tanto enriquecer las experiencias auditivas de los escuchas para producir mejores obras de arte.

II. Metodología

Este algoritmo pretende ser un paneo surround altamente personalizable que no permita el control sobre una distribución de niveles a cada parlante en un sistema multicanal como los paneos tradicionales, sino que permita controlar la posición aparente de una fuente sonora en un sistema de reproducción. Para simular estos marcadores psicoacústicos, se debe recurrir a técnicas de procesamiento de señales, y para que ello ocurra, será necesario el desarrollo de algoritmos que puedan determinar todas las variables necesarias para el procesamiento a partir de un número selecto de variables. Esto eventualmente dará vida a una ecuación que será codificada en un lenguaje de programación como C++ para construir un VST/AU plug-in.

A. Diseño Algorítmico

Los procesos de espacialización han sido en general muy diferentes y dependientes de los sistemas de reproducción utilizados; las ideas del vBAP no suelen funcionar tan bien en otros sistemas más sencillos como el surround o el Binaural difícilmente se traduce con claridad al estéreo. Por esta razón es importante desarrollar las mezclas con diversas herramientas de espacialización, y en ese sentido, todos los algoritmos presentados en la parte 1 pueden aplicarse en gran medida a una mezcla y a un sistema surround. Sin embargo, surgen importantes complicaciones incluso desde el paneo más sencillo. SurPan, busca plantear un método diferente de lograr el paneo en surround para permitir mayor control sobre el funcionamiento de los IID dentro de este sistema.

Usualmente, un paneo surround viene de algoritmos en donde se conjugan tres paneos tradicionales. Uno frontal, uno

trasero y uno frontal trasero se usan en conjunto para definir el nivel de cada parlante dentro del sistema (adicionalmente se conjuga también un nivel para el fader central y un fader para el canal LFE). Sin embargo, una correlación inversa no es posible en un sistema así. Esto significa que empezando con estos cuatro valores (los tres paneos tradicionales y el valor del fader del canal central) es bastante difícil entender en dónde, con exactitud, será percibida la fuente sonora.

Para resolver esto, una aproximación radicalmente diferente se requiere. Este algoritmo se basa en la posición que pretende simular. Así, para definir los parámetros de espacialización se utiliza un campo escalar para mostrar la relación entre el nivel de un parlante en el sistema en relación con la fuente sonora virtual que se pretende espacializar. Evidentemente esta relación implica que si la fuente sonora se encuentra exactamente encima del parlante, la ganancia del mismo debería ser unitaria (0dB) y en cuanto se empieza a alejar de la posición del parlante la ganancia debería decaer suavemente de manera concéntrica en relación con el parlante mismo.

Esto es de alguna forma similar a lo que ocurre en una distribución normal o gaussiana en donde el centro de un numero de muestras tiene amplitud máxima y a medida que el centro se aleja, al amplitud decae de manera similar a una curva exponencial cuadrática como lo muestra la siguiente figura. (William Briggs, 2013).

$$y = e^{-x^2}$$
 (1)

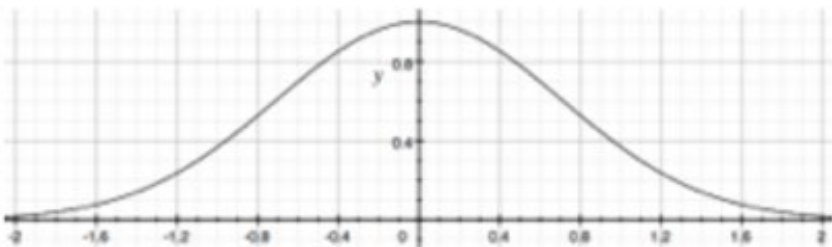


Figura 1. Distribución Normal

Sin embargo, para este caso particular, la expansión en R2 de una distribución es requerida. Esto generará un campo escalar z definido en relación a posiciones xy y dentro de un sistema cartesiano; esto significa que cada dupla x y y definen un valor de altura z así:

$$z = f(x,y) = b^{-x^2-y^2}$$
 (2)

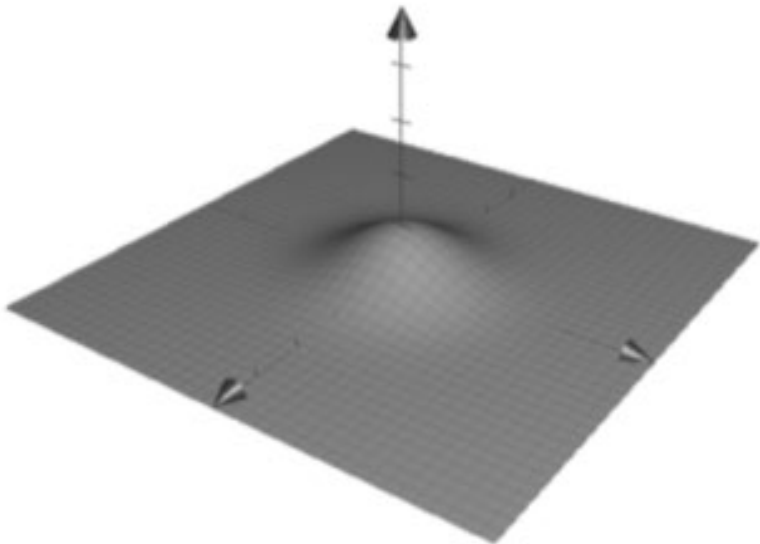


Figura 2. Campo Escalar 1

Dadas las posiciones reales de un sistema surround, múltiples funciones como esta podrían ser desplazadas a las posiciones de los parlantes con el objetivo de obtener control sobre los parlantes en estas posiciones.

Debido a que estas posiciones corresponden a puntos sobre una misma circunferencia pero a diferentes ángulos, estos desplazamientos podrían ser definidos en relación con las posiciones en coordenadas polares para hacerlas dependientes de los ángulos mismos.

$$z = b^{-(x-l)^2-(y-w)^2}$$
 (3)

Para un sistema surround, los parlantes se distribuyen como sigue: (Ballou, 1997): De manera correspondiente, las posiciones están definidas en coordenadas polares como sigue:

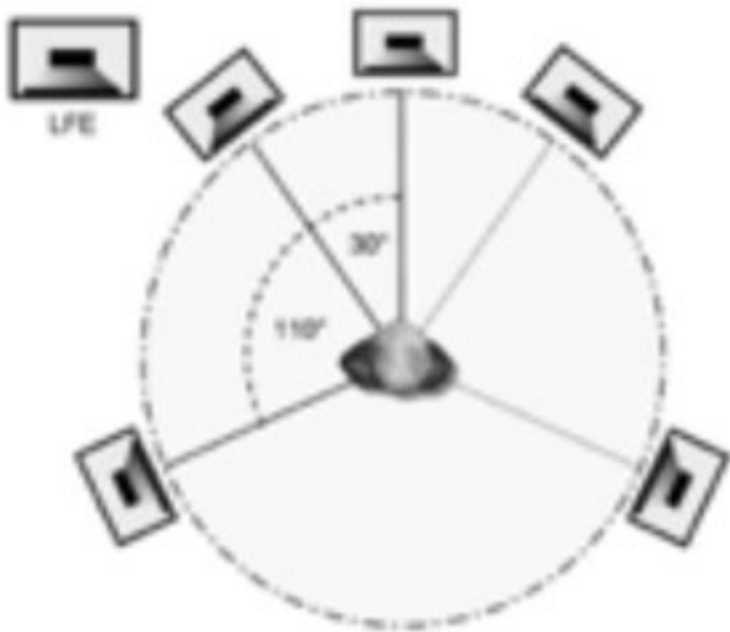


Figura 3. Parlantes en un sistema Surround

TABLA I
ÁNGULOS EN UN SISTEMA SURROUND

Parlante	Angulo	x	y
L	60	dcos(60)	dsin(60)
R	90	dcos(120)	dsin(120)
C	120	dcos(90)	dsin(90)
LFE	—	—	—
Ls	210	dcos(110)	dsin(110)
Rs	340	dcos(340)	dsin(340)

Y por lo tanto los desplazamientos reales de las funciones estarían definidos como sigue:

$$z = b^{-(x-dcos\theta)^2-(y-dsin\theta)^2}$$
 (4)

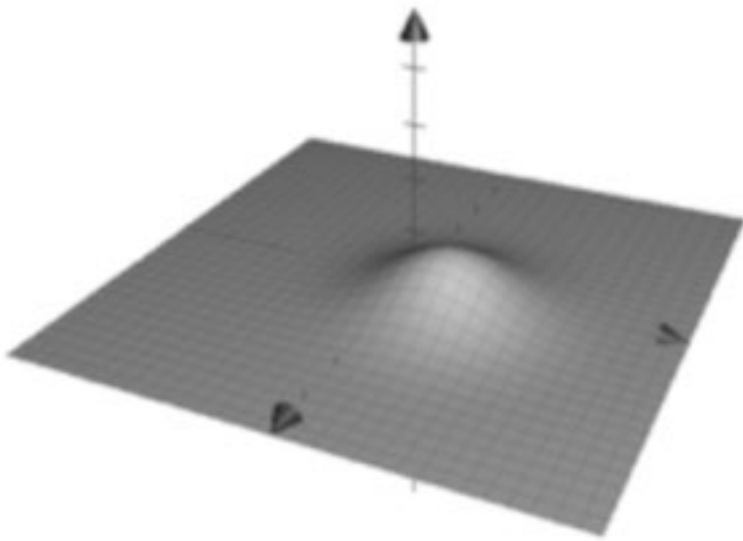


Figura 4. Campo Escalar 2

Debido a la alta proximidad entre los parlantes frontales en un sistema surround, una ponderación horizontal (o si se quiere tangencial) de las pendientes obtenidas de cada función sería deseable. Esto permitiría controlar el exceso de nivel al generarse un movimiento en el eje xy la falta del mismo al generarse un movimiento en el eje y. Adicionalmente, una ganancia general sería deseable en caso de que el nivel general termine siendo demasiado o muy poco. Esto se introduce como sigue:

$$z = a \cdot b \frac{(x-d\cos\theta)^2}{r_x^2} - \frac{(y-d\sin\theta)^2}{r_y^2} \quad (5)$$

