

(12) SOLICITUD INTERNACIONAL PUBLICADA EN VIRTUD DEL TRATADO DE COOPERACIÓN EN MATERIA DE PATENTES (PCT)

(19) Organización Mundial de la Propiedad
 Intelectual
 Oficina internacional



(43) Fecha de publicación internacional
 23 de Julio de 2009 (23.07.2009)

PCT

(10) Número de Publicación Internacional
WO 2009/090281 A1

(51) Clasificación Internacional de Patentes:
 H04S 3/00 (2006.01)

(21) Número de la solicitud internacional:
 PCT/ES2008/070246

(22) Fecha de presentación internacional:
 30 de Diciembre de 2008 (30.12.2008)

(25) Idioma de presentación: español

(26) Idioma de publicación: español

(30) Datos relativos a la prioridad:
 P 200800112 17 de Enero de 2008 (17.01.2008) ES

(71) Solicitante (para todos los Estados designados salvo US): AURALIA EMOTIVE MEDIA SYSTEMS, S.L. [ES/ES]; Paseo Mikeletegui, 7, 20009-SAN SEBASTIÁN (GUIPÚZCOA) (ES).

(72) Inventor; e

(75) Inventor/Solicitante (para US solamente): PORTAS AR-RONDO, Ivan [ES/ES]; C/ Mantulene 5 3ª B, 20009 San Sebastian (ES).

(74) Mandatario: PONS ARIÑO, Angel; Glorieta Ruben Dario 4, 28010 Madrid (ES).

(81) Estados designados (a menos que se indique otra cosa, para toda clase de protección nacional admisible): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[Continúa en la página siguiente]

(54) Title: METHOD OF CONVERTING 5.1 SOUND FORMAT TO HYBRID BINAURAL FORMAT

(54) Título: PROCEDIMIENTO DE CONVERSIÓN DE FORMATO SONORO 5.1 A BINAURAL HIBRIDO

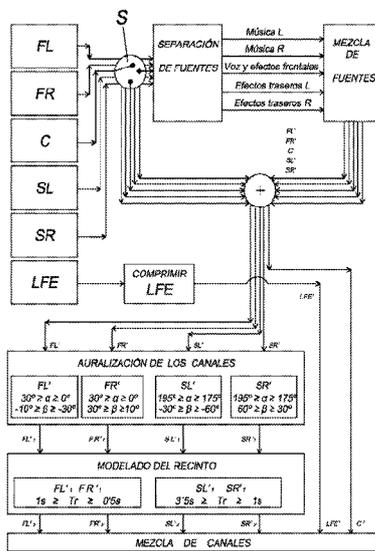


FIG. 3

SEPARACIÓN DE FUENTES SEPARATION OF SOURCES
 Música L Music L
 Música R Music R
 Vos y efectos frontales Voice and front sound
 Efectos traseros L Surround L
 Efectos traseros R Surround R
 MEZCLA DE FUENTES MIXING OF SOURCES
 COMPRIMIR LFE COMPRESS LFE
 AURALIZACIÓN DE LOS CANALES AURALIZATION OF CHANNELS
 MODELADO DEL RECINTO MODELLING OF ENCLOSURE
 MEZCLA DE CANALES MIXING OF CHANNELS

(57) Abstract: Method of converting 5.1 sound format to hybrid binaural format, comprising obtaining the signals from the FL, FR, C, SL, SR and LFE channels in 5.1 format which it is desired to convert into hybrid binaural format; auralizing the FL, FR, SL and SR channels in the following positions: FL: elevation from 0° to 30°, azimuth from -10° to -30°; FR: elevation from 0° to 30°, azimuth from +10° to +30°; SL: elevation from 175° to 195°, azimuth from -30° to -60°; SR: elevation from 175° to 195°, azimuth from +30° to +60°, thus obtaining the signals FL₁, FR₁, SL₁ and SR₁; modelling the response from the enclosure on the basis of the signals, introducing a reverberation effect; and mixing the signals FL₂, FR₂, SL₂ and SR₂ obtained in the previous step with the original LFE and C signals to obtain the two left and right output signals.

(57) Resumen: Procedimiento de conversión de formato sonoro 5.1 a binaural híbrido que comprende obtener las señales de los canales FL, FR, C, SL, SR y LFE del formato 5.1 que se desea convertir en formato binaural híbrido; auralizar los canales FL, FR, SL y SR en las siguientes posiciones; FL: elevación de 0° a 30°; azimut de -10° a -30°. FR: elevación de 0° a 30°; azimut de +10 a +30°. SL: elevación de 175° a 195°; azimut de -30° a -60°. SR: elevación de 175° a 195°; azimut de +30° a +60°, obteniéndose las señales FL₁, FR₁, SL₁ y SR₁; modelar la respuesta del recinto a partir de las señales introduciendo un efecto de reverberación; y mezclar las señales FL₂, FR₂, SL₂ y SR₂ obtenidas en la operación anterior junto con las señales originales LFE y C para obtener las dos señales de salida izquierda y derecha.

WO 2009/090281 A1



(84) Estados designados (*a menos que se indique otra cosa, para toda clase de protección regional admisible*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), euroasiática (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europea (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Publicada:

- *con informe de búsqueda internacional*
- *antes de la expiración del plazo para modificar las reivindicaciones y para ser republicada si se reciben modificaciones*

PROCEDIMIENTO DE CONVERSIÓN DE FORMATO SONORO 5.1 A
BINAURAL HÍBRIDO

DESCRIPCIÓN

5

OBJETO DE LA INVENCION

El objeto principal de la presente invención es un procedimiento para convertir sonido en formato sonoro 5.1, habitualmente utilizado para registro y reproducción sonora digital de contenido cinematográfico, en formato binaural híbrido.

