



FIA 2018

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

GRABACIÓN Y PRESENTACIÓN DE PAISAJES SONOROS EN ENTORNOS DE LABORATORIO

PACS: 43.38.Md

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

Palabras Clave: paisajes sonoros, audio espacial, grabación, renderización.

ABSTRACT

The evaluation of soundscapes is mainly carried out through field surveys, using soundwalking methodologies. They are costly processes in time and depend on variables such as the concentration of the participants. The recreation of soundscapes in laboratory environments using virtual reality tools allows a more controlled evaluation, being able to focus the subject's attention on the acoustic aspects.

This communication establishes a methodology for the recording and presentation of soundscapes in laboratory environments, using first-order ambisonics format for recording and spatial audio rendering techniques for presentation.

RESUMEN

La evaluación de paisajes sonoros se realiza principalmente mediante encuestas de campo, utilizando metodologías de soundwalking o paseos sonoros. Son procesos costosos en tiempo y sujetos a variables como la concentración de los participantes. La recreación de paisajes sonoros en entornos de laboratorio mediante herramientas de realidad virtual permite una evaluación más controlada, pudiendo focalizar la atención del sujeto en los aspectos acústicos.

En esta comunicación se establece una metodología para la grabación y presentación de paisajes sonoros en entornos de laboratorio, empleando formato ambisonics de primer orden para la grabación y técnicas de renderización de audio espacial para la presentación.

1. INTRODUCCIÓN

El paseo sonoro (soundwalk) es una de las principales herramientas de investigación en el campo de los paisajes sonoros e incluye la experimentación de una secuencia de lugares caracterizados

FIA 2018

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

por diferentes atmósferas. En su sentido más amplio, es el acto de caminar a través de un escenario con un enfoque crítico en la escucha de los sonidos que se perciban. Se utiliza para análisis cuantitativos y cualitativos [1], al realizar cuestionarios de evaluación del paisaje sonoro y recopilar datos objetivos, es decir, mediciones y grabaciones, para su posterior análisis [2].

La evaluación de paisajes sonoros en entornos de laboratorio elimina posibles factores de sesgo que son impredecibles in situ. Las nuevas tecnologías ofrecen herramientas para investigar variables específicas relativas a la percepción en ambientes muy controlados, creando configuraciones potencialmente multisensoriales [3]. Sin embargo, la denominada validez ecológica [4] de los datos recopilados bajo tales condiciones ha sido cuestionada a menudo, aunque también se han publicado resultados prometedores [5] que aseguran que entornos sonoros grabados y reproducidos mediante ambisonics pueden ser ecológicamente válidos para la evaluación de paisajes sonoros.

Tras un análisis de la tecnología disponible actualmente para la grabación y presentación de entornos sonoros, en esta comunicación se presenta el proceso implementado para la evaluación de paisajes sonoros en el Laboratorio de Acústica de la Universidad Pública de Navarra, junto con un ejemplo de su aplicación.

2 . TÉCNICAS DE GRABACIÓN Y PRESENTACIÓN DE PAISAJES SONOROS

Una recreación verosímil de un paisaje sonoro debe imitar nuestra percepción de la escena sonora tridimensional mediante el empleo de adecuadas técnicas de grabación, procesamiento y presentación. Los dispositivos actuales para la grabación y reproducción tienen suficiente fidelidad como para lograr una interpretación adecuada de las componentes frecuenciales y temporales del campo sonoro. Sin embargo, la interpretación de la dimensión espacial del sonido es más compleja pues exige un intercambio constante de información entre los estímulos recibidos en cada oído. Para localizar correctamente una fuente en el espacio, las propiedades de intensidad, fase y retardo deben presentarse en ambos oídos con precisión, por lo que la complejidad de la reproducción se incrementa. Por tanto, es necesario prestar especial atención a las técnicas de grabación y presentación de audio espacial desde una perspectiva tecnológica, incidiendo en su relación con la percepción del paisaje sonoro [6].

