

**Simulación acústica y gráfica.
Prototipo de navegación de entornos virtuales**

REFERENCIA PACS: 43.55.-Ka Computer simulation of acoustics in enclosures, modeling

Montell, Radha¹; Segura, Jaume²; Giménez, Alicia¹; Planells, Ana¹; Barba, Arturo¹; Cerdá, Salvador³; Lacatis, Radu¹; Vera, Lucía²; Martínez, Bibiana²

1 E.T.S.I.Industriales., Univ. Politécnica de Valencia, Camí de Vera s/n, radmonse@fiv.upv.es, agimenez@fis.upv.es, arbarse@doctor.upv.es, rala1@doctor.upv.es

2 Institut de Robòtica, Universitat de València, Polígon de la Coma s/n, jsegura@uv.es, vemalu@uv.es, bmartine@robotica.uv.es

3 E.T.S.I.Arquitectura, Univ. Politécnica de Valencia, salcerjo@mat.upv.es

Abstract

Graphical and acoustical simulations allow us to have complete information about the studied environments. In this work, we have developed a prototype that integrates graphical and acoustical simulation features allowing the user to navigate in first person through the simulated environment and experience a visual response equivalent to that obtained in the real one. Simultaneously, we obtain the acoustic response for the user's position, which has been precalculated in a number of points through acoustic simulation techniques combined with spatial sound techniques. This prototype has been developed to obtain subjective responses from surveys performed in a set of places that have been acoustically characterized from the objective and subjective point of view. To that end, we have used virtual reality environments (Powerwall, CAVE) to allow for greater immersion in the simulated environment.

Resumen

Los entornos de simulación inmersiva gráfica y acústica permiten tener información suficientemente completa sobre el entorno en estudio. En este trabajo se ha desarrollado un prototipo que integra las características de la simulación gráfica y acústica de recintos, el cual permite al usuario navegar en primera persona a través del entorno simulado y experimentar una respuesta visual equivalente a la que se obtendría en el entorno real. Simultáneamente, se obtiene la respuesta acústica correspondiente a la posición del usuario, la cual ha sido precalculada en una serie de puntos establecidos mediante técnicas de simulación acústica combinadas con técnicas de sonido espacial. Este prototipo se ha desarrollado para ser utilizado en la obtención de respuestas subjetivas mediante encuestas en recintos que han sido caracterizados acústicamente tanto desde el punto de vista objetivo como subjetivo. Para ello, se han empleado entornos de realidad virtual (Powerwall, CAVE) las cuales permiten mayor grado de inmersión en el entorno simulado.

El prototipo se ha desarrollado dentro del proyecto BIA2008-05485.

1. INTRODUCCIÓN

Las técnicas de simulación han sido ampliamente utilizadas en diferentes campos y con diferentes objetivos. Actualmente son imprescindibles ya que permiten observar el comportamiento del objeto de estudio sin ejecutar físicamente el experimento y por tanto, sin ningún riesgo. En este trabajo se combinan técnicas de simulación gráfica y acústica.

Las técnicas de simulación gráfica son utilizadas actualmente en numerosos campos de investigación (medicina, aviación, industrias, arquitectura, etc.). La aplicación de éstas ha permitido muchos avances gracias a que permiten la visualización de aquello sobre lo que se está trabajando y por tanto, una mayor comprensión. En el campo de la acústica también aportan numerosas ventajas. La simulación gráfica se ha utilizado para observar y comprender los complejos resultados de la simulación acústica de recintos [1][2]; para percibir el entorno de estudio de forma más realista [3]; en la creación de entornos virtuales [4]; en la reconstrucción de edificios [5][6]; entre otros.

De forma equivalente, las técnicas de simulación acústica permiten obtener la respuesta acústica de un espacio, abierto o cerrado, a partir de una serie de parámetros. Al igual que la simulación gráfica, éstas también han sido aplicadas en numerosos campos.

La simulación acústica, en todas sus fases (desde la fase de modelado y predicción hasta la fase de post-procesado y auralización), tiene numerosas aplicaciones tanto en el campo del diseño acústico de salas (como entrenamiento para arquitectos, acústicos, técnicos de audio y músicos), diseño de entornos para videojuegos y realidad virtual [7] (incrementando su nivel de realismo), diseño de habitáculos en la industria de la automoción, estudios en psicoacústica y calidad sonora, etc. [8]. Por otra parte, la auralización combinada con la telemática también presenta múltiples aplicaciones en sistemas de telepresencia [9][10], así como en aplicaciones de HDTV, etc.

