

METODOLOGÍA PARA LA EVALUACIÓN SUBJETIVA DE LA RECREACIÓN VIRTUAL AUDIOVISUAL DE UN RECINTO COMPARÁNDOLA CON LA REPRODUCCIÓN DE LA EXPERIENCIA CAPTADA IN-SITU

PACS: 43.55.Ka

Fernández Villa, Mario; Navarro, Juan Miguel.
Universidad Católica San Antonio de Murcia.
Avd. de los Jerónimos, s/n
30.107 Guadalupe (Murcia). España
Tel: +34 968 278 825
Fax: +34 968 278 581
E-Mail: mafernandez@alu.ucam.edu, jmnavarro@ucam.edu

ABSTRACT

The virtual recreation of an enclosure seeks to provide the viewer a perception as realistic as possible. This article presents a framework proposal to assess the experience with a virtual recreation of a particular room compared with a binaural audio recording and stereoscopic video of the real enclosure. The various stages of the method include the collection, modeling, simulation and reproduction of both types of captures.

RESUMEN

La recreación virtual de un recinto busca proporcionar al espectador una percepción lo más realista posible. En este artículo se presenta una propuesta de marco de trabajo para valorar la experiencia obtenida con una determinada recreación virtual de imagen y sonido comparándola con una grabación del audio binaural y vídeo estereoscópico del recinto real captada mediante un maniquí que integra ambos captadores. En este artículo se presenta las tareas a realizar en las distintas fases del método propuesto que incluyen la captación, modelado, simulación y reproducción de los dos tipos de capturas.

1. INTRODUCCIÓN

En las últimas décadas han sido numerosos los estudios, artículos y conferencias que se han dedicado a la simulación virtual de recintos ya sea de vídeo 3D de un local o "Paseo virtual" [1], o bien una simulación de audio binaural [2], utilizando para ello varias vías o métodos [3, 4], ya sea a través de ordenadores, altavoces, auriculares, amplificadores y en general todo tipo de dispositivos electrónicos, un ejemplo de simulación de audio binaural sería, por ejemplo, la recreación de una clase impartida por un profesor en un aula determinada.

La simulación acústica de salas comenzó en los sesenta con maquetas a escala pero la simulación de la absorción era compleja además la creación de las maquetas era costosa y los resultados posteriores no eran muy precisos. Más adelante en los noventa surgieron los ordenadores y los programas informáticos que lo hicieron todo más fácil. Actualmente los modelos basados en acústica geométrica alcanzan una gran eficiencia y un importante grado de acercamiento a la realidad [5].

En este trabajo se presenta una propuesta de metodología para la captura y comparación subjetiva de la sensación visual y sonora obtenida mediante una experiencia real y una experiencia sintetizada de un escenario en 3D. Para ello, se compara la captación real de audio y video 3D, con la simulación a través de software y hardware de ese sonido y video 3D en un recinto singular [6]. Como ejemplo de aplicación se ha dispuesto un aula de la universidad en la cual se ha realizado un estudio de medidas y características para que a través de software se genere una recreación virtual 3D, tanto de video [7] como de audio [8]. Por otro lado usando una cabeza de maniquí [9] se captarán señales de audio y video 3D para posteriormente compararlas con la recreación virtual obtenida y así comprobar qué grado de similitud posee nuestra recreación en referencia a la realidad.

Finalmente se procederá a la comparación de los resultados obtenidos en ambos casos de estudio. Para la correcta realización del estudio se someterá a los sujetos seleccionados a la ejecución de unos test de escucha que nos aportarán todos los datos que necesitamos para valorar las similitudes y diferencias entre el video y sonido 3D grabado en el recinto real y el sintetizado mediante software y hardware.

Este documento está dividido en los siguientes apartados. En la sección siguiente se describen las diferentes fases que componen la metodología propuesta. A continuación se presentan unos experimentos realizados usando esta metodología junto con una discusión de sus resultados. Finalmente se enuncian las conclusiones del trabajo.