El esquema habitual (Figura 1) del proceso de grabación y renderización de audio espacial consiste en un entorno original donde las señales son registradas con micrófonos específicos, una escena virtual en la que puede parametrizarse la espacialización y, finalmente, un entorno destino donde la escena se renderiza mediante auriculares o altavoces.

A diferencia de la visión, que captura información principalmente en la dirección de enfoque, nuestros oídos reciben una compleja mezcla de sonidos directos y reflexiones procedentes de diferentes direcciones. Nuestro cerebro analiza estos sonidos y los conecta mentalmente con las fuentes sonoras que los han generado, evocando una determinada escena sonora de forma que incluso podemos concentrarnos en sonidos específicos y su procedencia. Para localizar la dirección de llegada de un sonido en el plano horizontal utiliza fundamentalmente la diferencia entre los estímulos recibidos en un oído y otro, tanto de tiempo de llegada (Interaural Time Difference, ITD) como de nivel (Interaural Level Difference, ILD). Para percibir elevación, el pabellón auditivo genera un filtrado de las altas frecuencias que depende de la dirección y ángulo de llegada. Asimismo, las reflexiones en los hombros también producen una ligera alteración a frecuencias algo más bajas que somos capaces de detectar cuando la señal a localizar carece de componentes a altas frecuencias. Finalmente, la percepción de distancia incluye entre sus mecanismos la diferente intensidad, la atenuación del aire, que es superior a altas frecuencias, o la relación entre el sonido directo y la reverberación.

FIA 2018

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

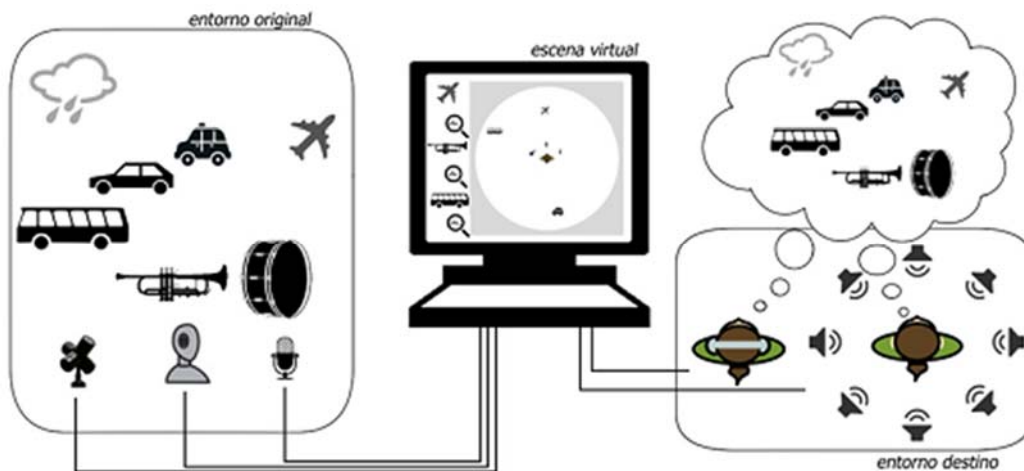


Figura 1. Proceso de grabación y renderización de audio espacial

2.1 Grabación

Para generar ambientes sonoros en un laboratorio, los investigadores suelen utilizar grabaciones binaurales reproducidas por auriculares [7]. Una grabación binaural captura directamente el sonido recibido en los oídos izquierdo y derecho, en el tímpano si se registran las señales con una cabeza binaural (dummy head) o a la entrada del conducto auditivo si se utiliza un micrófono binaural. Es el formato más similar a la audición humana cuando se reproduce a través de auriculares calibrados, pues incorpora automáticamente los mecanismos que necesita el cerebro para evocar una escena, registrando directamente las alteraciones que sufre el sonido al interactuar con nuestro cuerpo. La principal ventaja de esta técnica es que no requiere procesamiento en su reproducción mediante auriculares. Los principales inconvenientes son que la orientación de la cabeza no puede ser modificada con posterioridad, limitando la interactividad de la presentación, y que las HRTF inherentes a las grabaciones (del operario o de la cabeza binaural) difieren de las del oyente.