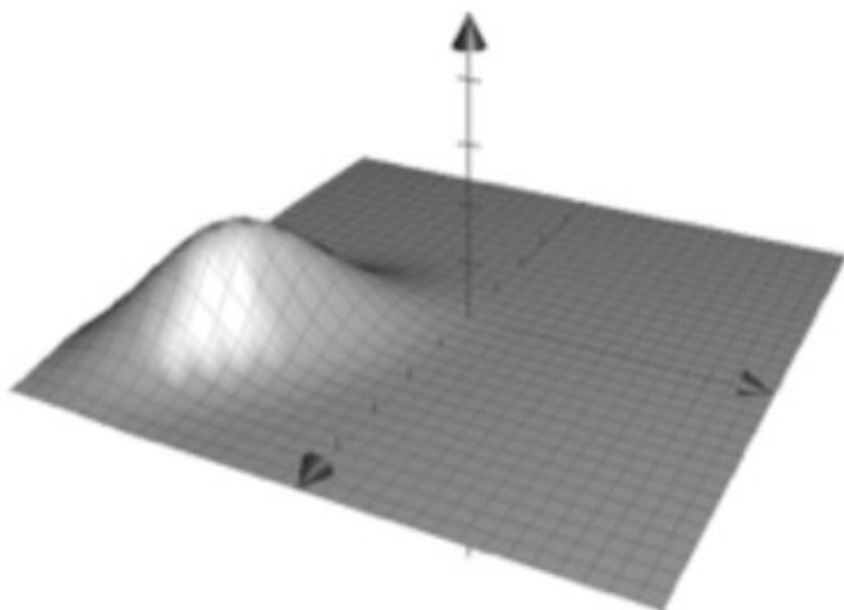


Figura 5. Campo Escalar 3

No obstante, ahora es obvio que lo que inicialmente eran círculos como curvas de nivel de las funciones, ahora son elipses que deberían ser rotadas de manera que apunten de alguna manera al centro del círculo. Esto se logra generando una rotación de cada parlante en relación con el ángulo que define su posición dentro del círculo con la siguiente técnica de rotación axial:

$$x = x' \cos \gamma + y' \sin \gamma \quad (6)$$

$$y = y' \cos \gamma + x' \sin \gamma \quad (7)$$

Finalmente la función rotada y transformada se ve así:

$$z = a \cdot b \frac{(-(x' \cos \gamma + y' \sin \gamma) - d \cos \theta)^2}{r_x^2} - \frac{(-(y' \cos \gamma + x' \sin \gamma) - d \sin \theta)^2}{r_y^2} \quad (8)$$

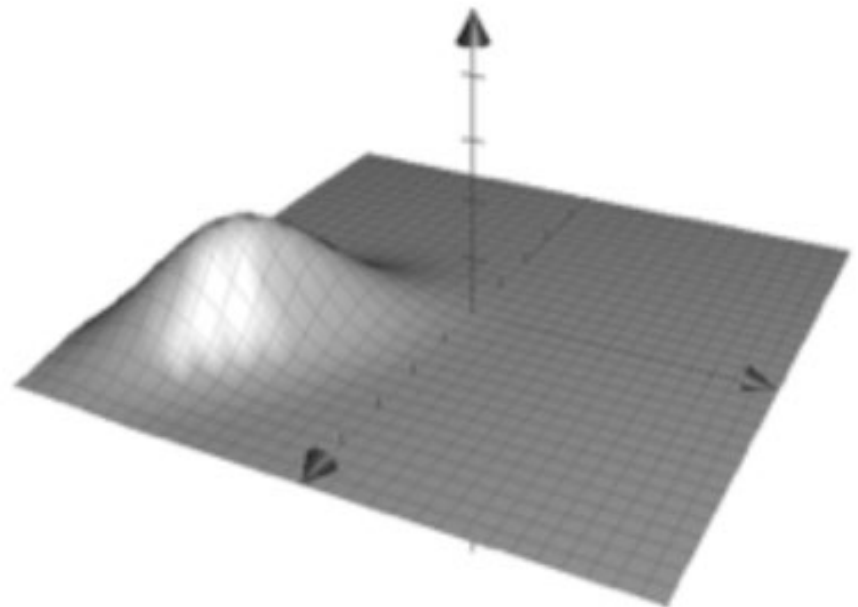


Figura 6. Campo Escalar 4

Finalmente esta función corresponde a la ganancia aplicada a cada uno de los parlantes si alguien está tratando de generar una fuente sonora aparente en un único punto (x,y). Aún cuando la ganancia de cada parlante debe ser calculada de manera independiente es importante entender las relaciones entre todos los parlantes y por lo tanto el conjunto completo de ganancias por canal para un sistema surround sería como sigue:

$$z = a \cdot [\frac{(-(x'_1 \cos \gamma_1 + y'_1 \sin \gamma_1) - d \cos \theta_1)^2}{r_{x1}^2} - \frac{(-(y'_1 \cos \gamma_1 + x'_1 \sin \gamma_1) - d \sin \theta_1)^2}{r_{y1}^2} + \frac{(-(x'_2 \cos \gamma_2 + y'_2 \sin \gamma_2) - d \cos \theta_2)^2}{r_{x2}^2} - \frac{(-(y'_2 \cos \gamma_2 + x'_2 \sin \gamma_2) - d \sin \theta_2)^2}{r_{y2}^2} + \frac{(-(x'_3 \cos \gamma_3 + y'_3 \sin \gamma_3) - d \cos \theta_3)^2}{r_{x3}^2} - \frac{(-(y'_3 \cos \gamma_3 + x'_3 \sin \gamma_3) - d \sin \theta_3)^2}{r_{y3}^2} + \frac{(-(x'_4 \cos \gamma_4 + y'_4 \sin \gamma_4) - d \cos \theta_4)^2}{r_{x4}^2} - \frac{(-(y'_4 \cos \gamma_4 + x'_4 \sin \gamma_4) - d \sin \theta_4)^2}{r_{y4}^2} + \frac{(-(x'_5 \cos \gamma_5 + y'_5 \sin \gamma_5) - d \cos \theta_5)^2}{r_{x5}^2} - \frac{(-(y'_5 \cos \gamma_5 + x'_5 \sin \gamma_5) - d \sin \theta_5)^2}{r_{y5}^2}]$$

Basado en esto, una correlación entre las variables obtenidas y los parámetros físicos sería posible. Gracias a la misma, es evidente que a corresponde a una ganancia general; sin embargo, otras variables no son tan evidentes. b determina las pendientes de las funciones y por lo tanto determina la distancia efectiva para cada parlante. Debido a esto, a medida que el valor de b disminuye (siendo siempre mayor a 1) sería cada vez más fácil para un parlante tener señal en común con el resto de los parlantes haciendo de este parámetro algo similar al factor de divergencia encontrado en paneos surround tradicionales. Por otro lado, d representa la separación entre los parlantes y por lo tanto representa la separación entre las señales mismas. Esto hace que d, de alguna manera, sea una función inversa a lo que la divergencia representa pero de una manera más general.

Lo que este algoritmo permite, eventualmente, es el posicionamiento de una fuente en surround de una manera precisa en relación con la percepción auditiva. Sin embargo, su verdadera fortaleza viendo del hecho de que es completamen-

te independiente de un sistema surround como tal. Así, es relativamente sencillo trasladar la idea de este algoritmo a cualquier otra configuración de parlantes (inclusive configuraciones no co-planares de parlantes).

Finalmente es importante agregar un factor de proximidad para incrementar la ganancia a medida que el sonido se acerque al centro del círculo que es en donde en últimas se generará la percepción. Esto se hace incluyendo una última función al centro del círculo pero con mucha menor pendiente con el objetivo de modificar el sistema completo.

B. Programación e Implementación

Para la implementación de este algoritmo en forma de plug-in, una librería de código para C++ se requiere para programar toda la estructura que, más que procesar audio, permitirá la conformación de un plug-in VST o AU. Esta también incluirá la GUI (Graphic User Interface) y todo el código necesario para el manejo de la información dentro del plug-in. Sin embargo, para este propósito, solamente el código correspondiente al procesamiento de audio será expuesto. Sin embargo, es importante notar que dicha librería será JUCE debido a que permitirá un procedimiento de seteo sencillo e incluirá todas las clases necesarias para este propósito. En especial, la función `processBlock` será llamada periódicamente para entregar buffers de audio al procesador y entregar el buffer procesado cada vez que el host del plug-in lo solicite. De esta manera, esta será la función principal para analizar para entender el código correspondiente a cada procesador.