Teniendo en cuenta la potencia predictiva que tienen los algoritmos utilizados en simulación acústica y su ajuste a las medidas, se puede observar el resultado final antes de llevar a cabo la construcción o modificación, permitiendo, de esta forma, corregir posibles errores durante la fase de modelado y no ya *in situ*, rediseñar experimentos, etc.

Existen diferentes grupos trabajando en el campo de la integración de simulación gráfico-acústica aplicada a los diversos usos anteriormente citados. Entre los más destacados en Europa se encuentran: el grupo de la Universidad de Chalmers (con M. Kleiner y Dählenback), el grupo de la Universidad de Aquisgrán (con M. Vörländer), el grupo de la Universidad Técnica de Dinamarca (con J.H. Rindel), el grupo de la Universidad de Tecnología de Helsinki (con L. Savioja, T. Lokki y V. Pulki), el grupo de la Universidad de Parma (con A. Farina) y el de la Universidad de Ferrara (con R. Pompoli).

Para tener una visión de la magnitud de esta área de investigación en acústica, mostramos proyectos con una temática similar, de aplicación de simulación gráfica y acústica para rehabilitación de edificios, incluidos en convocatorias de proyectos del programa marco europeo: proyecto ERATO (<http://server.elektro.dtu.dk/www/oldat/erato/>) y proyecto CAHRISMA (<http://server.elektro.dtu.dk/www/oldat/cahrisma/>). Además, la producción científica ha sido creciente en esta temática.

En este contexto, nuestro proyecto pretende ser un complemento para la evaluación subjetiva de las salas implicadas. Nuestro propósito no es centrarnos en la tecnología exclusivamente, sino utilizarla para comprobar si los resultados de un proyecto anterior (BIA2003-09306) aportan conocimiento para la evaluación subjetiva de salas de audición (aplicado concretamente a edificios del patrimonio histórico-artístico nacional) y comprobar la validez de las respuestas subjetivas obtenidas a partir de la simulación respecto a las obtenidas en los entornos reales.

El prototipo presentado permite navegar a través de entornos virtuales de salas y edificios del patrimonio histórico-artístico de la Comunidad Valenciana, experimentando una respuesta visual y acústica equivalente a la que se obtendría en el entorno real. A continuación se expondrá la metodología desarrollada, junto con las diferentes alternativas del prototipo. Se detallará nuestra experiencia en la aplicación realizada del prototipo y las dificultades encontradas. Y finalmente, se presentarán las conclusiones derivadas de la aplicación.

2. METODOLOGÍA

Las técnicas de simulación gráfica nos permiten visualizar la sala a estudiar de forma realista. Para ello, la sala ha sido modelada a partir de sus planos arquitectónicos y texturizada partiendo de fotos de la propia sala, teniendo en cuenta la iluminación real de la misma.

La simulación acústica se ha realizado con el programa CATT Acoustics [11]. A partir del modelo construido con los planos de la sala, se ha realizado la conversión al formato de geometría del CATT. En primer lugar, se ha comprobado la corrección del modelo (debe ser un espacio cerrado, sin aberturas; sin superficies no planas o con las normales inversas; sin definiciones de vértices y planos repetidos; etc.) y se ha depurado el modelo hasta conseguir un informe de depuración del CATT que nos indique que no existen errores (Figura 1). Éste puede llegar a ser un proceso arduo, ya que la dificultad en la corrección del modelo geométrico depende en gran medida de la experiencia del usuario del programa. A continuación, se ha ajustado el modelo mediante los parámetros obtenidos a partir de las medidas de la sala real y se ha obtenido la respuesta impulsiva de la sala. Seguidamente, se ha auralizado la señal en una serie de puntos establecidos obteniendo así una serie de ficheros de audio.

```

CATT-Acoustic v8.0j debug:
-----
DUPLICATE PLANE ID's      : no !
DUPLICATE CORNER ID's    : no !
DUPLICATE CORNERS        : no !
SINGLE-CONNECTED CORNERS  : no !
INACCURATE PLANE CORNERS : no !
EDGES CUTTING/TOUCHING   : no !
POSSIBLY REVERSED PLANES : no !
-----

```

Figura 1: Informe de depuración del CATT de un modelo sin errores.

