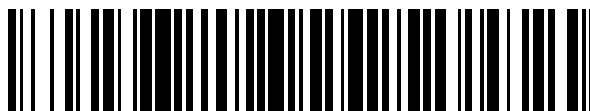


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 690 164**

51 Int. Cl.:

H04S 3/00 (2006.01)

H04R 3/12 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.06.2010 E 10167042 (0)**

97 Fecha y número de publicación de la concesión europea: **08.08.2018 EP 2285139**

54 Título: **Dispositivo y método para convertir una señal de audio espacial**

30 Prioridad:

08.01.2010 NO 20100031

25.06.2009 EP 09163760

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

19.11.2018

73 Titular/es:

**DTS LICENSING LIMITED (100.0%)
Hamilton House 2, National Technology Park
Castletroy Limerick, IE**

72 Inventor/es:

BERGE, SVEIN

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 690 164 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Dispositivo y método para convertir una señal de audio espacial

5 Campo de la invención

La invención se refiere al campo de procesamiento de señales de audio. Más específicamente, la invención proporciona un procesador y un método para convertir una señal de audio multicanal, tal como una señal de campo de sonido de formato B, en otro tipo de señal de audio multicanal adecuada para la reproducción a través de auriculares o altavoces, conservando al mismo tiempo información espacial en la señal original.

Antecedentes de la invención

El uso de mediciones, grabaciones y la reproducción de formato B en la provisión de reproducciones acústicas más ideal que capturan parte de las características espaciales de una reproducción de audio son bien conocidos.

En el caso de la conversión de señales de formato B a múltiples altavoces en una matriz de altavoces, no es un problema bien reconocido debido a la difusión de fuentes de sonido virtuales individuales durante un gran número de elementos de altavoz de reproducción. En el caso de la reproducción binaural de señales de formato B, las aproximaciones inherentes en el campo de sonido de formato B pueden llevar a una localización menos precisa de las fuentes de sonido y a una pérdida de la sensación de extracorriente que es una parte importante de la experiencia de reproducción binaural.

El documento US 6.259.795 de Lake DSP Pty Ltd. describe un método para aplicar HRTF a una señal de formato B que es particularmente eficaz cuando la señal está destinada a ser distribuida a varios oyentes que requieren diferentes rotaciones de la escena auditiva. Sin embargo, esa invención no aborda cuestiones relacionadas con la precisión de la localización u otros aspectos de la calidad de la reproducción del sonido.

El documento WO 00/19415 de Creative Technology Ltd. aborda el problema de la calidad de reproducción del sonido y propone mejorarlo utilizando dos señales de formato B separadas, una asociada con cada oreja. Esa invención no introduce tecnología aplicable al caso en el que solo está disponible una señal de formato B.

El documento US 6.628.787 de Lake Technology Ltd. describe un método específico para la creación de una señal multicanal o binaural a partir de una señal de campo de sonido de formato B. La señal del campo de sonido se divide en bandas de frecuencia, y en cada banda se determina un factor de dirección. En función del factor de dirección, las señales de los altavoces se calculan para cada banda haciendo un barrido de las señales para controlar los altavoces más cercanos. Además, los componentes de la señal residual se distribuyen a las señales del hablante por medio de técnicas de decodificación conocidas.

El problema con estos métodos es que la estimación de dirección es generalmente incorrecta en el caso en que más de una única fuente de sonido emita un sonido al mismo tiempo y dentro de la misma banda de frecuencia. Esto lleva a una localización imprecisa o incorrecta cuando hay más de una fuente de sonido presente y cuando los ecos interfieren con el sonido directo de una sola fuente.

45 Sumario de la invención

En vista de lo anterior, se puede ver como un objetivo de la presente invención es proporcionar un procesador y un método para convertir una entrada de audio multicanal, tal como una entrada de campo de sonido de formato B en una salida de audio adecuada para la reproducción a través de auriculares o por altavoces, conservando al mismo tiempo la información espacial sustancial contenida en la entrada multicanal original.