Debido al hecho de que el procesador `surPan` sólo toma en cuenta las ganancias relativas entre los diferentes parlantes de un sistema surround, el `processBlock` se ve reducido simplemente a modificar las amplitudes de las señales entrantes; así, sólo se requiere calcular dichos valores de ganancia para cada parlante cada vez que `x` o `y` cambian y por lo tanto estos valores no serán calculados para cada muestra sino, en un sentido general, cada vez que un parámetro cambia. Esto da la siguiente versión simplificada del `processBlock`

```
void SurpanAudioProcessor::processBlock
(AudioSampleBuffer& buffer,
MidiBuffer& midiMessages)
{
    const float* in = buffer.getReadPointer(0);
    float* outL = buffer.getWritePointer(0); float* outR
    = buffer.getWritePointer(1); float* outC
    = buffer.getWritePointer(2); float* outLfe = bu
    ffer.getWritePointer(3); float* outLs =
    buffer.getWritePointer(4); float* outRs =
    buffer.getWritePointer(5); int bufSize =
    buffer.getNumSamples();
    float input;
```

En este último, los arrays requeridos son declarados; no obstante, en esta situación en particular, sólo existe una entrada monofónica y una salida surround. Debido a esto se requiere de una variable de entrada tipo `float` que almacene temporalmente cada muestra proveniente del buffer entregado por el host. Esto es particularmente importante porque si alguna de las salidas se define con base al buffer de entrada directamente, esta definición cambiará prematuramente la entrada debido a que dichos buffers estarán funcionando a través de pointers y por lo tanto la primera definición de la salida del primer canal destruiría toda la información necesaria para calcular el resto de las salidas.

```
for (int i = 0; i < bufSize; i++)
{
    input = in[i];

    outL[i] =
    input * gainL * gain * proximity ;
    outR[i] =
    input * gainR * gain * proximity ;
    outC[i] =
    input * gainC * gain * proximity ;
    outLs[i] =
    input * gainLs * gain * proximity ;
    outRs[i] =
    input * gainRs * gain * proximity ;
    outLfe[i] =
    input * gainLfe * gain;
} }
```

Finalmente cada una de los valores de las salidas se calculan a través del producto de la entrada con la ganancia general, la ganancia particular de cada canal y el factor de proximidad. Sin embargo, la parte compleja resulta de calcular las ganancias de cada canal a través del campo escalar propuesto anteriormente durante el diseño algorítmico. Estas ganancias discretas serán calculadas con el siguiente código.

```
void SurpanAudioProcessor::setParameter
(int index, float newValue)
{
    switch (index)
    {
        case bypassParam : bypass = newValue; break;
        case gainParam    : gain = newValue; break;
        case xParam       : x = newValue;
            gainL = gainPerChan(x, y, slope, separ,
            angL_, angL, radT, radN);
            gainR = gainPerChan(x, y, slope, separ,
            angR_, angR, radT, radN);
            gainC = gainPerChan(x, y, slope, separ,
            0, angC, radT, radN);
            gainLs = gainPerChan(x, y, slope, separ,
            angLs_, angLs, radT, radN);
            gainRs = gainPerChan(x, y, slope, separ,
            angRs_, angRs, radT, radN);
            proximity = gainPerChan
            (x, y, 1.1, 0, 0, 0, 25, 25);
```

Acá, evidentemente la función `gainPerChan` es la clave para cada ganancia. Debido a esto, dicha función se muestra a continuación:

```
double gainPerChan
(double x, double y, double slope, double separ,
double ang_, double ang, double radT, double radN)
{
    double gainPerChan =
    (pow(slope,
    (((- pow((x * cos(ang_) - y * sin(ang_)) -
    separ * cos(ang - ang_), 2)) /
    (pow(radT, 2)))
    - pow((x * sin(ang_) + y * cos(ang_)) -
    separ * sin(ang - ang_), 2) /
    (pow(radN, 2)))));
    return gainPerChan; }
```

La anterior corresponde con lo que en el diseño algorítmico era el desarrollo de una distribución normal en R^2 para determinar la ganancia de cada parlante.

En esta función, los parámetros `separ` y `slope` el conjunto específico de ángulos interactúan resultando en un compor-

tamiento singular del parámetro de ganancia para cada parlante.

```
proximity =  
gainPerChan (x, y, 1.1, 0, 0, 0, 25, 25);
```

Finalmente, el factor de proximidad es calculado a través de la misma función dándole valores especiales que le permitan servir de ponderación general del sistema en relación con el centro del círculo de un sistema surround. Adicionalmente, una segunda ganancia general se implementa para permitir que el usuario pueda compensar excesos o escasez de nivel para ciertas configuraciones del sistema.

En general, la manera más simple de entender el proceso establecido por SurPan corresponde a designar a cada parlante una ganancia específica correspondiente a la altura de un punto en R2 debido a la influencia de una función sencilla asignada a cada parlante (esto significa el “montículo” correspondiente a la posición de cada parlante). Lo último debe computarse con la altura obtenida debido a la proximidad al centro del plano cartesiano XY debido a que éste representará la ganancia general calculada automáticamente. Finalmente una última ganancia se aplica dependiendo de las necesidades del usuario.

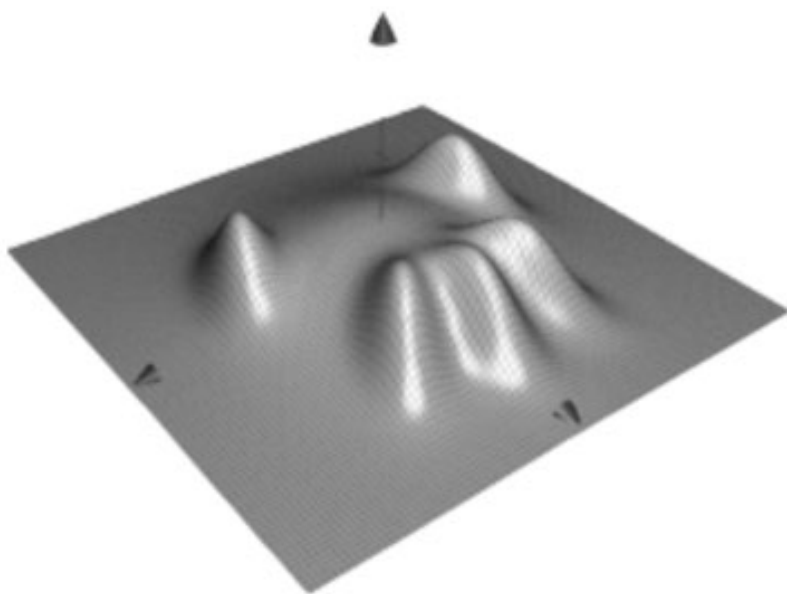


Figura 7. Campo Escalar Representativo de los parlantes en un sistema Surround

III. Resultados

En general, la intención es la de desarrollar herramientas comprensivas para que un ingeniero de mezcla pueda sacar provecho de los procesos de espacialización multicanal durante el proceso de mezcla. SurPan se presenta como un punto de partida para esto debido a que sólo concierne a las diferencias Interaurales de intensidad; sin embargo, tiene suficiente potencial para poder manipular otros marcadores psicoacústicos. Adicionalmente, es extremadamente importante notar que si bien SurPan fue diseñado para sistemas surround, algorítmicamente es perfectamente transferible a sistemas más grandes como VBAP o Ambisonics. Por esta razón, SurPan es especialmente importante en plantear bases para futuros desarrollos en el campo. Finalmente, representa una ruptura del paradigma del paneo en sistemas surround implementado un paneo real en R2 y no una versión estirada de un paneo lineal como ha pasado hasta ahora.

IV. Conclusiones

Al final, queda un plug-in de audio completamente dedicado a la espacialización multicanal con inmensas posibilidades creativas. Todo esto se da a través de la implementación analítica de definiciones básicas de IIDs y la posibilidad de incluir ITDs y eventualmente quizás la posibilidad de incluir afectaciones espectrales a través de HRTFs. Con suerte, estas implementaciones podrían mejorarse en el futuro y quizás este tipo de procesamiento pueda impactar directamente en las prácticas de mezcla comunes para mejorar una parte particular de la mezcla que no ha sido tan desarrollada por los ingenieros de mezcla actuales.

V. Referencias

- [1] Ballou, G. (1997). Handbook for Sound Engineers - The New AudioCyclopedia (Second Edition ed.).
- [2] Beranek, L. L. (1996a). Acoustics.
- [3] Carrión, A. (1998). Diseño acústico de espacios arquitectónicos.
- [4] William Briggs, L. C., Bernard Gillett. (2013). Calculus for Scientists and Engineers.
- [5] William E. Boyce, R. C. D. (2000). Ecuaciones diferenciales y problemas con valores en la frontera.).

AVANCES DEL AUDIO EN LATINOAMÉRICA.



Audio Engineering Society
Sección Colombia



CRITERIA: MEDICIÓN DE RESPUESTA AL IMPULSO Y OBTENCIÓN DE PARÁMETROS ACÚSTICOS

Yonny Quiceno / Alejandro Mejia / Juan F.Londoño / Antonio Escamilla

Yonny Quiceno*, Alejandro Mejia*, Juan F. Londoño*, Antonio Escamilla**

***Facultad de Ingenierías, Universidad San Buenaventura, Medellín, Colombia**

****Escuela de Ingenierías, Universidad Pontificia Bolivariana,**

Medellín, Colombia

yonnyquiceno@gmail.com

alejo_k28@hotmail.com

juanlondono6@gmail.com

antonio.escamilla@upb.edu.co

** Resumen—En este artículo se presenta “Criteria”, un software de código abierto para la medición de respuestas al impulso y el cálculo de parámetros acústicos en salas de concierto y propósito general, soportado en la norma ISO 3382-1:2009 y desarrollado en C++ usando el Framework JUCE. Se exponen los objetivos, la metodología para el desarrollo y las características de la aplicación. Se discute sobre algunas de las consideraciones de diseño mas relevantes y como estas afectan los cálculos de los parámetros acústicos; además se evalúa el desempeño del software comparando los resultados con otros software comerciales.*

Palabras clave— Cálculo de parámetros acústicos, deconvolución, medición de respuesta al impulso, tiempo de reverberación.