Otra forma de registrar las características espaciales del entorno sonoro es utilizar múltiples micrófonos. Las técnicas más clásicas utilizan diversas configuraciones para capturar sonido en formato estéreo o surround. Estas representaciones están directamente relacionadas con una disposición específica de los altavoces utilizados para la presentación porque son las propias señales correspondientes a cada altavoz las que se almacenan y transmiten. Otras técnicas emplean mayor número de micrófonos en forma de arrays. Como resultado, las grabaciones de los micrófonos ya no se utilizan directamente, sino que requieren un procesamiento posterior para acomodarlas a las diferentes configuraciones de los sistemas de reproducción. Estas técnicas son ampliamente utilizadas en localización y separación de fuentes, reducción de ruido, cancelación de eco, etc.

Finalmente, existen métodos para la grabación de audio espacial basados en técnicas de micrófonos coincidentes que producen señales que pueden ser transformadas en señales B-format o de formato ambisonics. Este formato se fundamenta en la descomposición del campo sonoro en funciones armónicas esféricas que constituyen una base ortogonal completa, lo que significa que el campo sonoro puede ser reconstruido partiendo del conocimiento de los coeficientes de cada función y viceversa. Aunque desarrollado en los años 70, ambisonics ha ganado peso recientemente desde que YouTube, Oculus VR ó Facebook lo adoptaron como un estándar para sus videos de 360 grados. En la actualidad, ambisonics se ha convertido en un estándar de audio espacial ampliamente utilizado en los sistemas de realidad virtual, entre otros



FIA 2018

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

factores por ser un sistema de registro centrado en la escena a renderizar. Su principal ventaja es que no contiene ningún requisito previo de configuración de altavoces para su presentación. Además, con un único registro y procesamiento posterior puede variarse la orientación, inclinación o rotación del campo sonoro en el punto analizado, algo complicado de lograr con el resto de sistemas.

2.2 Parametrización

En general, cualquier técnica multicanal de grabación junto con técnicas de procesado de señal permite obtener información sobre el campo sonoro en cuanto a localización de las fuentes sonoras, la señal "limpia" correspondiente a cada fuente y las propiedades del entorno sonoro, incluyendo el campo reverberante. Esta parametrización del campo sonoro mejora potencialmente la calidad de la reproducción, además de ofrecer un mayor control en la síntesis de paisajes sonoros, aunque dificulta su implementación en aplicaciones que requieren desarrollarse en tiempo real.

La codificación de audio direccional (Directional Audio Coding, DirAC) se revela como la representante de una nueva familia de enfoques que se basan explícitamente en mecanismos psicoacústicos [8]. DirAC es un método no lineal en el dominio tiempo-frecuencia para reproducir sonido espacial que explota el hecho de que los mecanismos para la percepción de sonido difuso son muy diferentes de aquellos relacionados con la percepción del sonido directo. De esa manera, una señal de audio espacial se descompone en componentes difusos y no difusos, y los dos grupos de señales son renderizados mediante diferentes técnicas.

2.3 Presentación

Los auriculares son especialmente adecuados para la presentación de señales binaurales ya que la señal en cada oído puede controlarse individualmente. Existen técnicas de presentación de señales binaurales que emplean altavoces, lo que en muchas situaciones se considera más conveniente. En este caso se deben aplicar filtros de cancelación cruzada (cross-talk cancellation) para corregir la interferencia entre los diferentes caminos altavoz-oído. La denominada presentación transaural manifiesta algunas limitaciones en cuanto a la zona de escucha (sweet spot), que es muy reducida. En cualquier caso, es importante que los movimientos del oyente, especialmente la rotación de la cabeza, sean rastreados en tiempo real (real-time head tracking) y considerados en la presentación.