Respecto a la presentación de la señal acústica, se han escogido dos técnicas diferentes: binaural y 5.1. En la presentación binaural, la señal se convoluciona con filtros HRTF i el receptor la percibe a través de auriculares. Esta técnica permite al receptor aislarse de los posibles ruidos que pueda haber en la sala de reproducción. En la presentación 5.1 la señal se transmite a través de 6 altavoces (5 altavoces que permiten la ubicación espacial de las fuentes a frecuencias medias y altas y 1 *subwoofer* para las frecuencias bajas). Ésta permite la recepción simultánea para múltiples receptores, generando una serie de fuentes virtuales que permiten a los receptores tener una sensación referencial de espacialidad que les permite ubicar las fuentes en el entorno. No obstante, los sonidos de la sala de reproducción pueden interferir con el audio simulado, por lo que se deberá tener especial cuidado en minimizar posibles ruidos.

El prototipo de navegador implementa un manejador de cámara que permite al usuario moverse a través de la sala en primera persona imitando el movimiento real a partir de movimientos básicos (avanzar, retroceder, cabeceo (girar derecha, girar izquierda) y balanceo

(mirar arriba, mirar abajo)) y teniendo en cuenta las colisiones con paredes y mobiliario. El control de movimiento se realiza intuitivamente mediante las teclas direccionales del teclado.

El prototipo se ha desarrollado en lenguaje C++, sobre las librerías OSG (OpenSceneGraph) [12] y OSGAL (OpenSceneGraph Audio Library) [13]. OSG es una potente librería para desarrollar aplicaciones gráficas 3D multiplataforma, escrita en C++ y OpenGL, la librería más usada en aplicaciones gráficas. La librería OSGAL proporciona una serie de herramientas para manejar sonido espacial en las aplicaciones OSG. Está escrita en OpenAL y se integra con OSG. Ambas librerías permiten aprovechar la potencia de OpenGL y de OpenAL ofreciendo un entorno de programación a más alto nivel y por tanto, más intuitivo.

El sonido reproducido en el navegador se obtiene mediante la combinación de las características de la librería acústica con el resultado de la simulación calculada por el CATT. OSGAL, junto con una correcta configuración de la tarjeta de audio, proporciona filtrado con funciones HRTF, de forma que este paso se puede realizar directamente con la librería. Asimismo, la librería acústica de OSG implementa la oclusión de sonidos causada por la geometría definida en OSG y define un modelo de atenuación.

2.1. Prototipo de Navegador Básico

Nuestro objetivo actual es desarrollar un navegador con la finalidad de utilizarlo para obtener respuestas subjetivas sobre el entorno proyectado. Para esto, se recrea la experiencia del espectador durante un concierto, en el cual éste se aposentaría en una butaca concreta en la que se mantendría durante toda la obra. En el entorno virtual, se predefinen varios puntos de vista establecidos sobre la zona de asientos, cada uno con su correspondiente auralización.

Ésta es la versión más simple del navegador. Para su ejecución es necesario haber procesado con anterioridad las auralizaciones para todos los puntos definidos. Éstas se encuentran almacenadas en el disco duro y cuando se establece una posición en las butacas, se reproduce el fichero de sonido correspondiente a la auralización en esa posición.

Mediante el programa CATT, obtenemos la señal acústica para una posición fuente-receptor fija. Si esta señal se convoluciona con filtros HRTF se obtiene, además, para una orientación determinada. No obstante, nuestra intención es que el receptor, una vez situado en su posición en la sala, sea capaz de orientar la cabeza mientras escucha la pieza musical auralizada. Para ello, se hace uso del filtrado HRTF que proporciona la librería OSGAL que, en tiempo real, permite obtener el sonido correspondiente para cada orientación.

En este prototipo, la inclusión de la librería OSGAL permite al usuario navegar a través de la escena y percibir los cambios en el sonido, según la distancia a la fuente o la existencia de oclusiones; no obstante, es una aproximación, ya que esta librería no tiene en cuenta reflexiones ni refracciones, y no maneja materiales.

La siguiente aproximación permite la navegación a través de la sala mientras se percibe una respuesta acústica más aproximada a la real.