Abstract— This article presents “Criteria”, an open source software for the measurement of impulse responses and room acoustic parameters in concert halls and ordinary rooms, based in the ISO 3382-1:2009 standard and programmed in C++ using the JUCE framework. The objectives, methodology for development and software characteristics are exposed. It discusses some of the most relevant design considerations and how these affect the acoustic parameters measurement. In addition, software performance is evaluated by comparing the results with other commercial software.

Keywords— Impulse response measurement, deconvolution, acoustic parameters measurement, reverberation time.

I. Introducción

Todo recinto cerrado tiene asociado un conjunto de parámetros acústicos objetivos medibles, que permiten su caracterización y están estrechamente ligados a la percepción subjetiva de la calidad acústica. El tiempo de reverberación ha sido uno de los parámetros acústicos más relevantes para caracterizar y acondicionar un recinto, sin embargo, existen otro tipo de características dentro una sala que también permiten analizar las cualidades acústicas del recinto, como lo son: las relaciones de energía, relaciones de energía lateral, ruido de fondo, entre otros.

La norma ISO 3382-1:2009, especifica diferentes métodos para medir el tiempo de reverberación en recintos. Es un estándar que aplica para auditorios, salas de conciertos, recintos para la palabra y la música. Describe también el procedimiento de medición, los equipos necesarios y el método para evaluar los datos. Además, se da información de la obtención de los parámetros acústicos del recinto que son derivados de la respuesta al impulso [1].

* Y. Quiceno, A. Mejia y J. Londoño, egresados de la Universidad de San Buenaventura, Medellín. Carrera 56C N° 51-110.

**A. Escamilla, Profesor Asociado en la Universidad Pontificia Bolivariana, Medellín. Circular 1 No. 70-01. (e-mail: antonio.escamilla@upb.edu.co)

La respuesta al impulso, de forma general, hace referencia a la reacción de un sistema dinámico en función del tiempo, en respuesta a una excitación externa. Matemáticamente, su importancia reside en que un impulso, modelado como una función Delta de Dirac, contiene partes iguales de todas las posibles frecuencias de excitación. Razón por la cual la respuesta al impulso define el comportamiento de un sistema lineal e invariante en el tiempo en función de la frecuencia, caracterizándolo por completo. En el caso de una sala, a partir de la respuesta al impulso, se puede obtener la curva de decaimiento energético y con ella obtener los parámetros acústicos que caracterizan el lugar, como lo son: la sonoridad, la claridad, el tiempo de reverberación, entre otros [1].

El desarrollo del software de medición de parámetros acústicos en salas, parte de un análisis de diferentes funcionalidades del software comercial disponible y del estado del arte en esta área [2] [3]. La investigación y el estudio de referentes teóricos, permitieron identificar las diferentes etapas de procesamiento que componen el software según lo sugieren los estándares ISO 18233:2006 e ISO 3382-1:2009, para la obtención de la respuesta al impulso [4] y para el cálculo de parámetros acústicos, respectivamente. Y posteriormente, desarrollar de forma modular, algoritmos para la generación de la señal de excitación, la obtención de la respuesta al impulso, el filtrado por bandas de octava y tercio de octava, el cálculo de la curva de decaimiento y el cálculo de parámetros acústicos. En este documento, se detalla la metodología para el desarrollo, implementación e integración de módulos y algoritmos; además se muestran resultados y conclusiones que evidencian un desempeño muy similar a otros software de amplio renombre comercial como Dirac® y Aurora®.

II. Desarrollo Metodológico

Los objetivos de diseño de una herramienta de software para el calculo de parámetros acústicos, nacen de un análisis del contexto local de la industria acústica y de la evaluación de las características de varios de los sistemas de medición de respuestas al impulso que han sido desarrollados con el mismo fin. El primer objetivo es desarrollar una herramienta de código abierto, multiplataforma, que pueda ser fuente de investigación en áreas como el procesamiento digital de señales y la acústica arquitectónica. Adicionalmente, que pueda desarrollarse de forma modular para expandir su funcionalidad de forma sencilla y generar código reutilizable en proyectos que quieran seguir esta misma filosofía.

Un segundo objetivo es diseñar un software que cumpla con las normativas internacionales ISO 18233:2006, ISO 3382-1:2009 e IEC 61260-1:2014 [5], de forma que los resultados que se obtengan puedan ser equiparables con software estándar en la industria y ser usados de forma confiable por consultores acústicos, investigadores y estudiantes. Por último, como tercer objetivo, diseñar una herramienta con pocas restricciones de hardware, para facilitar el uso de interfaces de audio de propósito general, y con una interfaz de usuario intuitiva, fácil de usar y gráficamente simple y agradable.

A. Generalidades

El software está diseñado para ser multiplataforma y se codificó usando el lenguaje de programación C++ y el framework JUCE, el cual es un amplio, extenso y potente conjunto de herramientas para la creación de aplicaciones de audio, publicado bajo Licencia Pública GNU [6]. Las clases que confor-

man la librería comprenden modelos para la integración de procesos que van desde la construcción de interfaces gráficas hasta la detección de hardware para aplicaciones de audio.

Para el cálculo de parámetros acústicos, la aplicación permite importar una respuesta al impulso previamente grabada o realizar una medición de respuesta al impulso a partir de configurar en primera instancia el hardware de captura y reproducción de audio y luego los parámetros de la señal de prueba, como se muestra en la Fig. 1. En cualquier momento, el usuario puede escoger entre filtros con resolución por octava o tercio de octava, para análisis en función de la frecuencia y calcular los parámetros temporales: DT, T_{10} , T_{20} , T_{30} y los parámetros energéticos: C_{50} , C_{80} , D_{50} , ST_{early} , ST_{late} , T_s y SNR, según como están definidos en la norma ISO 3382-1:2009, ver Fig. 2. Además, se pueden exportar los datos de los parámetros acústicos calculados como archivo de texto para facilitar su análisis por fuera de la aplicación.

A continuación se desglosa de forma detalla cada uno de los módulos que componen el sistema de medición, donde se exponen las consideraciones de diseño mas relevantes en cada etapa.

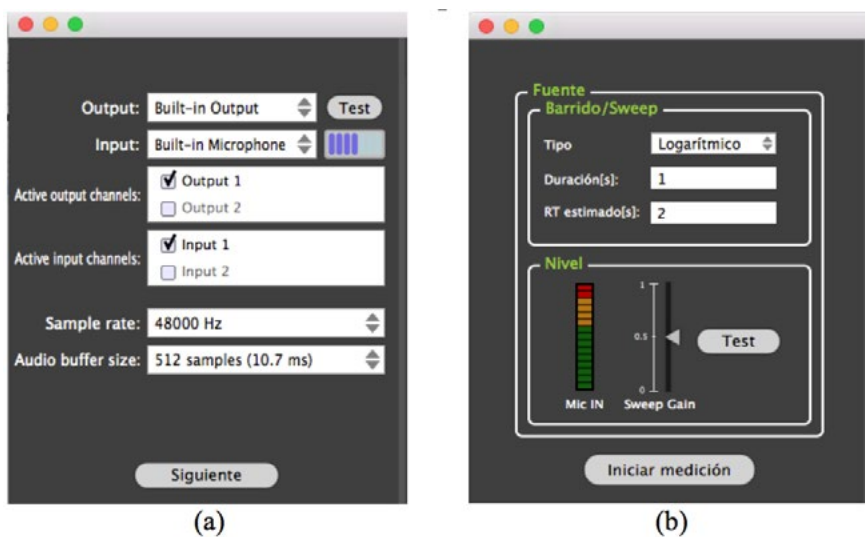


Fig. 1. Interfaz gráfica. (a) Configuración del dispositivo de audio, (b) Parámetros de la señal de prueba.

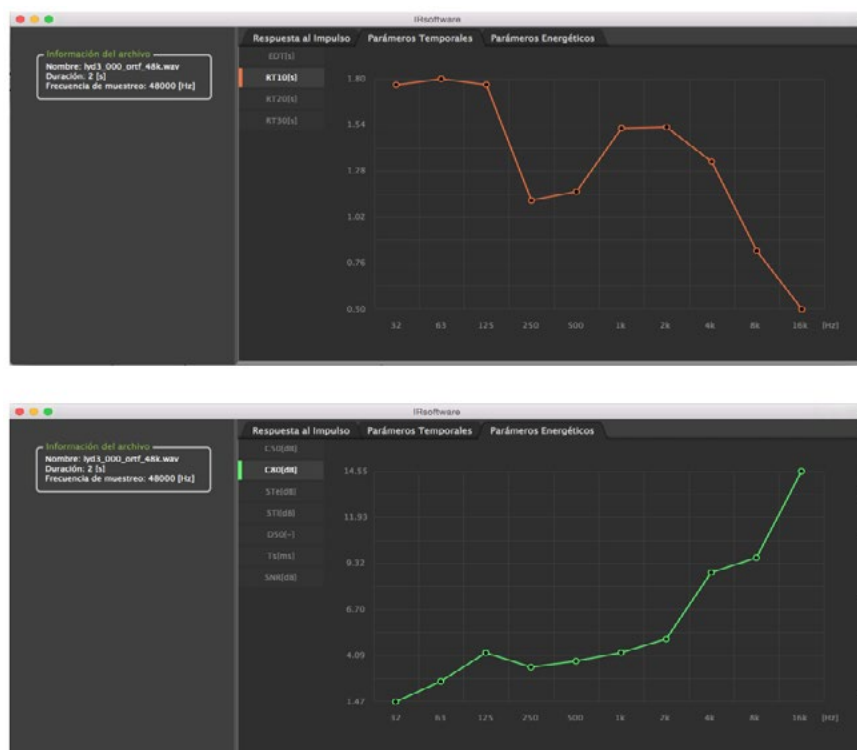


Fig. 2. Interfaz gráfica mostrando los resultados del cálculo de parámetros acústicos. (a) T_{10} , (b) C_{80} .