10

ANTECEDENTES DE LA INVENCION

15

Actualmente, el formato 5.1 representa el estándar para la reproducción sonora doméstica de cine. Un sistema sonoro en formato 5.1 está compuesto por seis canales de audio donde se mezclan en distintas proporciones las señales de música, voz, efectos sonoros, etc. Cada uno de los canales corresponde a un altavoz, y a su vez cada uno de los altavoces debe estar situado en una ubicación concreta con relación al usuario para conseguir una sensación sonora óptima.

20

Los altavoces principales (FL y FR en la Figura 1) forman idealmente un triángulo equilátero con la posición del usuario (O). Además, las rectas formadas por los altavoces envolventes (SL y SR) y el usuario (O) forman un ángulo de aproximadamente 110° con respecto al eje vertical (recta que une O y C). El altavoz LFE (Low Frequency Enhancement, en sus siglas en inglés), tiene por objeto realzar los sonidos graves para producir un efecto impactante en la reproducción. Su ubicación no es determinante, ya que la información que transmite tiene un espectro de frecuencias generalmente menor de 100 Hz, que tiene una naturaleza omnidireccional. Es decir, no se puede determinar de donde proviene el sonido.

25

30

Un inconveniente de los sistemas de audio basados en el formato 5.1 es

que la sensación sonora del usuario se deteriora rápidamente cuando éste no está situado en la ubicación óptima con respecto de los altavoces. La utilización de auriculares permite, sin embargo, una colocación óptima del usuario en todo momento, ya que los sistemas de reproducción sonora, al estar adheridos a la cabeza del usuario, no modifican su posición relativa respecto a su cabeza.

Sin embargo, el ser humano es un receptor sonoro volumétrico, es decir, procesa el sonido que llega a él a través de, por ejemplo, reflexiones creadas por los hombros y el torso, o difracciones creadas por el sonido al rodear la cabeza. La audición humana es por naturaleza binaural, donde la resultante de todo el proceso de recepción sonora termina en dos únicos canales: oído derecho y oído izquierdo. El término "binaural" hace referencia a la naturaleza de la audición humana, debido a que las personas somos capaces de captar toda la información espacial sonora a través de un único par de oídos.

Cuando no se tiene en cuenta esta fenomenología se suele producir el denominado "sonido intracraneal", como por ejemplo al escuchar sonido estéreo tradicional a través de auriculares. El sonido intracraneal consiste en la sensación de que las fuentes sonoras se encuentran en el interior del cráneo del usuario, en un punto situado entre los dos auriculares, por lo que el sonido estéreo tradicional no es un formato aconsejable cuando se tratan de representar de forma realista espacios sonoros tridimensionales.

Existen fundamentalmente dos formas de lograr reproducciones binaurales:

La primera de ellas consiste en sustituir el par de receptores puntuales que se utiliza habitualmente por receptores volumétricos, como maniqués, logrando de ese modo que el sonido que llega a ellos se procese de forma natural. De este modo se logra una grabación estéreo binaural, donde queda ya introducida toda la fenomenología descrita anteriormente.

La segunda se basa en realizar un procedimiento de auralización. Para

ello, se suele medir o modelar la respuesta de un determinado receptor (un maniquí o un ser humano, por ejemplo) a una señal de impulso proveniente de un determinado punto del espacio (habitualmente un ruido de banda ancha emitido desde un determinado punto alrededor del usuario). La patente US 2007213990 describe un método para transformar una señal estéreo bicanal tradicional en una señal binaural, centrándose en el tratamiento que debe sufrir la señal de entrada de cara a su preparación para ser transformada en sonido tridimensional. Específicamente, se describe cómo dividir la señal de entrada según distintas bandas de frecuencia para, una vez dividida la señal de entrada, auralizar cada sub-banda y finalmente unir las para formar los dos canales de salida en formato binaural.

DESCRIPCIÓN DE LA INVENCION

La presente invención describe un nuevo procedimiento para la auralización en tiempo real de audio en formato 5.1. Para lograr un resultado óptimo, cada canal es tratado y auralizado independientemente, de modo que es posible asignar parámetros acústicos específicos a cada uno de ellos con el objetivo de hacer más realista y espectacular la reproducción.

Las ventajas más importantes del procedimiento de la invención se pueden resumir en las siguientes:

- Se consigue una reproducción óptima en todos los casos, ya que, al estar los auriculares adheridos al usuario, la posición relativa entre el sistema de reproducción y el usuario no varía.

- El modelo híbrido que se describe, que combina la auralización de los canales FL, FR, SL y SR con los canales monofónicos originales C y LFE permite una mayor inteligibilidad de los diálogos, al no existir interferencias entre los canales frontales y el canal C, así como una inmersión superior debido a la constante referenciación inconsciente que realiza el cerebro entre el canal C

monofónico y los canales auralizados.

- 5 - El reajuste de las proporciones de los diferentes tipos de información, mediante la separación de fuentes y posterior remezclado, permite optimizar desde el inicio el contenido de los distintos canales para lograr un resultado óptimo.
- 10 - La colocación virtual específica de los canales FL y FR, así como el modelado del recinto específico, permiten un perfecto equilibrio con el canal de diálogos C, no interfiriendo con su inteligibilidad y dotando de la profundidad justa al plano frontal.
- 15 - La colocación virtual específica de los canales SL y SR, así como el modelado de un recinto específico diferente para los canales de los planos delantero y trasero, aportan una sensación de profundidad trasera impactante, dotando al sistema de planos diferenciados de reproducción sonora, creando de este modo una experiencia altamente inmersiva.
- 20 - El refuerzo del canal LFE permite recrear las sensaciones producidas por las componentes graves en las salas de cine, equilibrando el sistema de reproducción.