2.2. Prototipo de Navegador con Interpolación

Partiendo de la estructura del navegador básico, se puede realizar una serie de cambios que permitirán navegar a través de toda la sala obteniendo una respuesta acústica mejor simulada que en la versión anterior. En esta aproximación nuestro objetivo es obtener la respuesta impulsiva en cualquier punto de la sala, de forma que el usuario pueda moverse libremente y, al mismo tiempo, percibir correctamente la respuesta acústica en cada posición.

En este caso, se define un mayor número de posiciones, distribuidas de forma equidistante en forma de rejilla y cubriendo toda la zona de butacas, en las cuales se calculará la RI (respuesta impulsiva) de cada punto. A mayor número de puntos, y por tanto, mayor densidad, mejor aproximación a la RI real, pero será necesario más espacio en disco duro para almacenamiento

de las RI calculadas. La RI de cualquier posición se obtendrá como combinación de las RI más cercanas, aplicando un algoritmo de interpolación (Figura 2).

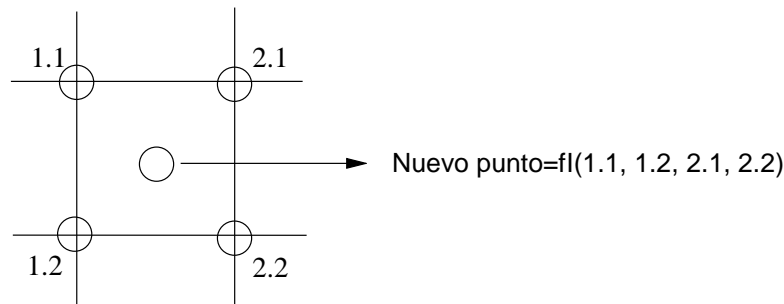


Figura 2: Rejilla de puntos sobre la zona de butacas. Función de interpolación (fl) aplicada a los puntos más cercanos.

A partir de la nueva RI, se calculará la auralización del nuevo punto, convolucionando la RI con la señal anecoica y con las funciones HRTF.

Esta aproximación tiene un coste espacial proporcional al número de puntos que formen la rejilla, ya que hay que precalcular y almacenar tantas RI como puntos tenga. Por otro lado, la función de interpolación tiene coste lineal y el algoritmo de convolución es en tiempo real, con lo que la respuesta acústica se obtiene en tiempo real, teniendo en cuenta que las RI han sido precalculadas con antelación y este coste no influye en el cálculo de la auralización.

Paralelamente, el programa CATT proporciona una aplicación en la cual implementa una variante de esta aproximación. Nuestra intención es poder comparar ambas implementaciones.

2.3. Prototipo de Navegador con Catt-Walker

El programa CATT incluye una aplicación para crear auralizaciones en tiempo real: CATT-Walker [14]. Al igual que en la aproximación anterior, se establece una distribución de puntos en donde se calcularán las respuestas impulsivas, en este caso en formato Ambisonics. A partir de las RI se realizan interpolaciones para obtener la RI en nuevas posiciones, y posteriormente, se auraliza. CATT-Walker permite obtener la señal auralizada en formato binaural o en Ambisonics.

Esta aplicación se complementa con una interfaz que permite utilizar el CATT-Walker como motor acústico, es decir, se puede controlar remotamente desde una aplicación que, en nuestro caso, corresponde al navegador. De esta forma, es posible establecer una estructura modular: el navegador se encarga de la parte gráfica y el CATT-Walker de la acústica, comunicándose ambos para intercambiar información acerca de la posición en la escena y de la auralización calculada. Esta aplicación permitiría que los módulos se encontrasen en ordenadores diferentes, aprovechando así la máxima potencia posible.

3. APLICACIÓN

Actualmente se ha completado y probado la primera aproximación de las presentadas, ya que ésta cubre nuestras actuales necesidades.

Se modeló una sala de usos múltiples de la Universidad Politécnica de Valencia, el Paraninfo, a partir de los planos arquitectónicos. Se texturizó y se ajustó la iluminación. Paralelamente y a partir del mismo modelo, se realizó la simulación acústica y la auralización en el programa CATT, obteniendo un conjunto de ficheros de audio correspondientes a las posiciones fijadas. Para esta sala se fijaron tres posiciones: en la primera fila, el asiento extremo derecho; en la última fila, el asiento extremo izquierdo; y el asiento central de la sala.