Despreciando la influencia de otros aspectos, como la visión, la vibración o la conducción ósea, una recreación perfecta de la presión sonora en los tímpanos daría como resultado una reproducción indistinguible (auténtica) de una escena sonora virtual. Sin embargo, desde el punto de vista de la percepción humana este complicado objetivo técnico no necesita cumplirse a la perfección para obtener una impresión perfectamente plausible [9]. Existen métodos menos complejos de representación de audio espacial con altavoces que utilizan ITD e ILD como mecanismos predominantes. Para una determinada posición, generalmente situada a la misma distancia de los altavoces, los mecanismos que emplea el cerebro para la localización pueden modelizarse mediante paneo en amplitud (amplitude panning), o lo que es lo mismo, generando diferencias de nivel de reproducción entre los altavoces. El estéreo, que sigue siendo el método de presentación de audio más extendido, o los sistemas de sonido envolvente (surround sound) como el 5.1 o el 9.1, que incluye altavoces elevados, son algunas de las técnicas de paneo en amplitud más conocidas. En estos sistemas, la posición de los altavoces es fija. Una mejora natural son los sistemas de presentación de audio espacial independientes de la configuración de los altavoces empleada, como VBAP (Vector-Base Amplitude Panning) o Ambisonics. VBAP reformula mediante vectores las clásicas técnicas de paneo en amplitud permitiendo generalizar el estéreo entre pares de altavoces circundantes a tríos de altavoces, añadiendo por tanto una

FIA 2018

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

dimensión adicional, la elevación. Ambisonics, por su parte, utiliza para la renderización todos los altavoces disponibles para presentar una sola fuente virtual.

Con el objetivo de lograr una presentación equilibrada en un área de escucha más amplia que los sistemas anteriores, en las últimas décadas se han desarrollado métodos que sintetizan el campo sonoro en áreas extensas mediante arreglos formados por una gran cantidad de altavoces. Estos enfoques, que pueden alcanzar varios cientos de canales o incluso más, se conocen como técnicas de síntesis de campo sonoro (Sound Field Synthesis, SFS) [10]. Las que reciben más atención en la actualidad son Wave Field Synthesis (WFS) y Near-field Compensated Higher Order Ambisonics (NFC-HOA). Para WFS existe tanto una comunidad de investigación activa como algunos productos comerciales. Por su parte, NFC-HOA se emplea principalmente en instituciones de investigación.

3. IMPLEMENTACIÓN

A la hora de implantar un sistema de grabación-presentación de paisajes sonoros en un laboratorio es importante garantizar la calidad en la renderización final de la escena sonora registrada. Un equipamiento sofisticado, además de un elevado coste económico, precisa contar con personal técnico cualificado. Por su parte, una simplificación excesiva en cualquier fase del proceso dificultaría registrar y reproducir un entorno sonoro con la suficiente precisión perceptual. El dispositivo implementado en el Laboratorio de Acústica de la Universidad Pública de Navarra se considera que mantiene el equilibrio entre ambos aspectos. Por un lado mantiene el rigor científico suficiente para respaldar la validez ecológica de futuras investigaciones. Por otro, los protocolos de registro, procesamiento y presentación se han simplificado y automatizado, aumentando la reproducibilidad del proceso.

En la Figura 2 se muestra el diagrama de flujo del procedimiento implementado. El paisaje sonoro se registra in situ mediante un micrófono ambisonics de primer orden - Sennheiser Ambeo - equipado con cuatro cápsulas cardioides posicionadas en forma de tetraedro. Las señales obtenidas (4ch A-format) se almacenan en una grabadora multipistas - zoom F4 - previamente calibrada en laboratorio. Cada registro se acompaña con una fotografía o vídeo 360 de la localización.

