



Sociedad de Ingeniería de Audio

Artículo de Congreso

Congreso Latinoamericano de la AES 2011
30 de Agosto a 1º de Septiembre de 2011
Montevideo, Uruguay

Un generador de Paisajes Sonoros.

Marco Colasso Soca

Centro de Investigación en Tecnologías de Audio
Pontificia Universidad Católica de Chile
Santiago de Chile, 7550000, Chile
marco.colasso@gmail.com

RESUMEN

Quad Soundscape Generator es un generador de paisajes sonoros interactivo para interfaces tangibles. Se plantea la utilización de diferentes clasificaciones referentes a parámetros perceptivos del sonido y el entorno con el fin de llegar a un manejo holístico del sistema por parte del usuario, complementándose con técnicas de mapeo adaptativas que son controladas por un motor lógico que buscan alejarse del control racional analítico utilizado a diario en la creación y manipulación sonora.

0 INTRODUCCIÓN

Durante estos últimos años hemos presenciado un considerable aumento en la cantidad de dispositivos dedicados a la interacción hombre-maquina. Un ejemplo de esto son las interfaces tangibles¹, éstas han tenido una excelente aceptación en el mundo de la música y el sonido ya que proponen un modo directo e intuitivo para la interacción hombre-maquina[1].

Por otro lado, contamos con diferentes métodos para la clasificación y análisis del sonido que nos permiten mejorar notoriamente el manejo y entendimiento del mismo. De este modo, los sonidos pueden ser organizados a partir de: características acústicas, perceptuales, emocionales, dependiendo de cómo son generados, o finalmente conforme a su significado en un determinado contexto.

El presente artículo describe brevemente la construcción de una interfaz tangible basada en una mesa y se centra en el diseño sonoro para la composición asistida de paisajes sonoros², basado en una síntesis de diferentes propuestas para el análisis y la clasificación de los sonidos con el fin de llegar a un manejo holístico del sistema.

0.1- Diseño Sonoro.

Hunt y Kirk [2] hablan de la existencia de dos modos de pensamiento, dos modos cognitivos para una performance musical con instrumentos digitales: modo analítico y modo holístico. Pensamiento analítico es cuando el usuario analiza constantemente las opciones que entrega el computador, los parámetros se estudian uno a uno, en orden secuencial y lógico, contrastando con el modo holístico más indicado para tareas creativas, en donde el

usuario percibe el objeto como un todo, el conjunto es más importante que los pequeños detalles, priorizando el pensamiento global. De esta forma, varios "flujos de información" relacionados entre si pueden ser percibidos al mismo tiempo.

Alejarse de un modo de control analítico es el objetivo del diseño sonoro, buscando de esta forma fortalecer la idea de una interacción directa e intuitiva propuesta por este tipo de controladores[1], por un lado organizando los objetos tangibles con diferentes taxonomías que refieren a la percepción de los sonidos y el entorno (sección 3), y por otro buscando un control conjunto del ambiente generado entre el usuario y la maquina (sección 4).

0.2- Paisaje Sonoro.

Soundscape (Paisaje Sonoro en español), es definido como "Un ambiente sonoro que pone énfasis en el modo en la forma que es percibido y comprendido por el individuo o por una sociedad. El paisaje sonoro depende entonces de la relación entre la persona y cualquier entorno que lo rodee. El término puede referirse a ambientes reales o a construcciones abstractas, tales como composiciones musicales y montajes en cinta, especialmente cuando se lo considera como un ambiente artificial" [3].

La experimentación con este tipo de material sonoro, nos permite generar y manipular una amplia gama de sonidos, que van de lo concreto a lo electrónico, de sonidos naturales a sonidos artificiales. A pesar de que este proyecto tiene por finalidad la creación de paisajes para instalaciones interactivas cómo lo hicieron Barton y Priscila McLean en Rainforest Image [4], diferentes aplicaciones pueden beneficiarse esta disciplina como por ejemplo la creación de ambientes virtuales, videojuegos, realidad aumentada, proyectos como Tapestry [5] y Soundscapes Modeling [6] son ejemplos relacionados.

¹ Interfaz en la cual el usuario interactúa con la información digital a través del medio físico.