El navegador integra el modelo gráfico y las auralizaciones, de forma que el usuario puede situarse en cada una de las posiciones definidas y escuchar la pieza musical. Al mismo tiempo, puede orientar el punto de vista manteniendo una correcta percepción acústica.

Las pruebas del navegador se realizan en dos entornos de realidad virtual inmersiva: CAVE (Cave Automatic Virtual Environment) y Powerwall.

El CAVE es un espacio compuesto por cuatro superficies en forma de cubo retroproyectadas (tres paredes y suelo). En el CAVE se proyectan imágenes estereoscópicas que, con la ayuda de unas gafas polarizadas sincronizadas con los proyectores, son percibidas como 3D. Además de las gafas, el usuario lleva unos sensores que captan sus movimientos en la sala y permiten al ordenador actualizar la imagen proyectada de acuerdo al nuevo punto de vista. Esta sala incorpora una instalación de altavoces 5.1 y permite la utilización de auriculares.

La Powerwall (Figura 3) es una gran pantalla retroproyectada de alta definición (2x5 metros) que, al igual que el CAVE, emite imágenes estereoscópicas que se visualizan con la ayuda de gafas polarizadas sincronizadas con los proyectores. La Powerwall se encuentra en una sala con capacidad para 25 personas situadas en una grada de 4 filas. Actualmente, ésta dispone de un sistema de altavoces 5.1 para emitir sonido espacial.



Figura 3: Powerwall con proyección del Paraninfo.

Ambas salas tienen sus ventajas e inconvenientes. En el CAVE se dispone de un sistema de seguimiento del movimiento que aumenta la sensación de inmersión en el entorno y que transmite las posiciones tanto al ordenador encargado de los proyectores como al navegador, que actualiza el sonido auralizado en tiempo real. Debido a estos sensores, sólo puede hacer uso del CAVE una persona a la vez. Por otro lado, los proyectores que se utilizan en el CAVE tienen una resolución baja (1024 x 1024). La Powerwall, en cambio, dispone de unos proyectores de alta resolución que permiten apreciar con claridad todos los detalles de la sala modelada. Ésta tecnología, al no disponer del sistema de seguimiento, puede ser utilizada simultáneamente por varias personas. No obstante, no todos los asientos pueden ser ocupados, ya que la reproducción en el sistema de audio 5.1 sólo se percibe correctamente en el centro de la sala, y por tanto, el espacio de recepción correcta es de aproximadamente 10 personas. Además, en la reproducción por altavoces hay que tener en cuenta que el sonido de la sala de reproducción (aire acondicionado, otras personas hablando, sonidos externos a la sala, etc.) interferirá con el sonido emitido auralizado. La solución propuesta para los dos últimos problemas es utilizar auriculares. Estos aislarán el sonido de la sala y permitirán una correcta percepción acústica en toda la audiencia.

En las pruebas realizadas, se proyectó la aplicación y se mostró a un grupo de personas, clasificados en expertos y no expertos musicalmente. A éstos se les mostró la encuesta y se les explicaron todos los puntos. A continuación, escucharon y vieron la simulación del Paraninfo a través del navegador desde las tres posiciones fijadas. Posteriormente, rellenaron la encuesta

indicando además, la preferencia por alguna de las posiciones, en cuyo caso, debían indicar la razón de la preferencia en términos acústicos.

El objetivo de este trabajo es obtener respuestas subjetivas mediante el uso del navegador propuesto en este artículo. En proyectos anteriores, se realizó un pase de encuestas en diferentes salas durante representaciones musicales con el objetivo de obtener un conjunto de parámetros que caractericen subjetivamente las salas y sirvan de evaluadores de la calidad. El objetivo en este caso es, además, comprobar si la simulación es subjetivamente válida. Es decir, que el sonido percibido por la audiencia es equivalente al que se percibiría en la sala real. De esta forma, el pase de encuestas en futuros estudios se podría realizar a partir de las simulaciones, lo que supondría un ahorro en tiempo y esfuerzo invertidos en ello.