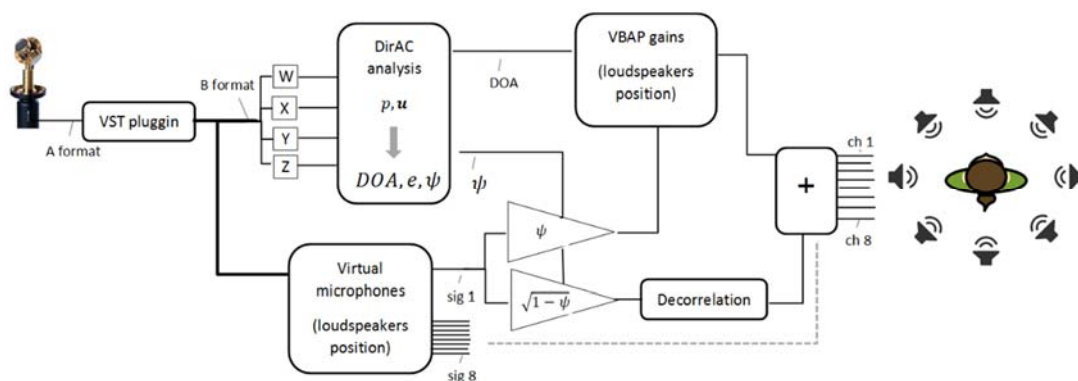


Figura 2. Diagrama de flujo del procedimiento implementado

La conversión a señales B-format se realiza mediante un plugin VST proporcionado por el fabricante, obteniendo finalmente una señal W proporcional a la presión p (omnidireccional) y tres al gradiente de presión en cada eje XYZ (figura de 8), pudiendo estimarse el vector velocidad de partícula u en el punto de registro. El sistema es portátil y no requiere una instalación

FIA 2018

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

compleja, lo que permite registrar numerosas localizaciones en un corto periodo de tiempo. El formato de audio resultante es compatible con el procesamiento de audio espacial de plataformas online. Esta discretización espacial del campo sonoro, además de permitir su posterior manipulación, mantiene su capacidad para adaptarse a futuras técnicas de renderizado que puedan desarrollarse.

A partir de la información relativa a la presión y a la velocidad se obtiene el vector de intensidad del campo sonoro \mathbf{i} , que expresa la dirección y la magnitud del flujo de energía sonora. Analizando este vector puede estimarse la dirección de llegada predominante (DOA) como:

$$DOA = \text{ang}(E[-\mathbf{i}]) \quad (1)$$

donde $E[\cdot]$ es el valor esperado, que se implementa con un promediado temporal. Por otro lado, se obtiene la densidad de energía e [11]:

$$e = \frac{\rho_0}{c} \|\mathbf{u}\|^2 + \frac{|p|^2}{2\rho_0 c^2} \quad (2)$$

donde ρ_0 es la densidad del aire, c la velocidad del sonido, y $\|\cdot\|$ el módulo, y finalmente se calcula el parámetro difusión ψ según [12]:

$$\psi = 1 - \frac{\|E[\mathbf{i}]\|}{cE[e]} \quad (3)$$

A partir de estos parámetros y de las señales WXYZ provenientes del micrófono B-format se realiza la síntesis de las señales enviadas a los altavoces. En este procedimiento la señal se divide en fragmentos consecutivos mediante ventanas temporales parcialmente solapadas. Cada una de ellas se procesa mediante FFT. El tamaño de la ventana, 20 ms con un hop size del 50%, permite codificar los valores de los parámetros direccionales con la resolución temporal y frecuencial de la percepción humana.

En la versión más simple del método DirAC únicamente se distribuye la señal del micrófono omnidireccional (W). Ésta se divide en dos flujos según el valor de ψ . La parte no-difusa (directa) se reproduce como si fuera una fuente virtual puntual situada en la dirección definida por el valor de DOA aplicando VBAP, es decir, utilizando únicamente dos o tres altavoces (en configuraciones 2D o 3D respectivamente) y mediante paneo en amplitud. Por otro lado, la parte difusa se reproduce por todos los altavoces que rodean al oyente tras aplicar filtros de decorrelación, lo que idealmente generaría el pretendido campo difuso. De esta manera se reduce notablemente la coherencia entre las señales aplicadas a los altavoces en la presentación, principal fuente de artefactos que se manifiestan en la renderización ambisonics tradicional [13].