25 En el presente documento, el término "auralizar" hace referencia al procesado de los diferentes canales para conseguir que el usuario tenga la impresión de que provienen de lugares del espacio concretos, consiguiéndose así una espectacularidad e inteligibilidad optimizadas.

30 Del mismo modo, el término "canal" hace referencia a la señal de cada uno de los altavoces que componen el formato de sonido 5.1 o el formato de sonido binaural híbrido. Así, hablaremos de los canales FL, FR, C, SL, SR o LFE, que son los canales de entrada en formato 5.1 y de los canales L y R, que son los canales de salida en formato binaural. Se utilizarán las letras "L" y "R" para

distinguir entre las posiciones de los canales situadas a la izquierda (left, en inglés) y derecha (right, en inglés) del usuario. También se utilizarán los términos “plano frontal” y “plano trasero” para hacer referencia a la posición de los canales frente al usuario o detrás del usuario, así como “plano lateral derecho” o “plano lateral izquierdo” para hacer referencia a la posición de los canales a los lados del usuario.

Por otro lado, el término “fuente” se refiere a una señal que contiene sonidos de un solo proceso físico, es decir, las fuentes serán, en general, música, voz y efectos.

Se define también el término “binaural híbrido” como un formato sonoro que mezcla canales auralizados con canales no auralizados o monofónicos. Concretamente, la presente invención mezcla los canales auralizados FL, FR, SL y SR con los canales no auralizados C y LFE.

De acuerdo con un aspecto de la presente invención, se describe un procedimiento de conversión de formato sonoro 5.1 a binaural híbrido, caracterizado porque comprende las siguientes operaciones:

1) Obtener las señales de los canales FL, FR, C, SL, SR y LFE del formato 5.1 que se desea convertir en formato binaural híbrido. La información que contienen estas señales es normalmente una mezcla de varias fuentes, donde:

FL: contiene principalmente música, y en menor medida voz y efectos.

FR: contiene principalmente música, y en menor medida voz y efectos.

C: contiene principalmente voz, y en menor medida música y efectos.

SL: contiene principalmente efectos, y en menor medida música.

SR: contiene principalmente efectos, y en menor medida música.

LFE: contiene únicamente graves.

2) Auralizar los canales FL, FR, SL y SR en las siguientes posiciones:

- FL: elevación de 0° a 30° ; azimut de -10° a -30° .
FR: elevación de 0° a 30° ; azimut de $+10^\circ$ a $+30^\circ$.
SL: elevación de 175° a 195° ; azimut de -30° a -60° .
5 SR: elevación de 175° a 195° ; azimut de $+30^\circ$ a $+60^\circ$.

obteniéndose como resultado las señales FL_1 , FR_1 , SL_1 y SR_1 .

10 Diremos que “auralizar” un canal en una posición determinada significa ubicar virtualmente ese canal de forma que la reproducción de las señales resultantes, una para el canal derecho y otra para el canal izquierdo, a través de unos auriculares producen la sensación en el usuario de que los sonidos de ese canal provienen de esa posición determinada del espacio.

15 Dicho de otro modo, auralizar es un proceso mediante el cual un canal carente de información espacial habitualmente monofónico, como en este caso, es decir, anecoico o seco, es procesado mediante un procedimiento llamado convolución, con la respuesta al impulso (respuesta en tiempo y frecuencia a un determinado estímulo acústico proveniente de un determinado punto del espacio)
20 de un determinado oyente.

 Sin embargo, debido a las diferencias físicas entre los diferentes usuarios (tamaño, distancia entre los oídos, etc.), no todos ellos responden de igual modo ante los nuevos canales FL_1 , FR_1 , SL_1 y SR_1 .

25 Para conocer la respuesta de cada tipo de usuario, se modela o se mide la respuesta de un determinado receptor (un maniquí o un ser humano por ejemplo) a una señal de impulso proveniente de un determinado punto del espacio (habitualmente ruido de banda ancha emitido desde un determinado
30 punto alrededor del usuario). Ésta respuesta al impulso del usuario se utiliza más tarde para procesar una fuente monofónica (sin información espacial) mediante un proceso de convolución, logrando así el efecto de escuchar dicha

fuente situada en el punto donde ha sido emitido el impulso.

Los inventores han descubierto que situar virtualmente los canales FL, FR, SL y SR dentro de los rangos angulares descritos anteriormente proporciona a todos los usuarios una sensación de espectacularidad óptima.

El motivo de que los rangos angulares de los altavoces frontales (FL y FR) no sean muy grandes es evitar la pérdida de inteligibilidad del canal de diálogos (C) debido a una imagen estéreo excesiva de la música, es decir, que la energía del canal FL vaya casi completamente a L y la energía de FR vaya casi completamente a R, y evitar la llegada de una gran cantidad de energía a los planos laterales, cerca de los oídos que interfiera en la localización de los canales del plano trasero (SL y SR).

El canal de diálogos (C) no se procesa en la operación de procesado de las señales de los canales FL, FR, SL y SR, ya que mantenerlo como fuente aporta dos grandes ventajas a la salida final del procedimiento.

La primera de ellas es ganar en inteligibilidad respecto al formato de entrada, ya que al mantener este canal intacto y auralizar los de los planos frontal (FL y FR) y trasero (SL y SR), los diálogos (C) quedan destacados en la posición central, reduciendo la fatiga auditiva para su seguimiento.