En un primer aspecto, la invención proporciona un procesador de audio dispuesto para convertir una señal de entrada de audio multicanal, tal como una señal de campo de sonido de formato B de tres o de cuatro canales, en un conjunto de señales de salida de audio, tal como un conjunto de dos señales de salida de audio dispuestas para auriculares o dos o más señales de salida de audio dispuestas para su reproducción a través de una matriz de altavoces. El procesador de audio está dispuesto para realizar un cálculo de descomposición de onda de plan paramétrico en la señal de entrada de audio multicanal como se define en la reivindicación 1 adjunta.

Este procesador de audio proporciona una conversión ventajosa de la señal de entrada de varios canales debido a la combinación extracción de descomposición de onda plana paramétrica de direcciones para fuentes de sonido dominantes para cada banda de frecuencia y la selección de al menos una posición de altavoz virtual coincidiendo con una dirección para al menos una fuente de sonido dominante.

Por ejemplo, esto proporciona una señal de altavoz virtual muy adecuada para la generación de una señal de salida binaural mediante la aplicación de funciones de transferencia relacionadas con la cabeza con las señales de los altavoces virtuales. La razón es que se asegura que una fuente de sonido dominante se representa en la señal de

altavoz virtual por su dirección, mientras que los sistemas de la técnica anterior con un conjunto fijo de posiciones de altavoces virtuales dividirán en general dicha fuente de sonido dominante entre las posiciones de altavoz virtual fijo más cercano. Cuando se aplican funciones de transferencia relacionadas con la cabeza, esto significa que la fuente de sonido dominante se reproducirá a través de dos conjuntos de funciones de transferencia relacionadas con la cabeza correspondientes a las dos posiciones fijas del altavoz virtual, que da como resultado una imagen espacial borrosa de la fuente de sonido dominante. De acuerdo con la invención, la fuente de sonido dominante se reproducirá a través de un conjunto de funciones de transferencia relacionadas con la cabeza correspondientes a su dirección real, dando como resultado una reproducción óptima de la información espacial 3D contenida en la señal de entrada original. La señal del altavoz virtual también es adecuada para la generación de señales de salida a altavoces reales. Se puede usar cualquier método que pueda convertir desde una señal y dirección de altavoz virtual a una matriz de señales de altavoz. Entre tales métodos se pueden mencionar

- Panorámica de amplitud
- Panorámica de amplitud de base vectorial
- Respuestas de micrófono virtual, incluidas las características de orden superior y diseños separados
- Síntesis de campo de onda
- Ambisónicos de orden superior

Por lo tanto, en una realización preferida, el procesador de audio está dispuesto para generar el conjunto de señales de salida de audio de tal manera que está dispuesto para la reproducción a través de auriculares o una matriz de altavoces, por ejemplo, aplicando funciones de transferencia relacionadas con la cabeza, u otras formas conocidas de crear efectos espaciales basados en una sola señal de entrada y su dirección.

La decodificación de la señal de entrada en el número de canales de salida representa

- determinar un conjunto de al menos una, tal como dos, tres o cuatro, posiciones de altavoz virtual seleccionadas de modo que una o más de las posiciones de altavoz virtuales coincidan al menos sustancialmente, tal como coincide exactamente, con la al menos una dirección dominante,
- decodificar la señal de entrada de audio en señales de altavoces virtuales correspondientes a cada una de las posiciones de altavoces virtuales, y
- aplicar una función de transferencia adecuada a las señales del altavoz virtual para mapear espacialmente las posiciones del altavoz virtual en el número de canales de salida que representan direcciones espaciales fijas.

A pesar de que dichas etapas pueden no estar directamente presentes en una implementación práctica de un procesador de audio o un software para ejecutarse en tal procesador, las posiciones de altavoces virtuales anteriores y las señales representan una analogía virtual para explicar una versión preferida de la invención.

El banco de filtros puede comprender al menos 500, tal como 1000 a 5000, preferiblemente filtros parcialmente solapantes que cubren el intervalo de frecuencias de 0 Hz a 22 kHz. Por ejemplo, se puede usar un análisis FFT con una longitud de ventana de 2048 a 8192 muestras, es decir, 1024-4096 bandas que cubren 0-22050 Hz. Sin embargo, se aprecia que la invención se puede realizar también con menos filtros, en caso de que se acepte un rendimiento reducido.