2. METODOLOGÍA PROPUESTA

En este apartado se describe de forma general la metodología, así como sus diferentes fases. A continuación en la Figura 1 se expone un diagrama de bloques que explica la metodología propuesta en este trabajo. En ella se diferencian una serie de fases que empiezan con la recogida de información y posteriormente se subdividen en 2 ramales, que finalmente convergen en un único ramal cuando se lleva a cabo la comparación de resultados y conclusiones.



Figura 1. Metodología propuesta.

El primer paso y común a ambos ramales es la recogida de información del recinto. A través de esta fase se recaba información geométrica y de las propiedades acústicas de la sala. Una acción útil a realizar es conseguir los planos digitalizados del recinto, de los cuáles se podría obtener la información a través de un visor, como por ejemplo el DWG True View de Autocad.

Las medidas de las diferentes alturas del recinto y de los elementos que en él se albergan se podrán obtener con una cinta métrica, distanciómetro láser o con cualquier otro sistema de medida. Otra técnica útil es la toma de diferentes fotografías para la posterior contribución a una recreación ajustada lo máximo a la realidad.

Una vez concluida la fase inicial nos centraremos en las siguientes, que se subdividen en 2 ramales. El ramal superior describe las fases a realizar para la captación real de video y audio, estas fases son: Preparación del recinto y sistema de captación, captación de video y audio 3D y finalmente la composición 3D real. Por otro lado el ramal inferior nos explica las fases que se desprenden de la recreación virtual del recinto. Dichas fases son las siguientes: Modelado del recinto, simulado del recinto y para terminar este ramal se lleva a cabo la composición del paseo virtual.

2. 1. Recinto Real

Tal y como se puede observar en la rama superior de la Figura 1 la adquisición de datos en el recinto real se descompone en las siguientes fases: Preparación del recinto y del sistema de captación, proceso de captación del vídeo y audio 3D y la preparación del sistema de reproducción. A continuación se describen con más detalle estas fases.

2. 1. 1. Preparación del recinto y del sistema de captación

Una vez que hemos tomado toda la información pertinente se inicia la preparación del recinto, para ello se hará una distribución de los puntos de medida, la colocación y calibración de la fuente sonora, disposición de amplificadores, tarjetas de sonido, cableado y todo tipo de hardware necesario, etc. Una vez que está el recinto preparado, se debe conseguir un sistema de captación apropiado, este sistema puede ser desde una cabeza de maniquí simple, pasando por un dummy head más sofisticado con orejas de silicona hechas de moldes reales

hasta incluso un sujeto real. Independientemente del sistema elegido, éste se convertirá en el portador de los micros y cámaras.

2. 1. 2. Captación de video y audio 3D

En esta fase es donde realmente se toman las medidas y se guardan resultados, para a continuación hacer un primer análisis de concordancia de los resultados obtenidos por si fuese necesario repetir alguna medida.



Figura 2. Maniquí fabricado para captación de video y audio.

El proceso de captación debe ser llevado a cabo mediante una cabeza artificial o una persona al cual se le instala un sistema de 2 cámaras situadas a la altura de los ojos y 2 micrófonos a la altura de los oídos. Tal y como se muestra en la Figura 2, estos irán alojados en el interior de unas orejas de silicona implantadas en la cabeza, y en el caso la persona de unos micros intraoído. El video es grabado en sendas tarjetas de memoria alojadas en las cámaras. Por otro lado el sonido es recogido por los micrófonos y a través de un circuito preamplificador y utilizando una tarjeta de sonido externa es almacenado en una grabadora. [10]

En el proceso de captación se pueden realizar dos tipos de tomas de datos, por un lado en puntos fijos y por otro con el receptor en movimiento:

- a. Oyente en movimiento. Utilizando el sistema de adquisición montado sobre un trípode con ruedas es posible realizar un recorrido pasando por unas zonas definidas en la fase anterior a la vez que se emite a través de la fuente sonora una grabación anecoica.
- b. Oyente fijo. Si suponemos puntos receptores fijos, donde la oyente no se mueve, se pueden adquirir las respuestas impulsionales de ambos oídos y una fotografía estereoscópica desde la localización y orientación del oyente. Además es posible realizar una captación de una grabación anecoica como en el tipo anterior.