La segunda ventaja reside en el hecho de constituir un punto de referencia auditivo para el cerebro, ya que mantener su naturaleza intracraneal hace ideal su combinación con los canales auralizados. De este modo, el cerebro compara constantemente la posición de este canal con los auralizados, haciendo la experiencia auditiva del usuario mucho más espectacular.

El canal LFE tampoco se procesa en esta operación del procedimiento debido a la naturaleza no direccional de las frecuencias que contiene, es decir, da la sensación de ser escuchado en todas las posiciones. Esta característica hace

que los altavoces destinados a la reproducción de este canal puedan ser situados prácticamente en cualquier lugar del recinto.

- 3) Modelar respuestas de recinto independientes para los planos frontal y trasero.

Se procesan los canales del plano frontal (FL_1 , FR_1) y trasero (SL_1 , SR_1) independientemente utilizando dos respuestas a impulso de recintos optimizados diferentes. El procesado por separado de los canales frontales y traseros aporta la ventaja de utilizar dos recintos virtuales diferentes, dotando de mayor profundidad únicamente a los canales traseros, que son los que poseen efectos más espectaculares. Una profundidad excesiva en los canales delanteros, sin embargo, dificultaría la inteligibilidad de los diálogos.

De acuerdo con realizaciones preferidas de la presente invención, la reverberación introducida en los canales FL_1 y FR_1 está dentro del rango de 0,5 segundos a 1 segundo, y la reverberación introducida en los canales SL_1 y SR_1 está dentro del rango de 1 segundo a 3,5 segundos.

Así, después de la operación de modelar la respuesta del recinto se obtienen como salida las señales del plano frontal FL_2 y FR_2 , y las señales del plano trasero SL_2 y SR_2

- 4) Mezclar las señales obtenidas en la operación anterior junto con las señales LFE y C originales para obtener las señales de salida del canal izquierdo y el canal derecho (L y R).

De acuerdo con una realización preferida de la presente invención, el procedimiento de conversión de formato sonoro 5.1 a binaural híbrido, comprende, previamente a la operación final de mezcla, comprimir la señal del canal LFE, obteniéndose una señal LFE'.

Otra realización preferida de la invención comprende, previamente a la

operación de auralización, las operaciones de:

5 a) Separar las señales de los canales FL, FR, C, SL, SR en las fuentes que los componen música L, música R, voz y efectos frontales, efectos traseros L y efectos traseros R. La separación se realiza utilizando un algoritmo de análisis de componentes independientes. Este análisis realiza una comparación de las distintas entradas (canales) que contienen información redundante en diferentes proporciones. Partiendo de la teoría de que varias señales se pueden considerar independientes si proceden de procesos físicos distintos, se logra aislar las distintas componentes, que en este caso son voz, música y efectos.

10 b) Mezclar las fuentes música L, música R, voz y efectos frontales, efectos traseros L y efectos traseros R para obtener las señales que constituirán la entrada a la operación subsiguiente de auralización de los canales. Esta operación de mezcla reconstruye las señales FL, FR, C, SL y SR con las proporciones óptimas de las fuentes que se separaron en la operación anterior.

20 De acuerdo con una realización preferida de la presente invención, la mezcla de las fuentes música L, música R, voz y efectos frontales, efectos traseros L y efectos traseros R para obtener los canales se realiza según los siguientes rangos porcentuales:

25 FL: 70-90% música L, 30-10% voz y efectos frontales
FR: 70-90% música R, 30-10% voz y efectos frontales
C: 70-90% voz y efectos frontales, 30-10% música L y R
SL: 70-90% efectos traseros L, 30-10% música L
SR: 70-90% efectos traseros R, 30-10% música R

30 El objetivo de estas dos operaciones opcionales es garantizar que cada canal en el proceso de auralización contenga la proporción adecuada de las distintas componentes, ya que la mezcla original en 5.1 estaba optimizada para su reproducción a través de 6 altavoces físicos, esquema completamente distinto a

un par de auriculares. A la hora de reproducir en auriculares, la información redundante característica de sistemas cuadrafónicos como el 5.1 obstaculiza la percepción de realismo espacial, y por eso es necesario este reajuste.

5 El canal de frecuencias graves LFE ya es en sí mismo una componente independiente, y por tanto su información no es redundante en los demás canales. Por este motivo no se incluye en las operaciones iniciales opcionales de separación y mezcla.

10 De acuerdo con otro aspecto de la invención, ésta se extiende también a programas de ordenador, en particular programas de ordenador en contenidos en una portadora, adaptados para llevar a cabo las operaciones del procedimiento descrito. El programa puede estar en forma de código fuente, código objeto o un código intermedio entre el código fuente y el código objeto, como una forma
15 parcialmente compilada, o de cualquier otra forma adecuada para implementar las operaciones de la invención.

La portadora puede ser cualquier dispositivo o entidad capaz de transportar el programa. Por ejemplo, la portadora puede comprender un medio de
20 almacenamiento, como una ROM, un CD ROM o cualquier otro medio de almacenamiento magnético, por ejemplo un disquete o un disco duro. Además, la portadora puede ser una portadora de transmisión, como una señal eléctrica u óptica que se pueda comunicar a través de cable eléctrico, óptico, por radio o de cualquier otro modo.

25 Alternativamente, la portadora puede ser un circuito integrado en el que está almacenado el programa, estando el circuito adaptado para efectuar las operaciones del procedimiento. En particular, podría ser un ASIC, una FPGA, un DSP, un microprocesador o un microcontrolador.