La unidad de separación de fuentes de sonido determina preferiblemente la al menos una dirección dominante en cada banda de frecuencia para cada marco de tiempo, tal como un marco de tiempo que tiene un tamaño de 2000 a 10000 muestras, por ejemplo, desde 2048 a 8192, como se ha mencionado. Sin embargo, se debe entender que se puede usar una actualización más baja de la dirección dominante, en caso de que se acepte un rendimiento reducido.

El número de altavoces virtuales debe ser igual o mayor que el número de direcciones dominantes determinadas por el cálculo de descomposición onda plana paramétrica. La cantidad ideal de altavoces virtuales depende del tamaño de la matriz de altavoces y del tamaño del área de escucha. En los casos en que se descubra que son ventajosos altavoces virtuales adicionales más allá de los determinados por la descomposición de onda plana paramétrica, las posiciones de los altavoces virtuales pueden determinarse mediante la construcción de una figura geométrica cuyos vértices se encuentran en la esfera de unidad. La figura está construida de modo que las direcciones dominantes coinciden con los vértices de la figura. De este modo se garantiza que las fuentes de sonido más dominantes en una banda de frecuencia se representan lo más espacialmente posible, lo que conduce a la mejor reproducción espacial de material de audio con varias fuentes de sonido dominantes distribuidas espacialmente, por ejemplo, dos cantantes o dos instrumentos musicales que se reproducen al mismo tiempo. Los vértices restantes determinan las posiciones de los altavoces virtuales adicionales. Sus ubicaciones exactas tienen poco efecto sobre la calidad de sonido resultante, siempre y cuando no haya ningún par de vértices demasiado cerca entre sí. Un cálculo específico que asegura una buena separación es el de simular cargas puntuales restringidas a colocarse sobre la superficie de una esfera. Como las cargas iguales se repelen entre sí, la posición de equilibrio de este sistema proporciona ubicaciones bien separadas en la esfera de unidad.

Como otro ejemplo, que es aplicable en el caso donde el número de direcciones dominantes es 1 o 2 y el número preferido de altavoces virtuales es 3 o 4, las siguientes construcciones geométricas son adecuadas para el cálculo de los vértices adicionales:

Número de direcciones dominantes	Número de altavoces virtuales	Método de construcción
1	3	Rotación de triángulo equilátero
2	3	Construcción de triángulo isósceles
1	4	Rotación de tetraedro regular
2	4	Construcción de tetraedro irregular con caras idénticas

5 Para generar una señal de salida multicanal, por ejemplo, dos o más canales adecuados para la reproducción a través de una matriz de altavoces, el procesador de audio puede comprender una unidad de sintetizador multicanal dispuesta para generar cualquier número de señales de salida de audio mediante la aplicación de funciones de transferencia adecuadas a cada una de las señales del altavoz virtual. Las funciones de transferencia se determinan a partir de las direcciones de los altavoces virtuales. Se conocen varios métodos adecuados para determinar tales funciones de transferencia.

15 A modo de ejemplo, se puede mencionar panorámica de amplitud, panorámica de amplitud de base de vector, síntesis de campo de ondas, características de micrófono virtuales y panorámica equivalente ambisónica. Todos estos métodos producen señales de salida adecuadas para la reproducción a través de una matriz de altavoces. También se puede optar por utilizar armónicos esféricos como funciones de transferencia, en cuyo caso las señales de salida son adecuadas para la decodificación mediante un decodificador ambisónico de orden superior. Otras funciones de transferencia también pueden ser adecuadas. Especialmente, dicho procesador de audio puede implementarse mediante una matriz de decodificación correspondiente a las posiciones del altavoz virtual determinadas y una matriz de función de transferencia correspondiente a las direcciones y al método de panoramización seleccionado, combinados en una matriz de transferencia de salida antes de aplicarse a las señales de entrada de audio. De este modo, se puede realizar un suavizado de las funciones de transferencia de dicha matriz de transferencia de salida antes de aplicarlas a las señales de entrada, lo que servirá para mejorar la reproducción de los sonidos transitorios.