2. 1. 3 Composición 3D Real

Para finalizar, la última fase consiste en el arreglo y la composición de todo lo grabado en el proceso anterior, es necesario el uso de software avanzado de edición no lineal para la correcta sincronización del audio y video 3D en un único archivo, para la posterior reproducción y prueba en la fase de comparación de resultados.

2.2. RECREACIÓN VIRTUAL

En este apartado se describen las fases mostradas en la rama inferior del diagrama de bloques de la Figura 1: Modelado del recinto, simulación del mismo y composición del paseo virtual.

2.2.1 Modelado del recinto

Para realizar la modelación 3D de un recinto se han de seguir varios procesos que se describen a continuación, entre ellos se encuentran la obtención de medidas de los recintos, análisis de los materiales de construcción y la creación del modelo con los distintos tipos software [11,12].

En primer lugar y partiendo de los datos recabados del recinto es necesario llevar a cabo el modelado del mismo. En el mercado actual existen multitud de aplicaciones de ordenador para el modelado en 3D, como por ejemplo: Autocad, SketchUp o EASE. Este último está especialmente diseñado para la simulación acústica de recintos, por lo que se puede usar en las siguientes fases de este proceso. A continuación se deben asignar colores o texturas a las superficies objetos del recinto para añadir realismo al modelado. Finalmente se creará una animación del recinto que muestre el mismo recorrido realizado en la captación real.

2.2.2 Simulación del recinto

Una vez finalizado el modelo y la animación de nuestro recinto se pasará a la fase de simulación acústica. Para realizar la parte de simulación siempre será necesario importar el recinto ya modelado, bien con otro software o el mismo de la simulación. Una vez cargado, el programa nos permitirá depurar todo tipo de errores y optimizar el diseño del recinto. Seguidamente, este tipo de software suelen ofrecer herramientas para asignar color y coeficientes de absorción y difusión a los materiales del recinto para así poder determinar en consecuencia su firma acústica en forma de respuesta impusional.

A día de hoy, con los modelos computacionales basados en la acústica geométrica se pueden conseguir condiciones cercanas a la realidad debido al incremento de la velocidad de cálculo, así como el desarrollo de nuevos algoritmos y métodos de análisis que incluyen los fenómenos de reflexión difusa y especular, atenuación del aire, difracción, etc. Los algoritmos más recientes producen respuestas al impulso que permiten la implantación de auralizaciones de alto realismo. Para poder crear el paseo virtual será necesario obtener la respuesta impusional en los mismos puntos de los pares fuente-receptor escogidos en la captación real.

2.2.3 Composición del paseo virtual

A través de la respuesta impulsional resultante del paso anterior tenemos que crear el fragmento de audio binaural para su integración con la animación creada. En la Figura 3 se muestra un diagrama de bloques con las tareas que se deben realizar para obtener unos archivos que se pueden reproducir mediante auriculares. En el caso de reproducción mediante altavoces sería necesario añadir algún algoritmo de cancelación de diafonía al final del proceso.

Las respuestas al impulso monofónicas deben convertirse a binaurales mediante el uso de las funciones de transferencia relacionadas con la cabeza (HRTF). Seguidamente llega la parte más delicada del proceso donde convolucionamos estas respuestas con el sonido anecoico utilizado en la captación real. Conforme el oyente virtual se mueve a una velocidad, igual a la del paseo real, las respuestas al impulso deben variar para adecuarse al par emisor-receptor. Finalmente, el archivo de sonido resultante se compondrá con la animación de forma sincronizada.