² Video prototipo: <http://vimeo.com/24359772>

1 CONTROLADOR

Al comenzar con este proyecto, se decidió optar por la utilización de una mesa tangible como controlador, ya que éste sistema basado en realidad aumentada, propone un modo de interacción directo e intuitivo entre el usuario y la información digital que no estaba presente en los diseños tradicionales basados en interfaces gráficas de usuario (GUI), ni en los controladores basados en faders y potenciómetros, donde perdemos demasiado tiempo recordando a donde están asignado y que es lo que hace cada parámetro [7][8].

El controlador consiste en una mesa multitáctil con la capacidad de reconocer forma, orientación y posición de objetos físicos aumentados digitalmente denominados objetos tangibles. Esta mesa (Figura 1) fue construida con una técnica llamada iluminación trasera difusa. Básicamente consta de un hexaedro que contiene en su interior leds infra-rojos capaces de iluminar la superficie de la forma más pareja posible. En este caso un material difusor es puesto en el medio de la superficie compuesta por dos vidrios, de modo que cuando un objeto toca la superficie éste refleja más luz que el difusor o los objetos que están en su fondo. Ésta luz es detectada por una cámara adaptada para captar solamente luz infra roja y finalmente la información es digitalizada para ser analizada por un tracker (software encargado del análisis topológico y detección de movimiento).

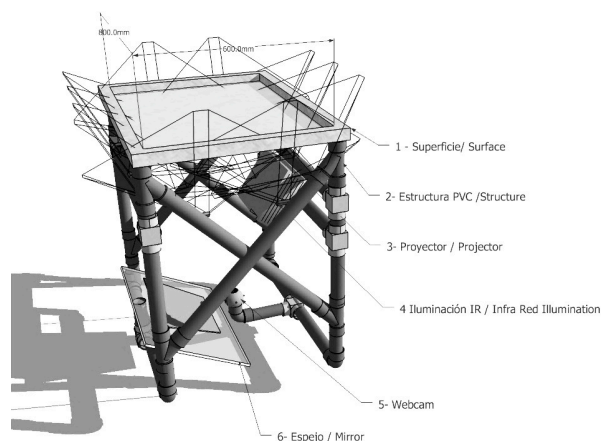


Figura 1 Modelado 3D del controlador construido que señala las diferentes partes del sistema.

El tracker entrega datos de identificación, posición y rotación en el caso de los objetos tangibles con marcadores fiduciales, que son comunicados a otra aplicación encargada de discriminar cual de los gestos realizados por el usuario son válidos, para la posterior generación del feedback visual y sonoro.

La construcción de esta interfaz fue posible gracias a los foros de la comunidad NUI group y el libro Multi-Touch Technologies [9].

2 SOFTWARE

Para esta primera versión del generador de paisajes sonoros, decidimos utilizar dos entornos de programación diferentes pero que tienen ventajas en común, son gratuitos y de código abierto, lo que ha contribuido a un valioso desarrollo por parte de comunidades a nivel mundial.

El reconocimiento de gestos, lógica, y feedback visual, utilizamos Processing [10], un lenguaje de programación

de texto especialmente diseñado para generar u modificar imágenes. La aplicación encargada del sonido se encuentra programada en Pure Data [11], un lenguaje de programación visual para audio y procesos gráficos a tiempo real. La comunicación entre programas se realiza por medio de los protocolos de comunicación TUIO y OSC.

Éstas aplicaciones son capaces de correr bajo cualquier sistema operativo, y cada día se encuentran más cerca de poder ser utilizados a través de cualquier browser desde Internet.

3 FAMILIAS DE OBJETOS TANGIBLES

Para lograr entender de forma general la elección de las taxonomías, agrupaciones de sonidos y funciones en familias de objetos tangibles, podemos pensar en una analogía entre los elementos utilizados para la composición del paisaje sonoro, con los que componen una imagen o paisaje visual. Los paisajes sonoros generados están compuestos por fondos (ambientes sonoros), figuras (eventos/objetos sonoros), y efectos que transforman las figuras (efectos sonoros).

3.1- Tangibles reproductores de ambientes sonoros.

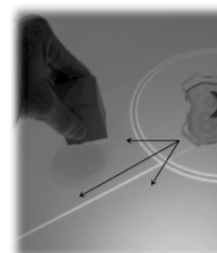
Esta primera familia de objetos tangibles son los encargados de reproducir ambientes sonoros previamente grabados, éstos pueden ser utilizados como el fondo del paisaje generado por el usuario.

Los ambientes han sido seleccionados de una biblioteca que hemos creado durante estos últimos cuatro años en base a grabaciones de ambientes de diferentes partes de Chile.