B. Medición con señales de barrido sinusoidal

Los sistemas modernos de medición de respuestas al impulso en recintos acústicos están basados en usar, como entrada del sistema, una señal de prueba conocida; y con esta, medir la respuesta de la sala [7] [8] [9]. Las condiciones para la elección de la señal de prueba, según se menciona en [2], deben ser: La perfecta reproducibilidad de la señal, la capacidad de maximizar la relación señal a ruido en la respuesta al impulso y permitir la eliminación de artefactos no-lineales que puedan aparecer en la respuesta al impulso. Tradicionalmente, se han usado 2 tipos de señales: Periódicas pseudoaleatorias con prácticamente las mismas propiedades estocásticas del ruido blanco, conocidas como señales MLS y/o señales de barrido sinusoidal de frecuencia cambiante en el tiempo (*Sine Sweep*).

La elección de señales MLS, en el contexto de mediciones de respuestas al impulso en recintos acústicos, ha mostrado que incluso bajo condiciones óptimas (suponiendo invarianza en el tiempo, ruido de fondo bajo, y promediando varias mediciones sincronizadas) no es posible obtener una respuesta al impulso con buen rango dinámico [3]. Autores como Farina [10] coinciden en que en mediciones que usan ruido como señal de excitación, la distorsión producida principalmente por el altavoz se dispersa por toda la respuesta al impulso y que a pesar de que la distorsión se puede reducir con menor nivel en la señal, esto produce más ruido de fondo que contamina el resultado.

Por otro lado, la técnica desarrollada por [10] de usar señales Sine Sweep para medir respuestas al impulso, permite obtener la respuesta al impulso lineal del sistema y separar la respuesta asociada a la distorsión armónica. El uso de señales de barrido sinusoidal permite excluir la distorsión usando señales de longitud mayor a la respuesta al impulso del recinto, quedando como única limitación para obtener una buena relación señal-ruido, el ruido de fondo en el lugar [3].

Como resultado del anterior análisis, se decidió usar señales Sine Sweep para la excitación del recinto. La interfaz de usuario permite escoger entre señales sinusoidales de barrido lineal o logarítmico, definir su duración e ingresar un tiempo de reverberación estimado en el recinto para optimizar el tiempo de captura de la respuesta del sistema. El barrido en frecuencia, que debe abarcar todo el rango audible (20Hz-20KHz), se extiende en ambos extremos para minimizar la influencia de los transitorios introducidos, tanto al comienzo como al final, por la función matemática para el cálculo de la señal sinusoidal.

C. Obtención de la respuesta al impulso

Para la obtención de la respuesta al impulso, la norma ISO 18233:2006, describe la metodología de las que son consideradas nuevas técnicas de medición en edificaciones por medio del procesamiento digital de señales. Según el tipo de señal de excitación, MLS o Sine Sweep, la técnica de procesamiento que relaciona la señal de excitación y la respuesta de la sala, difiere. Estos métodos son alternativos a los expuestos en la norma ISO 3382-1:2009, y plantean procesos indirectos para la obtención de la respuesta al impulso por medio de operaciones como la transformada de Hadamard,

en el caso de señales MLS; y la deconvolución, si se usan señales de prueba Sine Sweep [4].

Para el caso del uso de señales de excitación del tipo barrido sinusoidal, la obtención de la respuesta al impulso se obtiene por medio de una deconvolución directa o por medio de una deconvolución por división espectral. Se decidió optar por el método de deconvolución por división espectral para el desarrollo del software, dado que, a diferencia del método de deconvolución directa, implica únicamente el uso de transformadas rápidas de Fourier sobre la salida y la entrada del sistema, operación que se puede realizar con mucha eficiencia por medio del uso de la librería FFTW [11]. Por otra parte, la deconvolución directa requiere el cálculo adicional de un filtro inverso, y el cálculo de su convolución con la señal capturada, lo cual se traduce en mayor tiempo de procesamiento y mayor complejidad de implementación.

A manera explicativa, la Fig. 3 presenta el diagrama de bloques del algoritmo que se implementó para la obtención de la respuesta al impulso. En la técnica de la división espectral, la señal de excitación y la respuesta de la sala, son transformadas al dominio de la frecuencia, por medio de la transformada rápida de Fourier. Luego el espectro complejo de la respuesta de la sala se divide por el espectro complejo de la señal de excitación. Para obtener nuevamente una señal en el dominio del tiempo, se usa la transformada inversa de Fourier y se obtiene la respuesta al impulso, en donde la segunda mitad de la señal, correspondiente a tiempos de arriba negativos puede ser descartada. Como se describe en [3] este método también puede ser usado para remover los efectos de la distorsión armónica en el canal de excitación.

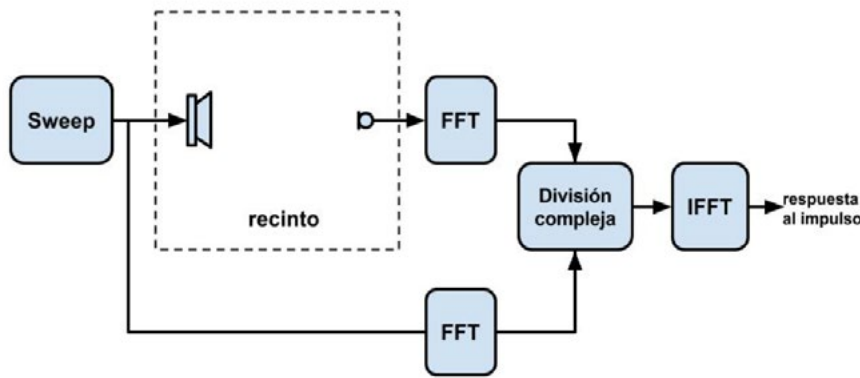


Fig. 3. Diagrama de bloques de deconvolución por división espectral.

D. Filtros por bandas de octava y tercios de octava

Para el diseño e implementación de un banco de filtros pasa-banda por octava y/o tercio de octava, la norma IEC 61260-1:2014 proporciona información en relación a los requisitos específicos de rendimiento para filtros análogos y digitales. De forma concreta, se definen frecuencias centrales, ancho de banda, atenuación del filtro y frecuencias de corte para la implementación de filtros digitales de clase 0, 1 o 2. A partir de las especificaciones dadas por la norma sobre filtros clase 0, se diseña un banco de filtros pasa-banda de respuesta al impulso infinita (IIR) y de orden 10, cuya ecuación en diferencias se muestra en (1).

$$y(n) = \sum_{k=0}^{10} b_k x(n-k) - \sum_{k=1}^{10} a_k y(n-k) \quad (1)$$

Al implementar filtros IIR se obtienen pendientes más pronunciadas y mayor atenuación en comparación al uso de filtros no recursivos (FIR). Adicionalmente, los filtros IIR permiten implementaciones más eficientes, con retrasos más cortos y menos memoria. En el caso particular de un filtro de orden 10, la selectividad de cada banda permite superar las especificaciones de la norma para filtros clase 0, como se muestra en la Fig. 4.

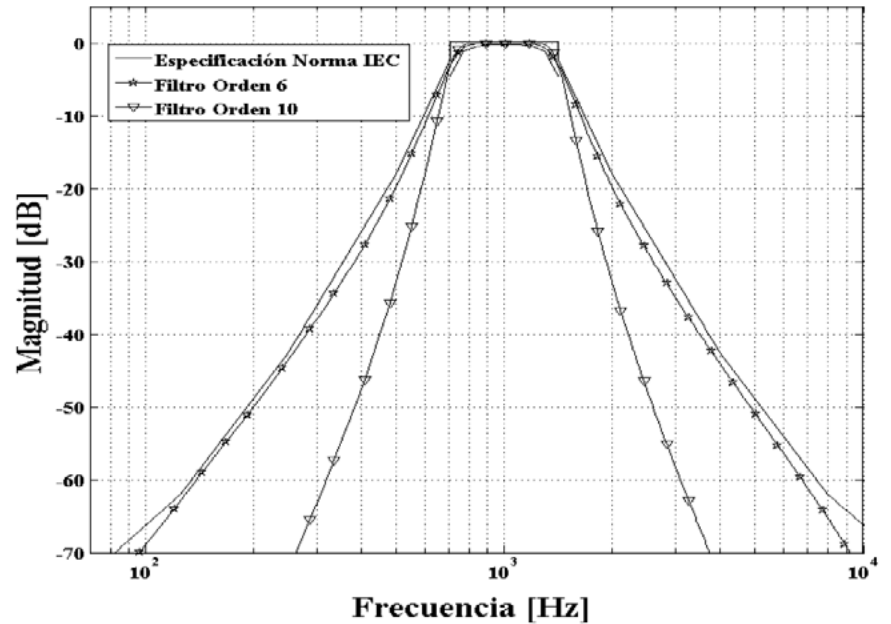


Fig. 4. Atenuación de un filtro de orden 6 y de orden 10 en comparación al límite de atenuación para un filtro de clase 0.