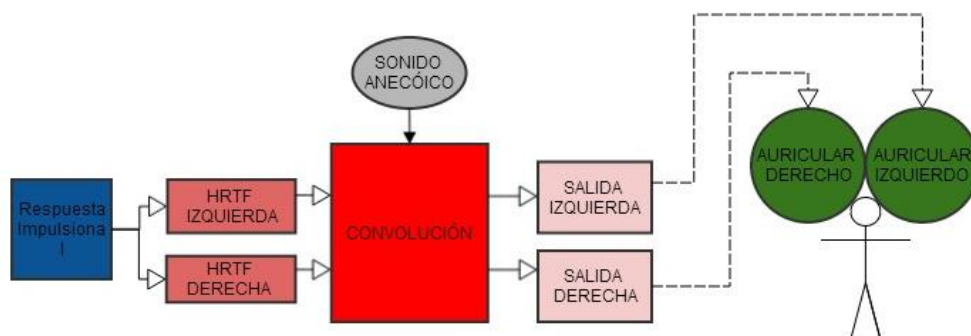


Figura 3. Creación del audio del paseo virtual para su reproducción con auriculares

2.3. Comparación y Evaluación de la Experiencia

Para el proceso de comparación es necesario contar con un sistema de reproducción de vídeo estereoscópico y de sonido binaural. En el caso de vídeo existen varios sistemas, tales como las pantallas de proyección, monitores o cascos de visionado. Para la recreación del audio se puede optar desde auriculares personales, hasta sistemas de altavoces con 2 o más canales.

La fase de evaluación consiste en reproducir diferentes experiencias del receptor en movimiento y de puntos fijos del receptor. Con las capturas realizadas se pueden reproducir de forma ciega las siguientes combinaciones: Vídeo y audio real, vídeo y audio virtual, vídeo real y audio virtual, y vídeo virtual y audio real. Para cada tipo de reproducción se dispondrá de una serie de sujetos para evaluar y comparar las sensaciones transmitidas por cada uno de los sistemas de reproducción. Las evaluaciones consisten en una serie de preguntas comparativas sobre la calidad subjetiva de la experiencia de audio binaural recibida.

3. EXPERIENCIA DE PRUEBA

Para validar este protocolo se presenta una experiencia realizada explicando más detalles y posibles mejoras y ayudas observadas.

3.1. Recinto Real

El recinto utilizado para esta prueba es un aula de la Universidad Católica de San Antonio, en él se distribuyeron 14 puntos situados estratégicamente para posicionar en ellos el receptor estático. La fuente se sitúa todo el tiempo al lado de la pizarra, en el centro, donde teóricamente estaría el profesor. Otra prueba es la que se hace con el sujeto en movimiento, el circuito realizado parte de la puerta del aula, llega al fondo de los pupitres por el primer pasillo y regresa hasta el principio por el último pasillo.

En la Figura 4 se muestra una imagen del video en movimiento grabado del aula en formato anáglifo. Un ejemplo de un paseo captado en aula se puede ver en [14].



Figura 4. Imagen de video 3D capturada en el recinto real.

3.2. Recreación Virtual

Para la recreación virtual hemos utilizado sketchup para el modelado y EASE para el simulado, a continuación se explican ambos procesos más ampliamente.

3.2.1 Modelado con sketchup

Los modelos de SketchUp se componen de dos cosas: aristas y caras. Los bordes son líneas rectas y las caras son las formas 2D que se crean cuando varias aristas forman un plano. Este programa posee como ventaja que permite exportar el modelo en diferentes formatos compatibles con las diferentes herramientas de simulación del mercado. Tal y como se muestra en la Figura 5, en primer lugar es recomendable construir un modelo básico del recinto con un nivel de detalle suficiente para la simulación acústica [13]. A continuación, para añadir más realismo al paseo virtual es necesario añadir más detalles y objetos singulares como mesas y sillas. A la hora de asignar las texturas de cada superficie u objeto se recomienda enlazar cada

capa con un color específico pues esto facilita la asignación del material correspondiente cuando se importe a una herramienta de simulación acústica, ver Figura 6.