Una implementación más avanzada del método propone aplicar de forma diferente el método descrito anteriormente para sintetizar las señales enviadas a cada altavoz. En lugar de distribuir la señal omnidireccional W , crea para cada canal disponible en la presentación una señal conforme a un micrófono virtual, con directividad de primer orden, alineado hacia el altavoz correspondiente (Virtual Microphone DirAC). Estas señales, que también se obtienen a partir de las originales WXYZ, son las que finalmente se distribuyen, mediante VBAP y filtros de decorrelación en función de los valores de los parámetros direccionales obtenidos en el análisis preliminar. Este método multicanal mejora la calidad de la reproducción cuando se compara con otros métodos de reproducción en test subjetivos [14] y es el que se ha implementado, con 8 canales, para la presentación de paisajes sonoros en el laboratorio. Esta configuración permite

FIA 2018

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

análisis de situaciones 2D, con los 8 altavoces rodeando al oyente, o 3D con cada uno de los altavoces en los vértices de un cubo. Se cuenta para ello con 8 altavoces Neumann KH120A junto con una interfaz de audio Focusrite 18i20 controlada mediante Matlab.



Figura 3. Fotografía de la sala para la renderización de audio espacial en el Laboratorio de Acústica de la Universidad Pública de Navarra

Finalmente, la sala (ver Figura 3) donde se presentarán los paisajes sonoros presenta acústica variable por medio de paneles desmontables de tres materiales absorbentes diferentes: foam, espuma de melamina, y algodón regenerado aglomerado con resina termoendurente, junto con tres trampas de graves GIK acoustics. Además, mediante unas cortinas de alta densidad se puede delimitar un volumen de unos 50 m³ dentro de una sala de 120 m³. Todo ello permite diferentes configuraciones de sala con el rango de tiempos de reverberación mostrado en la Figura 4.

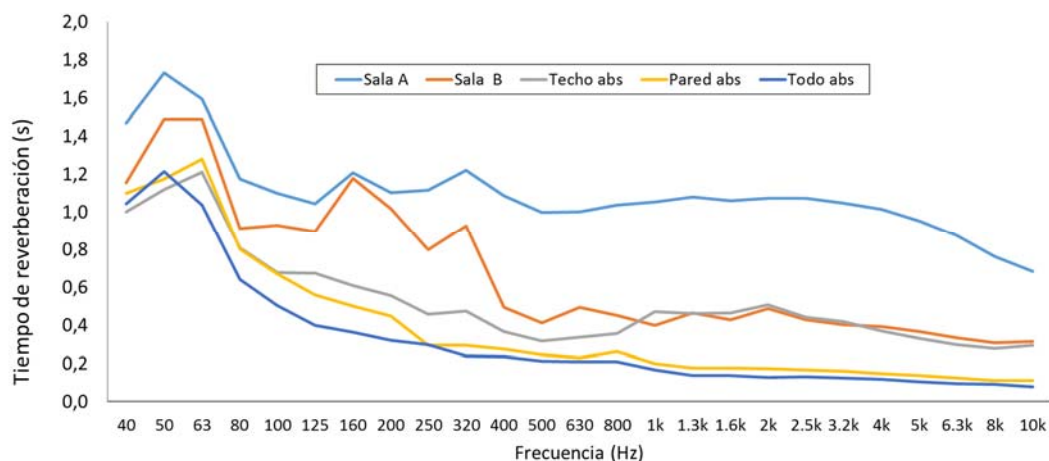


Figura 4. Tiempos de reverberación en bandas de 1/3 octava para diferentes configuraciones

4. APLICACIÓN PRÁCTICA: PARQUE FLUVIAL

El Parque Fluvial de Pamplona (Figura 5) es un paseo natural de recreo para viandantes y ciclistas de alto valor paisajístico que discurre a orillas de los ríos Arga, Elorz y Sadar. Cuenta

FIA 2018

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

con más de 17 kilómetros de longitud y recorre murallas, puentes históricos y pasarelas, parques, presas y molinos, granjas y huertas, merenderos, embarcaderos y espacios para la pesca. Su millón de metros cuadrados lo convierte en el gran pulmón verde de la ciudad de Pamplona.