25 Para generar una señal de salida de dos canales binaural, el procesador de audio puede comprender una unidad de sintetizador binaural dispuesta para generar una primera y segunda señales de salida de audio mediante la aplicación de funciones de transferencia relacionadas con la cabeza a cada una de las señales de los altavoces virtuales. Especialmente, dicho procesador de audio puede implementarse mediante una matriz de decodificación correspondiente a las posiciones del altavoz virtual determinadas y una matriz de función de transferencia correspondiente a las funciones de transferencia relacionadas con la cabeza, combinadas en una matriz de transferencia de salida antes de aplicarse a las señales de entrada de audio. De este modo, se puede realizar un suavizado de las funciones de transferencia de dicha matriz de transferencia de salida antes de aplicarlas a las señales de entrada, lo que servirá para mejorar la reproducción de los sonidos transitorios.

35 La señal de entrada de audio es preferiblemente una señal de audio multicanal dispuesta para la descomposición en componentes de onda plana. Especialmente, la señal de entrada puede ser una de entre: una señal de campo de sonido de formato B periférico o una señal de campo de sonido de formato B horizontal solamente.

40 En un segundo aspecto, la invención proporciona un dispositivo que comprende un procesador de audio de acuerdo con el primer aspecto. Especialmente, el dispositivo puede ser uno de: un dispositivo para grabar señales de sonido o video, un dispositivo para reproducir señales de sonido o video, un dispositivo portátil, un dispositivo informático, un dispositivo de videojuegos, un dispositivo de alta fidelidad, un convertidor de audio dispositivo y una unidad de auriculares.

45 En un tercer aspecto, la invención proporciona un método para convertir una señal de entrada de audio multicanal que comprende tres o cuatro canales, tales como una señal de campo de sonido de formato B, en un conjunto de señales de salida de audio, tales como un conjunto de dos señales de salida de audio (L, R) dispuestas para la reproducción de auriculares o dos o más señales de salida de audio dispuestas para la reproducción a través de una matriz de altavoces. El método se define mediante la reivindicación 14 adjunta.

50 El método puede implementarse en software puro, por ejemplo, en forma de un código genérico o en forma de un código ejecutable específico del procesador. Alternativamente, el método puede implementarse parcialmente en componentes electrónicos analógicos y/o digitales específicos y parcialmente en software. Todavía alternativamente, el método puede implementarse en un único chip dedicado.

55 Se aprecia que dos o más de las realizaciones mencionadas pueden combinarse ventajosamente. También se aprecia que las realizaciones y ventajas mencionadas para el primer aspecto se aplican también para el segundo y tercer aspectos.

Breve descripción de los dibujos

Realizaciones de la invención se describirán ahora, a modo de ejemplo solamente, con referencia a los dibujos.

- 5 La figura 1 ilustra componentes básicos de una realización del procesador de audio.
- La figura 2 ilustra detalles de una realización para convertir una señal de campo de sonido de formato B en una señal binaural,
- 10 La figura 3 ilustra una posible implementación del generador de matriz de transferencia al que se hace referencia en la figura 2,
- La figura 4 ilustra un proceso de selección de HRTF mejorado que se puede usar en la figura 2,
- 15 La figura 5 ilustra un dispositivo de audio con un procesador de audio de acuerdo con la invención, y
- La figura 6 ilustra otro dispositivo de audio con un procesador de audio de acuerdo con la invención.

Descripción de realizaciones

20 La figura 1 muestra un componente de procesador de audio con componentes básicos según la invención. La entrada al procesador de audio es una señal de audio multicanal. Esta señal se divide en una pluralidad de bandas de frecuencia en un banco de filtros, por ejemplo, en forma de un análisis de FFT realizado en cada uno de la pluralidad de canales. A continuación, se realiza una unidad de separación de fuente de sonido SSS en la señal de frecuencia separada. Primero, se realiza un cálculo de descomposición de onda plana paramétrica PWD en cada banda de frecuencia para determinar una o dos direcciones de fuente de sonido dominante. Las direcciones de la fuente de sonido dominante se aplican entonces a un algoritmo de cálculo de posición de altavoz virtual VLP que sirve para seleccionar un conjunto de direcciones de fuente de sonido virtual o altavoz virtual, por ejemplo, mediante la rotación de un conjunto fijo de direcciones de altavoces virtuales, tal que una o ambas, en caso de dos, direcciones de la fuente de sonido dominante coinciden con las direcciones de los altavoces virtuales respectivos. La operación precisa realizada por el VLP depende del número de estimaciones de dirección y del número deseado de altavoces virtuales. Ese número a su vez depende de la cantidad de canales de entrada, el tamaño de la matriz de altavoces y el tamaño del área de escucha. Un mayor número de altavoces virtuales generalmente conduce a una mejor sensación de envolvimiento para los oyentes en una posición central de escucha, mientras que un número más pequeño de altavoces virtuales conduce a una localización más precisa para los oyentes fuera de la posición central de escucha.