30

DESCRIPCIÓN DE LOS DIBUJOS

Para complementar la descripción que se está realizando y con objeto de ayudar a una mejor comprensión de las características de la invención, de acuerdo con un ejemplo preferente de realización práctica de la misma, se acompaña como parte integrante de dicha descripción, un juego de dibujos en donde con carácter ilustrativo y no limitativo, se ha representado lo siguiente:

Figura 1.- Muestra una vista de la ubicación de los altavoces físicos de un cine en un formato sonoro 5.1.

10

Figura 2.- Muestra un esquema explicativo de la posición de los ángulos de elevación (α) y de azimut (β).

Figura 3.- Muestra un esquema general de las operaciones del procedimiento de acuerdo con la presente invención.

15

REALIZACIÓN PREFERENTE DE LA INVENCION

Se parte del sonido original de una película en formato 5.1 que se desea convertir en binaural híbrido, que en este caso está grabado en un disco de tipo DVD. La Figura 1 muestra la posición de los altavoces de los canales en una sala de cine con relación a la posición en la que debe estar situado el usuario para una experiencia sonora óptima.

20

En este ejemplo, el procedimiento lo efectúa un ordenador que, en primer lugar, como se muestra en la Figura 3, obtiene a partir del DVD las señales de los canales originales en formato 5.1 (FL, FR, C, SL, SR, LFE). El canal LFE es separado para ser procesado en paralelo de modo independiente, sufriendo únicamente una compresión que da como resultado la señal LFE'.

25

30

En este ejemplo, se proporciona un selector (S) que permite que el usuario

seleccione o no las operaciones opcionales de extraer las fuentes de los canales originales y remezclarlas de acuerdo con unas proporciones nuevas para realzar la espectacularidad de la película. Para ello, se separan la fuentes (música L, música R, voz y efectos frontales, efectos traseros L y efectos traseros R), por ejemplo utilizando el algoritmo de separación de fuentes por análisis de componentes independientes 'FastICA', desarrollado por el HUT (Helsinki University of Technology), para volver a mezclarlos de acuerdo con unas proporciones nuevas optimizadas. En este ejemplo supondremos que la película es de acción, lo cual implica la existencia de una serie de características sonoras, como explosiones, tiros, ruido de motores, etc. Para conseguir la mayor espectacularidad posible en este tipo de películas, se han determinado las siguientes proporciones óptimas de mezcla:

FL': 80% música L + 20% voz y efectos frontales
FR': 80% música R + 20% voz y efectos frontales
C': 80% voz y efectos frontales + 20% música L y R
SL': 80% efectos traseros L + 20% música L
SR': 80% efectos traseros R + 20% música R

Una vez mezcladas las fuentes en los canales de este modo optimizado, se separa el canal de diálogos (C') del resto, los canales FL', FR', SL' y SR' son auralizados cada uno de ellos en una situación geométrica óptima para realzar la espectacularidad de la experiencia sonora del usuario. En este caso, se ha considerado que el oyente tiene las características de un usuario estándar basado en las respuestas al impulso de un maniquí Kemar.

A continuación se presentan las posiciones óptimas de los canales, descritas a través del ángulo de elevación (α) y del ángulo de azimut (β) que forman con el oyente:

FL': elevación 15°; azimut -20°
FR': elevación 15°; azimut 20°

SL': elevación 180°; azimut -40°

SR': elevación 180°; azimut 40°

La Figura 2 muestra la referencia de la ubicación de los ángulos de elevación y azimut, respectivamente α y β . Después de la operación de auralización, se obtienen las señales FL'₁, FR'₁, SL'₁ y SR'₁. A continuación, se procesan las señales FL'₁ y FR'₁ con la respuesta al impulso de un recinto similar a una sala de cine, con un tiempo de reverberación (T_r) de 0,5 segundos aproximadamente; y las señales SL'₁ y SR'₁ con la respuesta al impulso de otro recinto similar a una sala de cine diferente, con un tiempo de reverberación de 2 segundos aproximadamente.

Finalmente, se mezclan los canales obtenidos en la operación anterior, FL'₂, FR'₂, SL'₂ y SR'₂ con los canales LFE' y C' para obtener únicamente dos señales en formato binaural híbrido correspondientes a los canales L y R de unos auriculares.

REIVINDICACIONES

1. Procedimiento de conversión de formato sonoro 5.1 a binaural híbrido, caracterizado porque comprende las siguientes operaciones

5

obtener las señales de los canales FL, FR, C, SL, SR y LFE del formato 5.1 que se desea convertir en formato binaural híbrido;

auralizar los canales FL, FR, SL y SR en las siguientes posiciones:

10

FL: elevación de 0° a 30° ; azimut de -10° a -30° .
FR: elevación de 0° a 30° ; azimut de $+10^{\circ}$ a $+30^{\circ}$.
SL: elevación de 175° a 195° ; azimut de -30° a -60° .
SR: elevación de 175° a 195° ; azimut de $+30^{\circ}$ a $+60^{\circ}$,

15

obteniéndose como resultado las señales FL_1 , FR_1 , SL_1 y SR_1 ;

procesar independientemente las señales del plano frontal (FL_1 y FR_1) y las del plano trasero (SL_1 y SR_1), utilizando para ello, las respuestas a impulso de dos recintos virtuales diferentes, optimizados cada uno de ellos para dichos planos, obteniéndose como resultado las señales FL_2 , FR_2 , SL_2 y SR_2 ;

20

mezclar las señales FL_2 , FR_2 , SL_2 y SR_2 obtenidas en la operación anterior junto con las señales originales LFE y C para obtener las dos señales de salida izquierda y derecha.