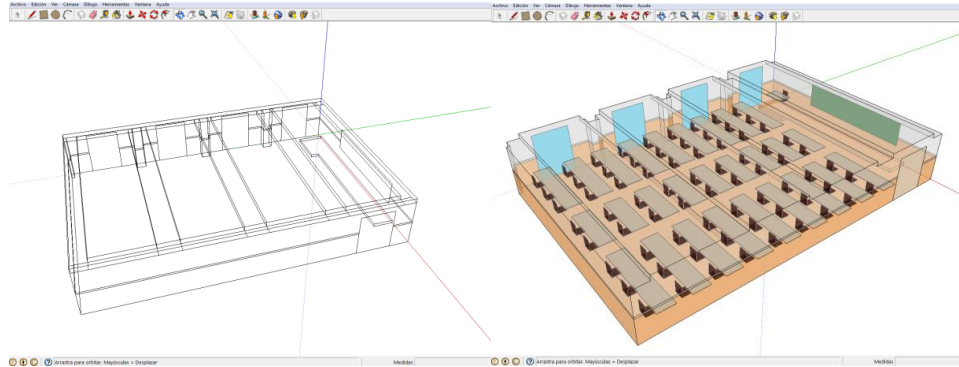


Figura 5. Modelo básico del recinto.

Figura 6. Modelo terminado del recinto.

3.2.2 Simulación con Ease

Una vez finalizado el modelado de nuestro recinto se pasará a la fase de simulación acústica con EASE (Enhanced Acoustic Simulator for Engineers), el cual ofrece un conjunto de herramientas muy valioso para la práctica profesional, desde el modelado de recintos hasta la simulación acústica de los mismos. Para la simulación con EASE importamos el recinto modelado en la fase anterior, una vez cargado, el programa nos permite depurar todo tipo de errores y optimizar el diseño del recinto, el resultado se observa en la Figura 7. Ease es una herramienta compleja y sofisticada que nos proporciona funciones para asignar el material y el color/textura a cada cara y objeto del recinto e incluso definir su orientación. El módulo EASEView es una herramienta que permite mejorar la impresión visual del modelo, dando la posibilidad de crear una representación foto-realista del mismo. Se basa en la incorporación al modelo de texturas fotográficas y fuentes de iluminación que proporcionan un realismo adicional al proyecto. Como se puede observar en la Figura 8 esta interfaz muestra las caras orientadas, las partes no visibles son la parte posterior de las caras, a las cuales no se les han asignado ningún material y textura, ya que no se utilizarán para la simulación.

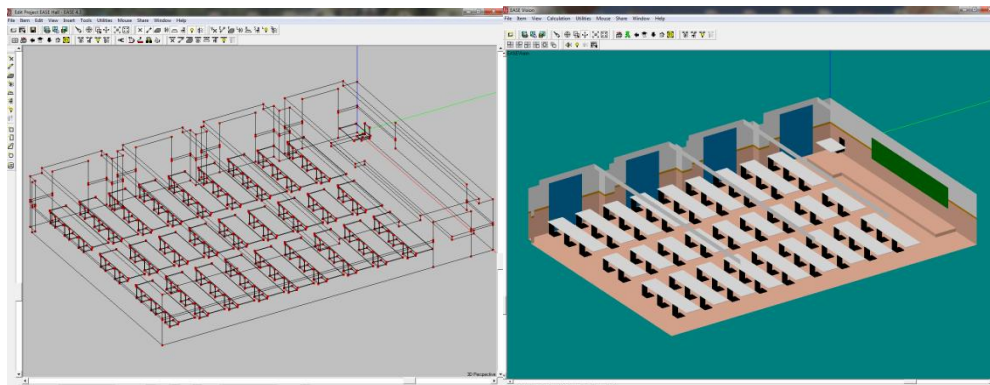


Figura 7. Modelo importado.

Figura 8. Recinto en EASEView.

La animación creada con Sketchup se integra con la auralización extraída con las diferentes respuestas al impulso de cada posición. Un ejemplo de un paseo virtual creado en el aula se puede ver en [14].

3.3. Comparación de resultados

En el proceso de reproducción, tanto para el caso real como el virtual, el video será visionado con un HMD (Head Mounted Display) que es un dispositivo de visualización similar a un casco, que permite reproducir imágenes creadas por ordenador. A este dispositivo se le incorpora unos auriculares de reproducción binaural.