4. CONCLUSIONES

En este artículo se presenta un prototipo de navegador que integra simulación acústica y gráfica, con el objetivo de servir como aplicación para la recogida de información subjetiva acerca de la calidad del sonido y gráficos simulados. Se ha precalculado la respuesta impulsiva y se ha auralizado en tiempo real, de forma que el usuario puede percibir los cambios instantáneamente. Se han presentado varias aproximaciones del navegador, cada una con diferentes objetivos. Finalmente, se ha descrito la experiencia de la aplicación del navegador en dos entornos de realidad virtual inmersivos: la Powerwall y el CAVE.

Actualmente, se encuentra implementada y en funcionamiento la primera aproximación. Este es un estudio en progreso en el que se está trabajando en la implementación de las otras dos aproximaciones y su comparación en términos de complejidad y rendimiento. Asimismo, se planea probar el navegador en una tercera tecnología de realidad virtual: el Visionario (entorno de proyección similar a la Powerwall, con pantalla semicilíndrica y tres cañones de proyección).

5. AGRADECIMIENTOS

Este trabajo ha sido subvencionado por el Ministerio de Ciencia e Innovación mediante el proyecto de investigación BIA2008-05485.

6. BIBLIOGRAFÍA

- [1] A. Stettner and D. P. Greenberg. 1989. Computer graphics visualization for acoustic simulation. In Proceedings of the 16th annual conference on Computer graphics and interactive techniques (SIGGRAPH '89). ACM, New York, NY, USA, 195-206. DOI=10.1145/74333.74353 <http://doi.acm.org/10.1145/74333.74353>
- [2] Tsingos, N., Funkhouser, T., Ngan, A., and Carlbom, I. 2001. Modeling acoustics in virtual environments using the uniform theory of diffraction. In Proceedings of the 28th Annual Conference on Computer Graphics and interactive Techniques SIGGRAPH '01. ACM, New York, NY, 545-552. DOI= <http://doi.acm.org/10.1145/383259.383323>
- [3] Pope, J. and Chalmers, A. 1999 Multi-Sensory Rendering: Combining Graphics and Acoustics. Technical Report. UMI Order Number: CS-EXT-1999-182., University of Bristol.
- [4] J. Huopaniemi, L. Savioja, T. Takala. 1996. DIVA virtual audio reality system. Proc. Int. Conf. Auditory Display (ICAD'96)
- [5] P. Dias, G. Campos, V. Santos, R. Casaleiro, R. Seco, and B.S. Santos. 3D reconstruction and spatial auralization of the Painted Dolmen of Antelas., Proc. SPIE 6805, 68050Y (2008), DOI:10.1117/12.766607
- [6] J.H. Rindel, A.C. Gade, M.L. Nielsen. The virtual reconstruction of the ancient Roman concert hall in Aphrodisias, Turkey. The Sixth International Conference on Auditorium Acoustics, 316-323. 2002.

- [7] Faria, Regis; Zuffo, Joao; “An Auralization Engine Adapting a 3D Image Source Acoustic Model to an Ambisonics Coder for Immersive Virtual Reality”. Audio Engineering Society Conference: 28th International Conference: The Future of Audio Technology--Surround and Beyond. June, 2006 (<http://www.aes.org/e-lib/browse.cfm?elib=13838>)
- [8] Mendel Kleiner & al; “Auralization: An overview”. Journal AES, Vol.41, No.11 (1993), pp 861-875.
- [9] Seong-Hoon Kang; Sung-Han Kim; “ Realistic audio teleconferencing using binaural and auralization techniques”. ETRI Journal, vol. 18, num. 1, April 1996.
- [10] Paul Henderson, Brian Lonsway, Richard J. Radke, Rendell Torres, Yasushi Shimizu, and Seema Jaisinghani, Coupled Auralization and Virtual Video for Full-Room Tele-Performance Environments. Proceedings of the International Workshop on Immersive Telepresence 2002 (in conjunction with ACM Multimedia 2002), December 6, 2002, Juan Les Pins, France (<http://www.ecse.rpi.edu/~rjradke/papers/radkeitp02.pdf>)
- [11] <http://www.catt.se/>
- [12] Open Scene Graph: www.openscenegraph.org/
- [13] Open Scene Graph Audio Library: <http://www.vrlab.umu.se/research/osgAL/>
- [14] BI Dahlenbäck, M Strömberg. Real time walkthrough auralization—the first year. Proceedings of the Institute of Acoustics, 2006