Las grabaciones de ambientes son clasificadas según la propuesta de R. Murray Schafer [12], quién presenta los extremos cómo paisajes HI-FI, y paisajes LO-FI. Los ambientes HI-FI son encontrados principalmente en zonas rurales, o lugares donde tenemos una buena relación señal-ruido, ésta nos permite distinguir los eventos sonoros con claridad discriminando la figura del fondo, minimizando las situaciones donde un sonido enmascara a otro. El LO-FI es todo lo contrario, y suele encontrarse en zonas urbanas con altos niveles de polución sonora.

Existen 2 objetos tipo LO-FI, uno encargado de ejecutar y controlar un reproductor de grabaciones estéreo, y otro para ejecutar grabaciones cuadrifónicas. También existen 2 objetos reproductores de ambientes clasificados como HI-FI con las mismas características.

Al momento de superponer las auras (círculo proyectado alrededor de los objetos cuando se encuentran activos) de un objeto LO-FI y un HI-FI, estos se dejan de escuchar para dar lugar a un tercer objeto llamado "LO to HI-FI". Estas son grabaciones de ambientes que pueden ser



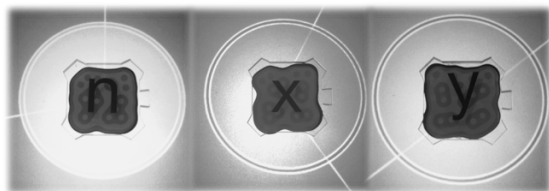
agrupadas por sus características de balance intermedio.

3.2- Tangibles que gatillan Eventos sonoros.

Un evento sonoro es definido cómo la partícula auto contenida más pequeña de un paisaje sonoro [12].

Los objetos tangibles encargados de reproducir y/o generar eventos sonoros están agrupados y marcados por letras que clasifican los sonidos según sus cualidades tipológicas. Basados en la clasificación propuesta por Pierre Schaeffer [13] ordenamos los sonidos dependiendo de su masa. Masa es definida como el modo en que el sonido ocupa el campo de las alturas, su espectro.

Masa Tónica: Sonido cuya altura es reconocible, definida, que podríamos reproducir, anotar o cantar. Su notación es la letra N.



Masa Compleja: Sonidos que no tienen esta posibilidad por carecer de una altura definida. Su notación es la letra X.

Masa Variable: Cuando el fenómeno perceptivo denominante es una combinación de masa tónica con masa compleja o posee una altura definida que varía en su altura. Su notación es la letra Y.

Existen cuatro objetos por cada categoría N, X e Y (Figura 2).

Figura 2 Podemos ver 3 de los 12 objetos tangibles de acrílico pertenecientes a la familia de eventos sonoros que son clasificados según su masa.

3.3- Satélites: Tangibles que transforman eventos en objetos sonoros.

Un objeto sonoro es “Todo fenómeno sonoro que se perciba como un conjunto, como un todo coherente, y que se oiga mediante una escucha reducida que lo enfoque por sí mismo, independientemente de su procedencia o de su significado” [14].

Los objetos satélites tienen la función de descomponer un evento sonoro en objetos sonoros, tanto en aspectos espectrales cómo en los espaciales. Cada evento puede ser descompuesto hasta en 3 objetos sonoros independientes. Esta función se realiza ingresando el objeto en el aura del evento, y desplazando el cubo en cualquier dirección (Figura 3). La familia de eventos/objetos sonoros son controlados físicamente por el usuario desplazando los objetos tangibles sobre la mesa, éstos gestos entregan datos que modifican diferentes parámetros como altura, textura, envolvente, dependiendo de cada evento. Un pequeño motor lógico explicado en la sección 4.3 es quien decide si en el momento en que el usuario coloca un evento sobre la mesa, éste gatillará un sampler o un sintetizador, que luego podrá ser descompuesto hasta por 3 satélites, estos sonidos son espacializados independientemente en una posición análoga del objeto en la mesa a la posición del sonido reproducido en la sala. Esto es gracias a un sistema codificador para formato ambisonics de 36 canales. La codificación puede ser de primer, segundo o tercer orden dependiendo del sistema multicanal dispuesto en la sala.

Figura 3 Se observa la transformación del evento en un objeto sonoro ingresando el satélite en el aura del evento, y desplazando el mismo en cualquier dirección.