Por medio de la herramienta para el diseño y análisis de filtros de MATLAB, llamada Filter Design & Analysis Tool, se obtienen los coeficientes de la función de transferencia de los filtros, para cada banda de frecuencia. Pensando en la eficiencia de su implementación, se decide reemplazar cada función de transferencia $H(z)$ de orden 10, por 5 secciones de segundo orden en serie, ya que los filtros IIR de orden superior pueden ser muy sensibles a la cuantización de sus coeficientes, y volverse inestables fácilmente. Este hecho es un problema mucho menos frecuente con filtros de segundo orden, por lo tanto, los filtros de orden superior se implementan típicamente de esta manera, según se muestra en (2).

$$H(Z) = \prod_{k=1}^L H_k(Z) = \prod_{k=1}^L \frac{b_{0k} + b_{1k}Z^{-1} + b_{2k}Z^{-2}}{1 + a_{1k}Z^{-1} + a_{2k}Z^{-2}} \quad (2)$$

E. Curva de decaimiento energético

La curva de decaimiento energético es una curva que muestra cómo decae la energía del recinto en función del tiempo. Esta curva es obtenida por medio de la integral inversa de Schroeder, que no es otra cosa que la integración inversa de la respuesta al impulso al cuadrado, como se indica en (3).

$$E(t) = \int_t^{\infty} p^2(\tau) d\tau = \int_{\infty}^t p^2(\tau) d(-\tau) \quad (3)$$

Ante condiciones de ruido de fondo alto durante la medición de la respuesta al impulso, es probable la obtención de pará-

metros acústicos erróneos, que se asocian a inconvenientes en el procedimiento del cálculo de la curva de decaimiento energético. Situación que hace necesaria la utilización de algún método que minimice la influencia del ruido de fondo. Según la norma ISO 3382-1:2009, existen dos métodos que minimizan la influencia del ruido capturado en la respuesta al impulso. Uno de estos métodos sugiere que, cuando el nivel de ruido de fondo es conocido, el límite inicial t_1 para la integral inversa de Schroeder, es la intersección entre una línea horizontal a través del ruido de fondo y una línea que representa la pendiente del decaimiento de la respuesta al impulso al cuadrado en escala logarítmica. La integración inversa se lleva a cabo entonces según (4).

$$E(t) = \int_t^{\infty} p^2(\tau) d\tau = \int_{t_1}^t p^2(\tau) d(-\tau) \quad (4)$$

El procedimiento por medio del cual el software detecta el límite t_1 , se basa en la comparación de la envolvente suavizada de la respuesta al impulso con un valor de ruido de fondo calculado a partir del último 10% de la señal capturada, el cual se asegura que corresponde a ruido de fondo, dados los criterios usados para la optimización del tiempo de captura. En la Fig. 5 se ilustra la manera en que se hace la comparación, donde t_1 correspondería al tiempo en el cual la envolvente suavizada interseca por primera vez al nivel de ruido de fondo.

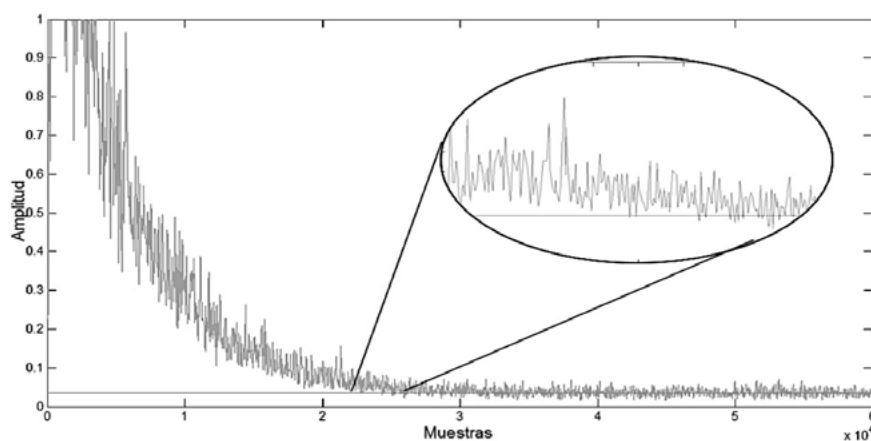


Fig. 5. Intersección de la envolvente suavizada de la respuesta al impulso con nivel de ruido de fondo (línea horizontal).

Habiendo determinado t_1 , se aplica una versión discreta de (4) sobre las respuestas al impulso filtradas, obteniendo así curvas de decaimiento energético para cada banda de frecuencia.

F. Cálculo de parámetros acústicos

Para el cálculo de los parámetros de criterio temporal se desarrolló un algoritmo que, partiendo de la curva de decaimiento energético, realiza una regresión lineal con los datos dentro del rango de evaluación de interés específico para cada parámetro. Por ejemplo, para el cálculo del T_{20} , se obtiene una función lineal ajustada a los datos del decaimiento energético entre -5dB y -25dB, que determina la tasa de decaimiento energético por unidad de tiempo. Posteriormente se debe extrapolar la función lineal correspondiente, hasta el valor donde se pueda evaluar un decaimiento de 60 dB y medir la cantidad de tiempo que finalmente se interpreta como un parámetro acústico (EDT , T_{10} , T_{20} y T_{30}). Es importante dejar claro que el subíndice numérico en los anteriores parámetros, indica explícitamente el rango en decibelios con

el que fue evaluada la tasa de decaimiento energético por unidad de tiempo, con la que luego se evalúa para cada parámetro temporal, un hipotético decaimiento de 60dB [1].

Para el cálculo de los parámetros de criterio energético (C_{50} , C_{80} , D_{50} , ST_{early} , ST_{late} , T_s) se realizó un algoritmo que utiliza las respuestas al impulso previamente filtradas, por bandas de octava o tercio de octava, y calcula las relaciones energéticas según las ecuaciones de la norma ISO 3382-1:2009. En este cálculo, es vital definir con mucha precisión los límites de integración, (según la definición de cada parámetro energético), en relación a la información de la respuesta al impulso comprendida entre la posición donde la envolvente de la señal sobrepasa significativamente el nivel del ruido de fondo y la posición donde la envolvente de la señal decae debajo del nivel del ruido de fondo. Dicho procedimiento permite que las integrales para el cálculo de los parámetros energéticos, se hagan sobre el rango adecuado de datos. Lo que garantiza que sólo se tengan en cuenta las contribuciones energéticas de la respuesta al impulso, y no las del ruido de fondo. Además, permite que la curva de decaimiento energético tenga una compensación por ruido que permite el cálculo más preciso de los parámetros temporales, según lo establecido en la ISO 3382-1:2009.

III. Resultados

Para el análisis de resultados se realizaron mediciones en varios recintos de la universidad de San Buenaventura Medellín, que presentan diferentes características en cuanto a superficies, volumen, nivel de ruido, uso, entre otros. En cada uno de estos recintos se realizaron procesos de medición haciendo uso de los software Critería, Dirac® y Aurora®, teniendo en cuenta que los parámetros de configuración de la medición con cada software fueran los mismos. Además, toda la cadena electroacústica para las mediciones fue la misma, con el fin de garantizar un punto de comparación bien establecido. A continuación se muestran algunos de los resultados más representativos.

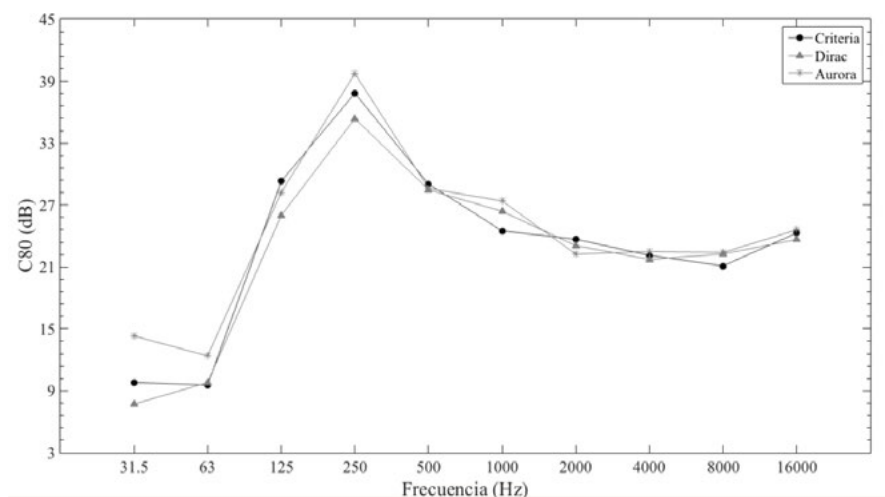


Fig. 6. Comparación parámetro C_{80} por octavas.