25

2. Procedimiento de conversión de formato sonoro 5.1 a binaural híbrido de acuerdo con la reivindicación anterior, caracterizado porque las respuestas a impulso de los recintos virtuales utilizados para el procesado del plano frontal y trasero, comprenden tiempos de reverberación de entre 0,5 s y 1 s para el primero, y de entre 1 s y 3,5 s para el segundo.

30

3. Procedimiento de conversión de formato sonoro 5.1 a binaural híbrido de acuerdo con cualquiera de las reivindicaciones anteriores, caracterizado porque comprende, previamente a la operación final de mezcla, una compresión del canal LFE.
- 5
4. Procedimiento de conversión de formato sonoro 5.1 a binaural híbrido de acuerdo con cualquiera de las reivindicaciones anteriores, caracterizado porque antes de la operación de auralización comprende las operaciones de:
- 10 separar las señales de los canales FL, FR, C, SL, SR en las fuentes que los componen: música L, música R, voz y efectos frontales, efectos traseros L y efectos traseros R;
- remezclar las fuentes estimadas en proporciones optimizadas para
- 15 procesos posteriores, reconstruyendo los canales FL, FR, C, SL y SR.
5. Procedimiento de conversión de formato sonoro 5.1 a binaural híbrido de acuerdo con la reivindicación anterior, caracterizado porque la operación de remezcla de las fuentes música L, música R, voz y efectos frontales, efectos
- 20 traseros L y efectos traseros R se realiza de acuerdo con los siguientes rangos porcentuales:
- FL: 70-90% música L, 30-10% voz y efectos frontales
- FR: 70-90% música R, 30-10% voz y efectos frontales
- 25 C: 70-90% voz y efectos frontales, 30-10% música L y R
- SL: 70-90% efectos traseros L, 30-10% música L
- SR: 70-90% efectos traseros R, 30-10% música R
6. Procedimiento de conversión de formato sonoro 5.1 a binaural híbrido de
- 30 acuerdo con la reivindicación anterior, caracterizado porque es llevado a cabo por uno dispositivo de entre los de la siguiente lista: un ASIC, una FPGA, un DSP, un microprocesador y un microcontrolador.

7. Programa de ordenador que comprende instrucciones de programa que provocan que un ordenador lleve a cabo las operaciones del método de acuerdo con cualquiera de las reivindicaciones anteriores.

5 8. Programa de ordenador de acuerdo con la reivindicación 7, caracterizado porque está almacenado en unos medios de almacenamiento.

9. Programa de ordenador de acuerdo con la reivindicación 7, caracterizado porque se transmite a través de una señal portadora.