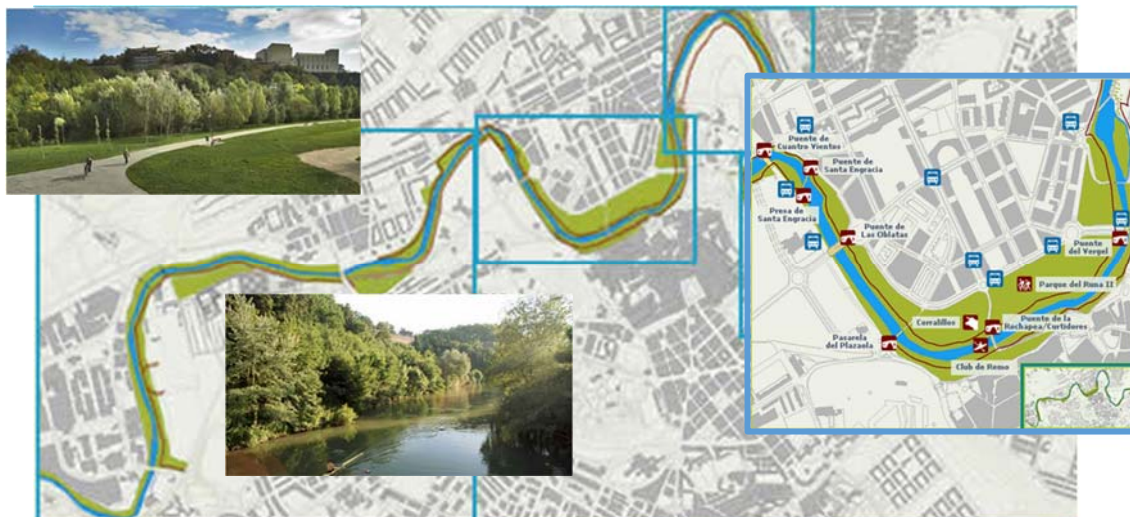


Figura 5. Parque fluvial de Pamplona

De cara a evaluar el procedimiento implementado, se registró en una mañana una muestra de 10 localizaciones representativas (Figura 6), procesando 30 segundos de grabación para cada una de ellas.



Figura 6. Dos de las localizaciones registradas. Puente de Santa Engracia (arriba) y Presa de la harinera (abajo).

Posteriormente, se realizó una valoración subjetiva en el laboratorio (Figura 7). Se complementó visualmente la presentación mediante unas gafas 3D Xiaomi Play 2VR. Además se comparó la calidad de la presentación multicanal con la de vídeos 3D en formato mp4 renderizados para una reproducción binaural clásica mediante auriculares y elaborados con FB360 spatial workstation.

FIA 2018

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



Figura 7. Presentación en el laboratorio de los paisajes sonoros grabados

Los resultados preliminares obtenidos fueron satisfactorios, mostrando una clara preferencia por la presentación multicanal frente a la binaural debido principalmente a un mayor grado de inmersión en el paisaje sonoro analizado. Sin embargo, es necesario un análisis pormenorizado del procedimiento completo, con el objetivo de minimizar la presencia de artefactos debidos al procesamiento de la señal que la técnica propuesta requiere.

5. CONCLUSIONES Y LÍNEAS FUTURAS

Se ha implementado una metodología para la grabación y presentación en entornos de laboratorio. El método utiliza como entrada señales B-format de primer orden obtenidas con un micrófono ambisonics. Con estas señales se extraen los parámetros “dirección de llegada, DOA”, y “grado de difusión”, que se emplean posteriormente en la renderización. Los valores de los parámetros direccionales se codifican con la resolución temporal y frecuencial de la percepción humana y luego se decodifican en el campo sonoro reproducido.

Aunque existen configuraciones de mayor complejidad, con el objetivo de realizar evaluaciones sistemáticas a gran escala se ha priorizado la portabilidad del dispositivo de registro, así como sistematizado el procesamiento de las señales para su presentación en un sistema multicanal. Las técnicas de grabación y reproducción juegan un papel crítico para lograr una alta validez ecológica de la investigación de paisajes sonoros realizada en un entorno acústico virtual. Los resultados preliminares obtenidos indican que merece la pena explorar más a fondo la validez ecológica de la fórmula implementada.