Dentro de cada caso se evaluaron 2 situaciones diferentes, la primera es que el receptor se mantenga fijo en un punto del recinto y así podemos medir la respuesta al impulso desde esa posición. La segunda situación es que el receptor se quede en movimiento continuo, a la misma velocidad, para este caso, se probó con una fuente anecoica que simuló varias situaciones habituales de una clase impartida por un profesor.

4. CONCLUSIONES

En este trabajo se ha presentado una propuesta de metodología general para realizar una evaluación subjetiva de la comparación entre la experiencia provocada por la recreación virtual audiovisual de un evento en un recinto y la reproducción de la experiencia captada in-situ. Para ello se han descrito las fases a realizar en ambos casos mostrando las alternativas que se pueden tomar en las diferentes tareas. También se plantea el proceso de evaluación y comparación entre ambas reproducciones. Finalmente se ha presentado una experiencia de prueba de esta metodología en una aula de clase. Los resultados preliminares de las comparativas muestran la relevancia del vídeo real en la sensación sonora de la auralización. En trabajos futuros se presentarán resultados más detallados sobre las observaciones y conclusiones obtenidas en varias pruebas que se están llevando a cabo.

5. BIBLIOGRAFÍA

[1] Moezzi, S. ; California Univ., San Diego, La Jolla, CA, USA ; Li-Cheng Tai ; Gerard, P., "Virtual view generation for 3D digital video", MultiMedia, IEEE (Volume:4 , Issue: 1) 2002.

[2] Levine, Scott; Leonard, Brett; King, Richard, "Binaural Room Simulation for Acoustic Testing", AES Convention:134 (May 2013), E-Brief Papers—Part 1.

[3] Czyzewski, Andrzej; Kosikowski, Lukasz; Kostek, Bozena , "Virtual Hearing Aid – A Computer Application for Simulating Hearing Aids Performance" , AES Convention: 122. Psychoacoustics, Perception, and Listening Tests. May 2007.

[4] Roginska, Agnieszka; Scott, Frederick, "Room-Dependent Preference of Virtual Surround Sound", AES Convention:124 (May 2008), Spatial Audio Perception and Processing.

- [5] Pellegrini, Renato; Horbach, Ulrich, "Perception-based Design of Virtual Rooms for Sound Reproduction", Subject:Spatial Audio
- [6] S. Hirzel Verlag, "Binaural Room Simulation" Acta Acustica united with Acustica, Volume 72, Number 4, December 1990, pp. 295-296.
- [7] M Schöllmann, M Fulland, H.A Richard, "Development of a new software for adaptive crack growth simulations in 3D structures" Institute of Applied Mechanics (FAM), University of Paderborn, Pohlweg 47-49, D-33098 Paderborn, Germany.
- [8] Noisternig, Markus; Sontacchi, Alois; Musil, Thomas; Holdrich, Robert, "A 3D Ambisonic Based Binaural Sound Reproduction System", AES Conference:24th International Conference: Multichannel Audio, The New Reality (June 2003), Multichannel Audio
- [9] Bill Gardner and Keith Martin, "HRTF Measurements of a KEMAR Dummy Head Microphone", MIT Media Lab Perceptual Computing Technical Report, May 1994.
- [10] Schärer, Zora; Lindau, Alexander, "Evaluation of Equalization Methods for Binaural Signals", AES Convention:126 (May 2009), Spatial Audio and Spatial Perception.
- [11] Hammershøi, Dorte, "Fundamental Aspects of the Binaural Recording and Synthesis Techniques", AES Convention:100 (May 1996), Psychoacoustics, Hearing, and Hearing Instruments
- [12] Aramustafaoglu Ku, Attila; Spikofski, Gerhard, "Binaural Room Scanning and Binaural Room Modeling", Subject:Surround Sound
- [13] S. Pelzer et al., "Room Modeling for Acoustic Simulation and Auralization Tasks: Resolution of Structural Details", Proc. DAGA Berlin (2010)]
- [14] Canal youtube de GRITA, Grupo de Investigación en Telecomunicaciones Avanzadas <http://www.youtube.com/channel/UCZ-NPJ8B-8m8HzlCHPeMvoA>