10

15

20

25

30

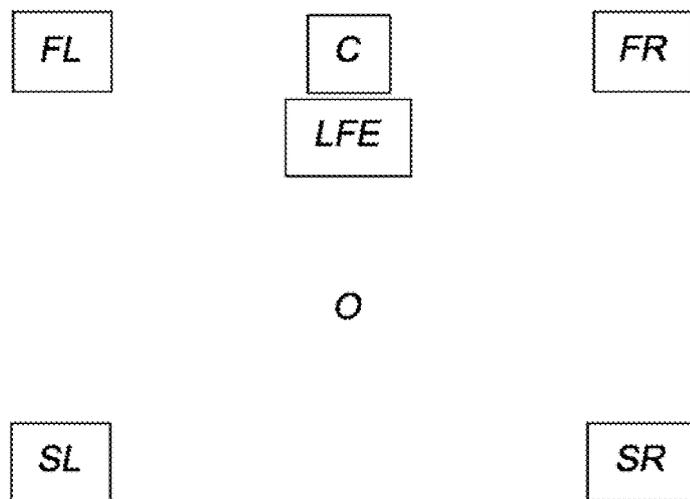


FIG. 1

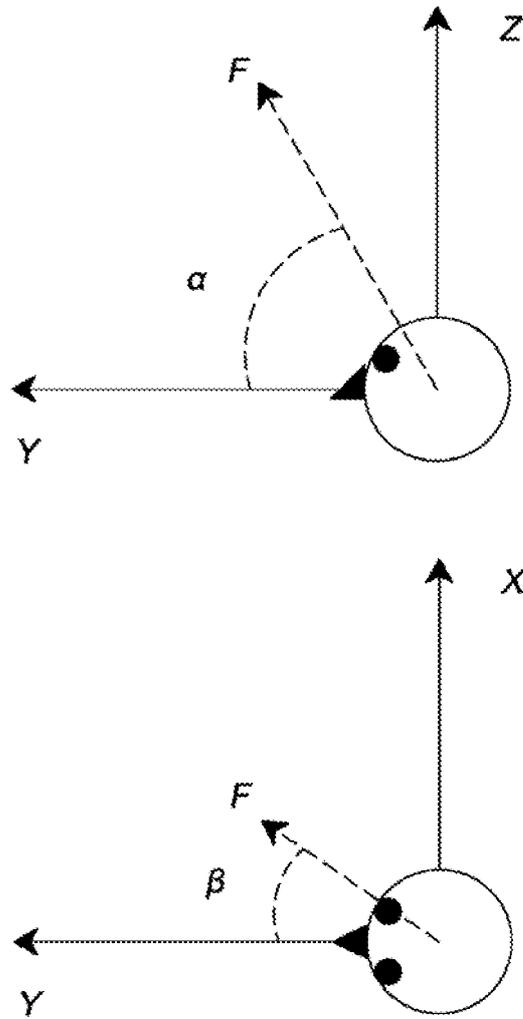


FIG. 2

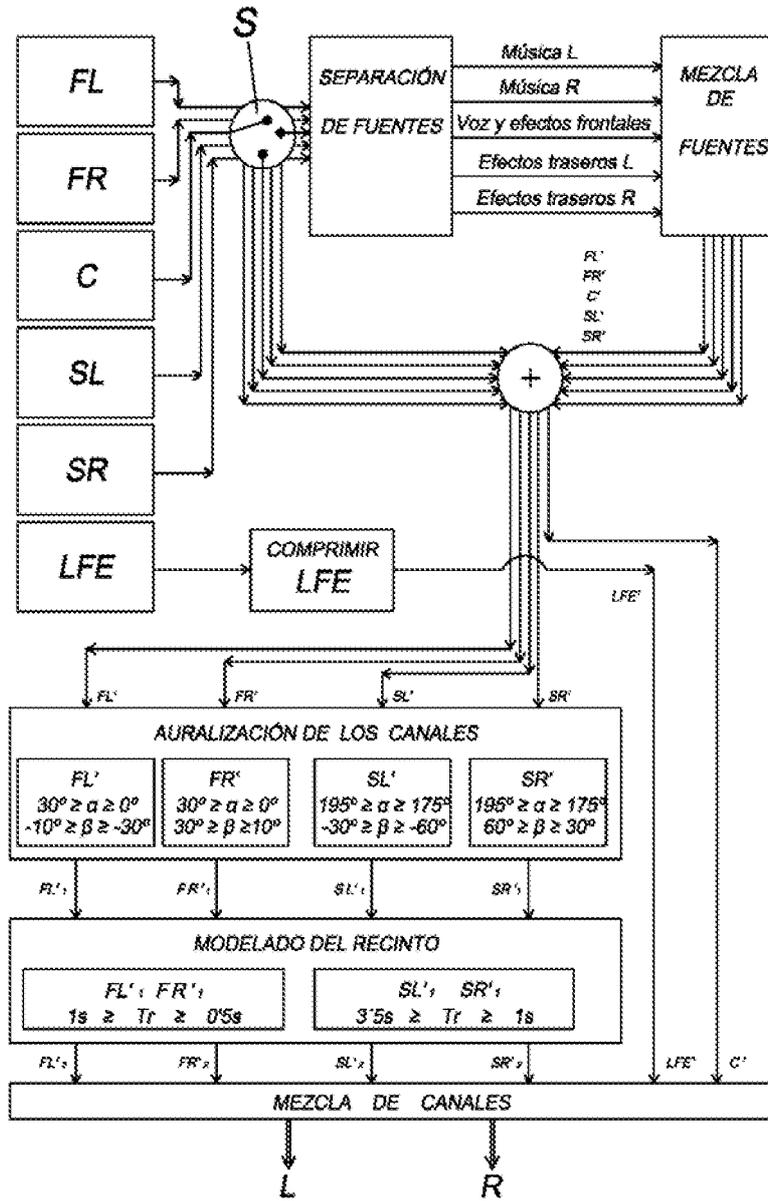


FIG. 3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/ ES 2008/070246

A. CLASSIFICATION OF SUBJECT MATTER

H04S 3/00 (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04S

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

INVENES, EPODOC, WPI, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2007123788 A2 (SRS LABS INC ; WANG WEN) 01.11.2007, paragraphs [0029-0058]; paragraph [0075]; figure 3, figure 6,	1-3,7-9
A		4-6
Y	US 6002775 A (WOOD et al.) 14.12.1999, the whole document.	1-3,7-9
Y	US 5742689 A (TUCKER et al.) 21.04.1998, column 1, line 14 - column 7, line 9; figure 4,	1-3,7-9
A		4-6
Y	EP 1816890 A1 (SONY CORP) 08.08.2007, figure 19, paragraphs [0154-0173];	1-3,7-9

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance.</p> <p>“E” earlier document but published on or after the international filing date</p> <p>“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>“O” document referring to an oral disclosure use, exhibition, or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other documents, such combination being obvious to a person skilled in the art</p> <p>“&” document member of the same patent family</p>
---	--

Date of the actual completion of the international search

11.May.2009 (11.05.2009)

Date of mailing of the international search report

(14/05/2009)

Name and mailing address of the ISA/
O.E.P.M.

Paseo de la Castellana, 75 28071 Madrid, España.
Facsimile No. 34 91 3495304

Authorized officer

M. Rivas Sáiz

Telephone No. +34 91 349 85 95

INTERNATIONAL SEARCH REPORT

International application No.

PCT/ES 2008/070246

C (continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of documents, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Technologies for Presentation of Sorround-Sound in Headphones.17.12.2007 [retrieved on 23.03.2009] Retrieved from the Internet .< http://www.headwize.com/tech/sshd_tech.htm >; < http://web.archive.org/web/20071212180538/http://www.headwize.com/tech/sshd_tech.htm >. the whole documetno.	1-9
A	Ciaramel the A. "BSS toolbox for delayed and convolved mixtures". Neural Networks, 2005. Proceedings. 2005 IEEE International Joint Conference on Montreal, Que., Canada. Vol 2, pages 1245 - 1250. the whole documetno.	4-6

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/ ES 2008/070246

Patent document cited in the search report	Publication date	Patent family member(s)	Publication date
WO 2007123788 A	01.11.2007	US 2007230725 A EP 2005787 A EP 20070754557 KR 20090007700 A	04.10.2007 24.12.2008 03.04.2007 20.01.2009
US 6002775 A	14.12.1999	WO 9833356 A AU 5826898 A TW 391149 B	30.07.1998 18.08.1998 21.05.2000 21.05.2000
US 5742689 A	21.04.1998	WO 9725834 A AU 1527197 A	17.07.1997 01.08.1997
EP 1816890 A	08.08.2007	JP 2006345480 A KR 20070079301 A CN 101014210 A US 2007183617 A	21.12.2006 06.08.2007 08.08.2007 09.08.2007 09.08.2007

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional N°
PCT/ ES 2008/070246

A. CLASIFICACIÓN DEL OBJETO DE LA SOLICITUD

H04S 3/00 (2006.01)

De acuerdo con la Clasificación Internacional de Patentes (CIP) o según la clasificación nacional y CIP.