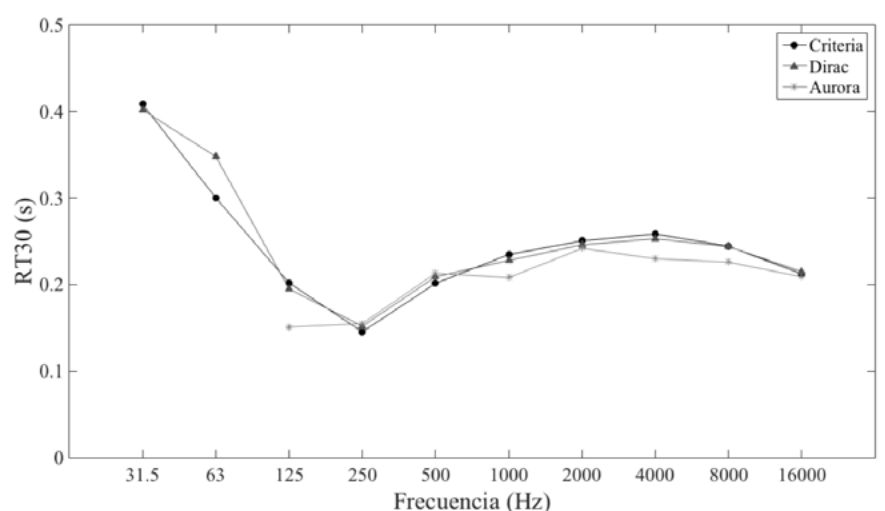


Fig. 7. Comparación parámetro RT_{30} por octavas.

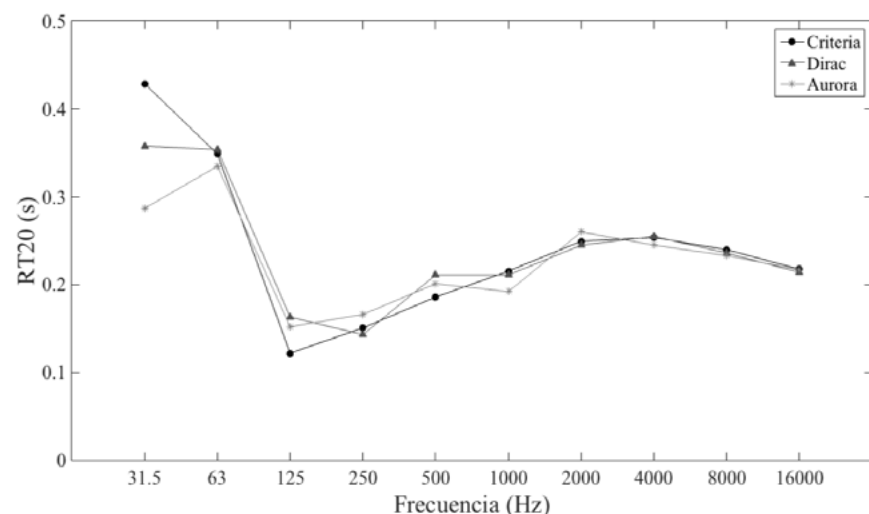


Fig. 8. Comparación parámetro RT20 con mayor ruido de fondo en la medición (aire acondicionado encendido).

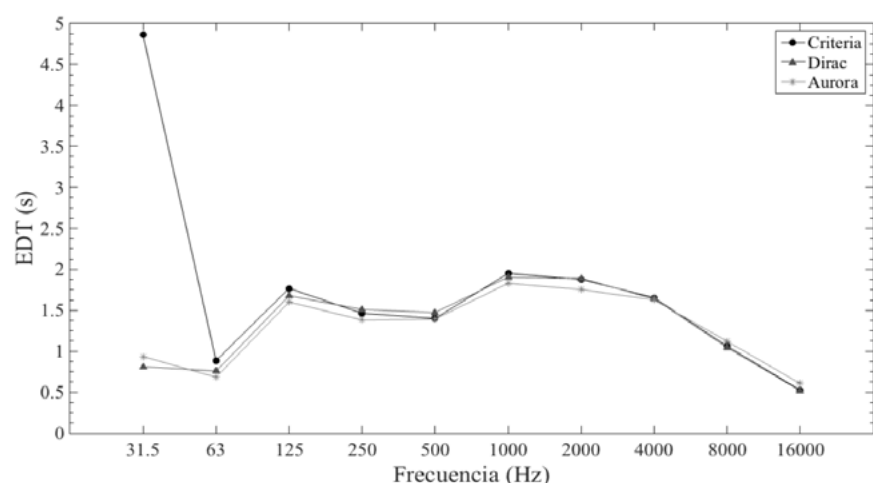


Fig. 9. Comparación para el parámetro EDT por octava.

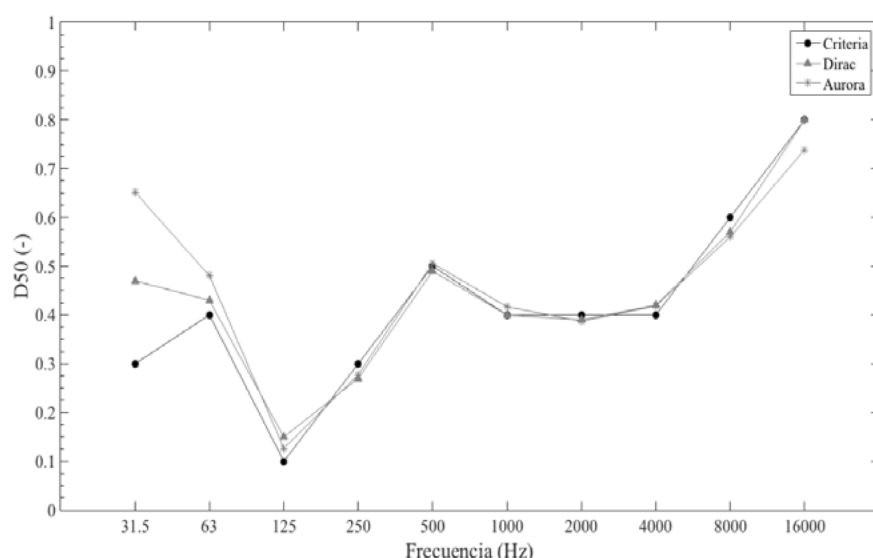


Fig. 10. Comparación parámetro D50 por octava con mayor ruido de fondo en la medición (aire acondicionado encendido).

Al analizar los resultados de las Figs. 6,7 y 9, se observa que los parámetros obtenidos por Critería convergen en conjunto con los obtenidos por Dirac® y Aurora®, presentando tendencias muy similares tanto en los parámetros temporales como en los energéticos. No obstante, se puede ver que en las bandas por debajo de 125 Hz los resultados no siguen una tendencia en particular. Esto podría deberse a que en la normativa no se establece un diseño de filtro específico, ocasionando posibles diferencias en el diseño de estos y es posible que los filtros usados por los diferentes software presenten diferencias en cada una de sus características, entre las cuales está el retraso de fase, el cual afecta en mayor proporción a las respuestas al impulso filtradas correspondientes a las bandas de baja frecuencia.

En los resultados presentados en las Figs. 8 y 10 se muestra que la presencia de ruido de fondo durante la medición, produce mayor dispersión entre los parámetros acústicos obtenidos usando los tres software mencionados anteriormente. Estos resultados son similares a los que se presentan en [12] y [13], donde se comparan resultados entre las bandas de octava de 125 Hz y 4000 Hz bajo diferentes técnicas de compensación de ruido y condiciones acústicas. Allí se evidencia que a pesar de los intentos de calibración que hacen los diferentes software a partir de respuestas al impulso idealizadas, en presencia de condiciones reales en la medición donde el sistema no es lineal e invariante en el tiempo, las disimilitudes en los parámetros acústicos calculados son evidentes.

IV. Conclusiones

Independientemente de la gran cantidad de variables involucradas en los diferentes procesos para la obtención de los parámetros acústicos, se obtienen resultados muy similares a los obtenidos con otros software de renombre comercial; lo cual demuestra que los procesos implementados apoyan su funcionamiento en un seguimiento adecuado de lo establecido en los referentes teóricos. Sin embargo, se presentan pequeñas variaciones en los parámetros acústicos obtenidos, que se deben a diferencias en el procesamiento posterior de la respuesta al impulso y no a su medición.

Otro tipo de pruebas, mostraron de forma preliminar, que procesos como el filtrado por bandas de frecuencia y los diferentes métodos de compensación de ruido que usan los software evaluados, son la principal causa de las diferencias en los parámetros acústicos. Situación que además evidencia que no hay un punto de referencia que se pueda tomar como verdad absoluta en la comparación de parámetros obtenidos con uno u otro software.

V. Referencias

- [1] Standard, ISO, «3382-1:2009. Acoustics - Measurement of Room Acoustic Parameters-part 1: Performance Spaces,» Genève: International Organization for Standardization, 2009.
- [2] G.-B. Stan, J.-J. Embrechts y D. Archambeau, «Comparison of different impulse response measurement techniques,» Journal of the Audio Engineering Society, vol. 50, n° 4, pp. 249-262, 2002.
- [3] S. Muller y P. Massarani, «Transfer-function measurement with sweeps,» Journal of the Audio Engineering Society, vol. 49, n° 6, pp. 443-471, 2001.
- [4] Standard, ISO, «18233: 2006. Acoustics - Application of New Measurement Methods in Building and Room Acoustics,» Genève: International Organization for Standardization, 2006.
- [5] IEC, «61260-1:2014. Electroacoustics-Octave-Band and Fractional-Octave-Band Filters,» International Electrotechnical Commission, 2014.
- [6] ROLI, Juce Framework, 2014.
- [7] B. Boren y A. Roginska, «Multichannel impulse response measurement in matlab,» de Audio Engineering Society Convention 131, 2011.