3.4- Tangibles que controlan efectos

Para ser congruentes con la idea de controlar el sonido en base a parámetros perceptivos, ha sido difícil encontrar clasificaciones referidas a efectos que logren realmente potenciar las sensaciones transmitidas desde los paisajes.

Por lo tanto decidimos personalizar esta clasificación, y basados en una experiencia anterior [15] utilizamos nombres referidos a factores externos que afectan un ambiente o paisaje de una ciudad, cómo el stress, la rutina, el smog e información.

Esta familia tiene la particularidad de requerir una inyección de energía por parte del usuario para poder funcionar. Los parámetros de control cómo posición, velocidad y aceleración de movimiento controlan los niveles de envío de cada evento sonoro al canal de efecto, y modifican los parámetros de los efectos.

El efecto stress consiste en un delay granular. Pequeños granos de audio son grabados en un buffer, leídos y reordenados antes de ser enviados a la salida de delay. Parámetros cómo la dirección de reproducción, largo de la muestra y pitch de cada grano pueden ser controlados por el usuario de forma indirecta.

El efecto rutina consiste en un delay con altura variable, en otras palabras pitch delay. El tiempo de delay es constantemente modulado por osciladores internos, posee un controlador de feedback que distorsiona suavemente la señal, y un control de 6 octavas de variación tonal.

Smog consiste en una transformación de filtros (morphfilter). Este proceso es más conocido en video, donde se trata de generar una serie de imágenes que permitan transformar suavemente una imagen a otra. Esta idea es llevada al sonido buscando una transformación de un objeto sonoro (filtro 1) hacia los datos espectrales tomados del ambiente reproducido (filtro 2). Podemos experimentar transiciones entre estos dos filtros, controlando la profundidad e intensidad de filtrado, o incluso invertir el filtro 2 para obtener diferentes resultados.

Info+ es la abreviación de la palabra Información. Este efecto consiste en un simple Reverb que es alimentado dependiendo de la velocidad de movimiento y su cercanía a los eventos sonoros.

4 MÓDULOS

4.1 Transporte

El tiempo es determinado por un valor de BPM (beats por minuto) de entrada, y un parámetro llamado “porcentaje de swing”. Éste crea una pequeña desviación temporal calculada de forma random, que tiene la función de “humanizar” el beat, logrando que el tiempo no se perciba de forma tan rígida, tan perfecta.

4.2 Control de Transporte.

Cada 2048 samples se realiza un análisis espectral de los ambientes sonoros reproducidos, el resultado obtenido del promedio de niveles para bajas frecuencias determina de forma proporcional el valor del tiempo en BPM, el resultado promedio para frecuencias medias y altas determinará el porcentaje de swing utilizado. De esta forma cuando suena un ambiente LO-FI aumenta el tiempo y disminuye el swing ya que posee un mayor nivel en frecuencias graves. En cambio cuando el ambiente es HI-FI (mayor componente de medios y agudos), el tiempo disminuye y el porcentaje de swing es mayor. Con esta lógica conseguimos un tiempo rápido y determinado en

ambientes LO-FI (artificiales), más lento e indeterminado para paisajes HI-FI (naturales).

4.3 Motor lógico (Logic Engine).

Durante la programación de este software, surge la necesidad de realizar un pequeño motor lógico que se encargue de tomar decisiones a tiempo real sobre las características de los sonidos que se están reproduciendo, afectando y transformando el mapeo (forma en que los controles del usuario se conectan a las variables del sonido [16]) dependiendo de tres parámetros ingresados por el usuario antes de comenzar la interacción:

-Tiempo de duración de la interacción (length): Duración en minutos de la interacción.

-Cantidad de partes (parts): Sirve para calcular una señal de cambio llamada "cambio de parte", estos son los momentos en que se realizan cambios de mapeo durante la interacción. Sobre el tiempo total ingresado, se divide en x cantidad de partes iguales.

- Gráfica de Equilibrio vs Tiempo (equilibrio/time).

La Figura 3 muestra una gráfica determinada por el usuario que controla el nivel de equilibrio de la interacción en el tiempo. Cuanto más cercano al cero se encuentre el valor de equilibrio, el ambiente generado tenderá a ser equilibrado, según Barry Truax [17] un ambiente equilibrado (cercano a un HI-FI) es donde el receptor puede emitir sonido tanto como las fuentes que están alrededor (Escucho, entiendo, devuelvo), y cuanto más cercano al uno (parte superior) el ambiente tenderá al desbalance (cercano a un LO-FI), donde el receptor pierde su propia presencia, por la existencia de fuentes que enmascaran todo lo demás.