B. SECTORES COMPRENDIDOS POR LA BÚSQUEDA

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)
H04S

Otra documentación consultada, además de la documentación mínima, en la medida en que tales documentos formen parte de los sectores comprendidos por la búsqueda

Bases de datos electrónicas consultadas durante la búsqueda internacional (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC, WPI, INSPEC

C. DOCUMENTOS CONSIDERADOS RELEVANTES

Categoría*	Documentos citados, con indicación, si procede, de las partes relevantes	Relevante para las reivindicaciones N°
Y	WO 2007123788 A2 (SRS LABS INC ; WANG WEN) 01.11.2007, párrafos [0029-0058]; párrafo [0075]; figura 3, figura 6,	1-3,7-9
A		4-6
Y	US 6002775 A (WOOD et al.) 14.12.1999, todo el documento.	1-3,7-9
Y	US 5742689 A (TUCKER et al.) 21.04.1998, columna 1, línea 14 - columna 7, línea 9; figura 4,	1-3,7-9
A		4-6
Y	EP 1816890 A1 (SONY CORP) 08.08.2007, figura 19, párrafos [0154-0173];	1-3,7-9

En la continuación del Recuadro C se relacionan otros documentos Los documentos de familias de patentes se indican en el Anexo

<p>* Categorías especiales de documentos citados:</p> <p>“A” documento que define el estado general de la técnica no considerado como particularmente relevante.</p> <p>“E” solicitud de patente o patente anterior pero publicada en la fecha de presentación internacional o en fecha posterior.</p> <p>“L” documento que puede plantear dudas sobre una reivindicación de prioridad o que se cita para determinar la fecha de publicación de otra cita o por una razón especial (como la indicada).</p> <p>“O” documento que se refiere a una divulgación oral, a una utilización, a una exposición o a cualquier otro medio.</p> <p>“P” documento publicado antes de la fecha de presentación internacional pero con posterioridad a la fecha de prioridad reivindicada.</p>	<p>“T” documento ulterior publicado con posterioridad a la fecha de presentación internacional o de prioridad que no pertenece al estado de la técnica pertinente pero que se cita por permitir la comprensión del principio o teoría que constituye la base de la invención.</p> <p>“X” documento particularmente relevante; la invención reivindicada no puede considerarse nueva o que implique una actividad inventiva por referencia al documento aisladamente considerado.</p> <p>“Y” documento particularmente relevante; la invención reivindicada no puede considerarse que implique una actividad inventiva cuando el documento se asocia a otro u otros documentos de la misma naturaleza, cuya combinación resulta evidente para un experto en la materia.</p> <p>“&” documento que forma parte de la misma familia de patentes.</p>
--	--

Fecha en que se ha concluido efectivamente la búsqueda internacional. 11.Mayo.2009 (11.05.2009)	Fecha de expedición del informe de búsqueda internacional 14 de Mayo de 2009 (14/05/2009)
---	---

Nombre y dirección postal de la Administración encargada de la búsqueda internacional O.E.P.M. Paseo de la Castellana, 75 28071 Madrid, España. N° de fax 34 91 3495304	Funcionario autorizado M. Rivas Sáiz N° de teléfono +34 91 349 85 95
--	---

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional N°

PCT/ES 2008/070246

C (continuación). DOCUMENTOS CONSIDERADOS RELEVANTES		
Categoría*	Documentos citados, con indicación, si procede, de las partes relevantes	Relevante para las reivindicaciones N°
A	Technologies for Presentation of Sorround-Sound in Headphones.17.12.2007 [recuperado el 23.03.2009] Recuperado de Internet .< http://www.headwize.com/tech/sshd_tech.htm >; < http://web.archive.org/web/20071212180538/http://www.headwize.com/tech/sshd_tech.htm >. Todo el documetno.	1-9
A	Ciaramella A. "BSS toolbox for delayed and convolved mixtures". Neural Networks, 2005. Proceedings. 2005 IEEE International Joint Conference on Montreal, Que., Canada. Vol 2, páginas 1245 - 1250. Todo el documetno.	4-6

INFORME DE BÚSQUEDA INTERNACIONAL

Información relativa a miembros de familias de patentes

Solicitud internacional N°

PCT/ES 2008/070246

Documento de patente citado en el informe de búsqueda	Fecha de Publicación	Miembro(s) de la familia de patentes	Fecha de Publicación
WO 2007123788 A	01.11.2007	US 2007230725 A EP 2005787 A EP 20070754557 KR 20090007700 A	04.10.2007 24.12.2008 03.04.2007 20.01.2009
US 6002775 A	14.12.1999	WO 9833356 A AU 5826898 A TW 391149 B	30.07.1998 18.08.1998 21.05.2000 21.05.2000
US 5742689 A	21.04.1998	WO 9725834 A AU 1527197 A	17.07.1997 01.08.1997
EP 1816890 A	08.08.2007	JP 2006345480 A KR 20070079301 A CN 101014210 A US 2007183617 A	21.12.2006 06.08.2007 08.08.2007 09.08.2007 09.08.2007