Aunque en el tema de los paisajes sonoros se está lejos de alcanzar el consenso científico sobre muchos aspectos (por ejemplo, indicadores acústicos o metodologías de evaluación) los entornos grabados y posteriormente reproducidos en un laboratorio pueden ser una herramienta muy eficaz. De esta manera se pueden controlar los estímulos presentados a los individuos, exponiendo a cada oyente exactamente al mismo estímulo. Esta capacidad de investigar la relación directa entre causa y efecto puede contribuir al desarrollo de hipótesis que sean validadas posteriormente in situ.

Asimismo, una recreación verosímil de paisajes sonoros (grabados o simulados) que permita su manipulación sería una poderosa herramienta en fases de planificación o diseño. El oyente podría contrastar en tiempo real sus interpretaciones con sus expectativas. Del mismo modo que la auralización se ha convertido ya en una herramienta habitual en el diseño de auditorios, las técnicas de grabación y presentación de paisajes sonoros pueden reducir la brecha existente entre investigadores y profesionales, permitiéndoles aplicar el concepto de paisaje sonoro en sus trabajos prácticos en un momento en que los descriptores e indicadores objetivos aún no se han establecido claramente.



FIA 2018

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

AGRADECIMIENTOS

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

REFERENCIAS

- [1] Jeon JY, Hong JY, Lee PJ. (2013) Soundwalk approach to identify urban soundscapes individually. *Journal of the Acoustical Society of America* 134(1):803-812
- [2] Aletta F, Kang J, Axelsson Ö. (2016) Soundscape descriptors and a conceptual framework for developing predictive soundscape models. *Landscape and Urban Planning* 149: 65–74.
- [3] Oberman T, Obad Šćitaroci B, Jambrošić K. (2018) Towards a virtual soundwalk. In *Handbook of Research on Perception-Driven Approaches to Urban Assessment and Design*; IGI Global: Hershey, PA, USA: 317–343.
- [4] Guastavino C, Katz BFG, Polack JD, Levitin DJ, Dubois D. (2005) Ecological validity of soundscape reproduction. *Acta Acustica united with Acustica* 91(2): 333-341
- [5] Davies WG, Bruce NS, Murphy JE. (2014) Soundscape reproduction and synthesis. *Acta Acustica united with Acustica* 100(2): 285-292
- [6] Hong JY, He J, Lam B, Gupta R, Gan WS. (2027) Spatial Audio for Soundscape design: recording and reproduction. *Applied Sciences* 7(6): 627
- [7] Axelsson Ö, Nilsson M, Berglund B. (2010) A principal components model of soundscape perception *Journal of the Acoustical Society of America*, 128(5):2836-2846
- [8] Pulkki V, Delikaris-Manias S, Politis A. (2018) *Parametric time-frequency domain spatial audio*. John Wiley & Sons.
- [9] Spors S, Wierstorf H, Raake A, Melchior F, Frank M, Zotter F. (2013) Spatial Sound With Loudspeakers and Its Perception: A Review of the Current State. *IEEE Proceedings* 101:1920-1938.
- [10] Ahrens J. (2012) *Analytic Methods of Sound Field Synthesis*, Springer-Verlag Berlin.
- [11] Stanzial D, Prodi N, Shiffner G. (1996) Reactive acoustic intensity for general fields and energy polarization. *Journal of the Acoustical Society of America* 99(4): 1868-1876
- [12] Merimaa J, Pulkki V. (2005) Spatial impulse response rendering 1: Analysis and synthesis. *Journal of the Audio Engineering Society* 52(12):115-1127
- [13] Solvang A. (2008) Spectral impairment of two-dimensional higher order Ambisonics. *Journal of the Audio Engineering Society* 56(4): 267-279
- [14] Vilkkamo J, Lokki T, Pulkki V. (2009) Directional audio coding: Virtual microphone-based synthesis and subjective evaluation. *Journal of the Audio Engineering Society* 57(9): 709-724