- [8] Farina, Angelo, Aurora software suite, 1996.
- [9] Bruel & Kjaer, Dirac - Room Acoustics Software - Type 7841.
- [10] A. Farina, «Simultaneous measurement of impulse response and distortion with a swept-sine technique,» de Audio Engineering Society Convention 108, 2000.
- [11] M. Frigo y S. G. Johnson, «FFTW: An adaptive software architecture for the FFT,» de Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on, 1998.
- [12] P. Fausti y A. Farina, «Acoustic measurements in opera houses: comparison between different techniques and equipment,» Journal of Sound and Vibration, vol. 232, p. 213–229, 2000.
- [13] B. Katz, «International round robin on room acoustical impulse response analysis software 2004,» Acoustical Society of America, vol. 5, n° 4, pp. 158-164, 2004.



CONFERENCIA LATINOAMERICANA

Resumen del evento

La conferencia anual AES Latinoamérica es el evento más importante del audio en América Latina, el cual se realizó en Colombia en su edición 2016, y en el que se reunieron conferencistas y panelistas reconocidos, además de fabricantes, expositores, profesionales, aficionados y todos los interesados en el mundo del sonido para ampliar su conocimiento, dar un vistazo a lo último del mercado en audio profesional, y asistir a múltiples actividades paralelas que le permitieron a los asistentes expandir su red de contactos profesionales.

Durante los tres días del evento se presentaron más de 15 conferencistas internacionales y 25 nacionales quienes hicieron parte de las diferentes actividades programadas tales como conferencias, demostraciones, paneles de discusión, ponencias y sesiones de escucha.

La Conferencia Latinoamericana AES 2016 se llevó a cabo del 21 al 23 de julio en las instalaciones de la Universidad de San Buenaventura en la ciudad de Bogotá-Colombia. Conferencias, talleres, paneles de discusión, demostraciones, competencias estudiantiles de diseño y grabación, críticas de grabación, visitas técnicas, documentos y carteles se presentaron en torno a cuatro temas principales durante los tres días.

- Acústica, electroacústica y la psicoacústica
- Refuerzo sonoro y Sonido en vivo
- Producción (grabación, mezcla y masterización)
- Post-producción (Cine, TV, Radio y videojuegos)

Durante la conferencia, se presentaron invitados de renombre internacional tales como:

Vance Powell (EE.UU.)
Mathew Waters (EE.UU.)
Ryan Hewitt (EE.UU.)
Eduardo Vaisman (ARG)
Bruce Olson (EE.UU.)
Robert Génèreux (CAN)
Alex Lameira (BRA)
Daniel Reis (BRA)
Mauricio "Magú" Ramírez (MEX)
Alejandro Bidondo (ARG)
Valeria Palomino (MEX)
Rodrigo Casassus (CHI)

Por otra parte se presentaron más de 25 profesionales locales como Ana Jaramillo, Mauricio Ardila, Ricardo Escallón, César Vilar, de Tony Peñarredonda, Marcela Zorro, Carlos Silva, Andrés Núñez, Camilo Silva, Andrés Landínez y Andrés Millán, entre otros.

Competencias de grabación y diseño

También se realizó la tradicional competencia de grabación y en ésta oportunidad se incluyó una competencia de diseño; la competencia tuvo el apoyo de importantes empresas del sector comercial de audio que contribuyeron con los premios a los ganadores y participantes, entre ellas encontramos a Audio Element con Izotope, AudioTechnica, Sennheiser, Retro Knob, Instituto Ene Audio, Genelec y VCR LTDA.

La siguiente es la lista de ganadores en las diferentes categorías.

RESULTADOS COMPETENCIAS AES 2016	
Competencia de Grabación	
Categoría 1 - Grabación Acústica Tradicional	
Ana Reyes	Primer Puesto
Daniela Varela	Segundo Puesto
Luis Chaparro	Tercer Puesto
Categoría 2 - Grabación Tradicional en Estudio	
Rodolfo Chaves	Primer Puesto
Carlos Socotá	Segundo Puesto
Alfredo Neira	Tercer Puesto
Categoría 3 - Grabación Moderna en Estudio	
Rafael Santamaría	Primer Puesto
Jossman Saac	Segundo Puesto
Jose Antonio Cordero	Tercer Puesto
Categoría 4 - Sonido para Medios Audiovisuales	
Esteban Calderón	Primer Puesto
Andrés Marthe González	Segundo Puesto
Daniel Castilla	Tercer Puesto

Competencia de Diseño	
Categoría 1 - Acústica	
Desierta	
Categoría 2 - Electroacústica	
Daniel Upegui	Primer puesto
Angie Guerrero	Segundo puesto
Nathalia Beltrán	Mención
Juan Sierra	Mención

MOMENTOS

DESTACADOS DEL EVENTO



Ceremonia de apertura a cargo del presidente de la AES Colombia Ing. Andrés Millan, también participaron las directivas académicas de la Universidad de San Buenaventura y como acto principal se le realizó un homenaje al productor e ingeniero Eduardo de Narváez por sus más de 25 años de trayectoria profesional.



La vice presidenta de la AES región Latino América Valeria Palomino entrega la placa en homenaje al productor Eduardo de Narváez.



Presentación del norteamericano Vance Powel: "El estudio analógico en la economía del DAW".



Eduardo Vaisman de Argentina en su presentación sobre "El audio team en video juegos AAA"



La colombiana Ana Jaramillo haciendo la presentación sobre "Programas de simulación acústica como herramienta para el análisis y diseño de espacios".



Mauricio "Magú" Ramírez de México platicando sobre el "Factor de Cresta".



El presidente del comité de estándares de la AES Bruce Olson (USA) haciendo su presentación acerca de el "Diseño de sistemas de sonido y el uso de programas de simulación"



Matt Waters (USA) presentando la conferencia llamada "Contando historias a través del sonido".



Ryan Hewit (USA) presentando la conferencia "Mezclar es un verbo: llevando su mezcla más allá de un simple balance".



Alejandro Bidondo de Argentina en su charla denominada "Midiendo la difusión del campo sonoro – Sound Field Diffusion Coefficient (SFDC)"



Ale Lameira (BRA) presentado "Nuevos desarrollos de guías de onda aplicadas a arreglos lineales".



La vicepresidente regional de la AES para Latinoamérica Valeria Palomino en su presentación sobre "Nuevas tendencias para grabación de ópera y orquestas en vivo"



Panel sobre Acústica en Auditorios: música vs palabra". Panelistas: Alejandro Bidondo (ARG), Andrés Núñez, Ana Jaramillo y Álvaro Tobón (COL), moderado por Lucas Serrano de Colombia.



Panel sobre Mezcla híbrida: consideraciones, flujos de trabajo y buenas prácticas, moderado por Andrés Millán (COL).



Jurado evaluador de la competencia de diseño revisando el proyecto ganador.



Panel de mastering moderado por Carlos Silva



Coctel de apertura de la conferencia realizado en el hall de la biblioteca de la Universidad de San Buenaventura.





COMITÉ AES COLOMBIA

El Comité de la Audio Engineering Society Sección Colombia, desde sus inicios a principios de la década de los 2,000, ha trabajado para generar espacios de intercambio de conocimiento, siempre con fines académicos, y enfocados a estudiantes, profesionales, docentes e investigadores. AES Colombia ha desarrollado importantes actividades dentro del marco Latinoamericano, partiendo de la 4ta Conferencia Latinoamericana del año 2010 en la Pontificia Universidad Javeriana, hasta la Conferencia Latinoamericana AES 2016 en las instalaciones de la Universidad de San Buenaventura Sede Bogotá; sin dejar a un lado la constante presentación de conferencias, talleres, paneles de discusión y seminarios individuales con panelistas nacionales e internacionales.

Ahora bien, con un comité renovado y conformado con profesionales y docentes con experticia en las diferentes

ramas del sonido y afines de miembros activos de Bogotá, Medellín y Bucaramanga, y con un subcomité de la ciudad de Medellín, nuestro país se consolida como una de las regiones más fuertes y activas de Latinoamérica. El comité de AES Colombia proyecta una expansión nacional de las asociación con la presentación de la Conferencia Anual AES Colombia 2017 en la ciudad de Medellín y con actividades en Bogotá, Bucaramanga y a lo largo de diversas ciudades del territorio nacional. Por lo anterior, se demuestra la constancia en actividades de las asociación y la enorme proyección que existe para la expansión del conocimiento a través de diversas actividades académicas y publicaciones como ésta, que mantendrá una periodicidad anual y basada en la presentación de los proyectos de investigación expuestos en las conferencias anuales de AES Colombia.

ING. ANDRÉS FELIPE MILLÁN SILVA
Presidente AES Colombia

PRÓXIMOS **EVENTOS AES**

CONFERENCIA AES COLOMBIA 2017



Audio Engineering Society
Sección Colombia

MEDELLÍN-COLOMBIA 24-26 DE AGOSTO DE 2017



142ND AUDIO ENGINEERING SOCIETY INTERNATIONAL CONVENTION

Convention: May 20-23, 2017 • Exhibition May 20-22, 2017
Maritim Hotel Berlin



EXHIBITS: OCT 18TH-20TH
PROGRAM: OCT 18TH-21ST

AES NEW YORK 2017

143RD INTERNATIONAL CONVENTION



AVANCES DEL AUDIO EN LATINOAMÉRICA.



Audio Engineering Society
Sección Colombia



UNIVERSIDAD DE
SAN BUENAVENTURA
SEDE BOGOTÁ

PATROCINA



AFMG
Ariant Felsel Media Group



Electro-Voice

infoComm
INTERNATIONAL



SENNHEISER



OLSON SOUND DESIGN



Meyer
Sound

APOYA



INSTITUTO ENE AUDIO



Read the difference
diffUSION
MAGAZINE