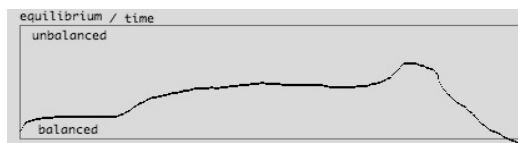


Figura 3 Gráfica ingresada por el usuario antes de comenzar la interacción que determina el equilibrio del ambiente en función del tiempo ingresado.

5 CONCLUSIONES Y TRABAJO FUTURO

Se explica y justifica el diseño de control para un software que permite generar y manipular paisajes sonoros a tiempo real.

La utilización de una perspectiva perceptiva y no tan técnica para el manejo de los sonidos, la implementación de un motor lógico que permita una ejecución y toma de decisiones conjunta por el usuario y el computador, y las variadas técnicas de mapeo incluidas en el proceso, son las principales apuestas para conseguir un manejo holístico del sistema.

En un futuro presentaremos la primera etapa de evaluación de resultados, pruebas con usuarios que nos permitirán determinar la efectividad de las taxonomías sonoras utilizadas e identificar aspectos susceptibles de mejora.

REFERENCIAS BIBLIOGRÁFICAS

[1] JORDÀ Sergi, GEIGER Günter, ALONSO Marcos, KALTENBRUNNER Martin. The reacTable: Exploring the Synergy between Live Music Performance and Tabletop Tangible Interfaces. Music Technology Group, Pompeu Fabra University, Barcelona, Spain 2007. 8p.

[2] HUNT Andy, KIRK Ross, Mapping Strategies for Musical Performance, University of York, U.K., 2000.

[3] TRUAX Barry, Handbook for acoustic ecology [en línea] Second Edition, Simon Fraser University, and ARC Publications. 1999 Disponible en <http://www.sfu.ca/sonic-studio/handbook/>

[4] CHADABE Joel, Electric Sound: The Past and Promise of Electronic Music, Prentice Hall, New Jersey, 1996. Páginas 330 y 331. 370 p.

[5] MISRA Ananya, COOK Perry, WANG Ge, TAPESTREA : Techniques And Paradigms for Expressive Synthesis, Transformation, and Rendering of Environmental Audio. <http://taps.cs.princeton.edu/>

[6] Kersten Stefan, Roma Gerard, Schirosa Mattia, Janer Jordi, Soundscape Modelling, Music Technology Group. Pompeu Fabra University, <http://www.mtg.upf.edu/technologies/soundscapes>

[7] BOVERMANN, Till. Tangible Auditory Interfaces, Combining Auditory Displays and Tangible Interfaces (Doctor en ciencias de la ingeniería)Universidad de Bielefeld. Alemania, Diciembre 2009. 202p.

[8] PATTEN James, RETCH Ben, ISHII Hiroshi, Audiopad, a tag based interface for musical performance. Tangible Media Group, MIT media Lab, Cambridge, Massachusetts, EEUU., 2002, 6p.

[9] NUI Group Authors, [Community Release] [Electrónico] Multi-Touch Technologies. 1era edición, 2009, 87p. <http://nuicode.com/projects/wiki-book/files>

[10] REAS Casey, FRY Ben, Processing: a programming handbook for visual designers and artists, The MIT Press Cambridge, Massachusetts London, England, 2007. 736p

[11] PURE DATA Portal. puredata.info/

[12] SCHAFER R. Murray. The Tuning of the World. Random House Inc, 1st edition, June 1977, 301p.

[13] SCHAEFFER R. Pierre. Solfege de L'Object Sonore Audio CD y Texto. 1998. 176p.

[14] CHION, Michel. El Sonido, música cine y literatura. Paidós Ibérica, Ediciones: 1999. 413p.

[15] Santiago Atmosférico, Pd Patch [En línea], 2006. <http://www.mcolasso.com/?p=11>

[16] CHADABE Joel, The Limitations of Mapping as a Structural Descriptive in Electronic Instruments. NIME proceedings, Dublin, Irlanda. 2002.

[17] TRUAX Barry. Acoustic Communication. by Ablex Publishing Corporation, Norwood, New Jersey. 1984. 244p.