35 A continuación, la señal de entrada se transfiere o decodifica DEC de acuerdo con una matriz de decodificación correspondiente a las direcciones de altavoces virtuales seleccionados, y funciones de transferencia relacionadas con la cabeza, opcionalmente, u otras funciones de transferencia dependientes de la dirección que corresponden a las direcciones de altavoces virtuales se aplican antes de que los componentes de frecuencia finalmente se combinen en una unidad de suma SU para formar un conjunto de señales de salida, por ejemplo, dos señales de salida en caso de una implementación binaural, o cuatro, cinco, seis, siete o incluso más señales de salida en caso de conversión a formato adecuado para la reproducción a través de una configuración de sonido envolvente de altavoces. Si el banco de filtros se implementa como un análisis FFT, la suma se puede implementar como una transformación IFFT seguida de una etapa de adición de superposición.

40 El procesador de audio se puede implementar de varias maneras, por ejemplo, en forma de un procesador que forma parte de un dispositivo, en el que el procesador está provisto de un código ejecutable para llevar a cabo la invención.

45 Las figuras 2 y 3 ilustran componentes de una realización preferida adecuada para convertir una señal de entrada que tiene características tridimensionales y está en un "formato B ambisónico". El sistema de formato B ambisónico es un sistema de posicionamiento de sonido de muy alta calidad que opera al descomponer la direccionalidad del sonido en componentes armónicos esféricos denominados W, X, Y y Z. El sistema ambisónico está diseñado para utilizar una pluralidad de altavoces de salida para recrear de forma cooperativa los componentes direccionales originales. Para obtener una descripción del sistema de formato B, se hace referencia a: <http://en.wikipedia.org/wiki/Ambisonics>.

60 Con referencia a la figura 2, la realización preferida está dirigida a proporcionar una espacialización mejorada de las señales de audio de entrada. Se ingresa una señal de formato B que tiene componentes X, Y, Z y W. Cada componente del conjunto de entrada de formato B se procesa a través de un banco de filtros correspondiente (1)-(4), cada uno de los cuales divide la entrada en varias bandas de frecuencia de salida (el número de bandas depende de la implementación, normalmente en el intervalo de 1024 a 4096).

65

Los elementos (5), (6), (7), (8) y (10) se replican una vez para cada banda de frecuencia, aunque solo se muestra uno de cada uno en la figura 2. Para cada banda de frecuencia, las cuatro señales (una de cada banco de filtros (1)-(4)) son procesadas por un elemento de descomposición de onda plana paramétrica (5), que determina el menor número de ondas planas necesarias para recrear el campo de sonido local codificado en las cuatro señales. El elemento de descomposición de onda plana paramétrica también calcula la dirección, fase y amplitud de estas ondas. La señal de entrada se indica w, x, y, z , con los subíndices r e i . A continuación, se supone que los canales están escalados de forma que la amplitud máxima de una onda de plano único sea igual en todos los canales. Esto implica que el canal W puede tener que escalarse por un factor de 1, $\sqrt{2}$ o $\sqrt{3}$, dependiendo de si la señal de entrada se escala de acuerdo con las convenciones SN3D, FuMa o N3D, respectivamente. El campo de sonido local puede en la mayoría de los casos ser recreado por dos ondas planas, como se expresa en las siguientes ecuaciones:

$$\begin{bmatrix} w_1 \\ x_1 \\ y_1 \\ z_1 \end{bmatrix} e^{i\phi_1} + \begin{bmatrix} w_2 \\ x_2 \\ y_2 \\ z_2 \end{bmatrix} e^{i\phi_2} = \begin{bmatrix} w_r \\ x_r \\ y_r \\ z_r \end{bmatrix} + \begin{bmatrix} w_i \\ x_i \\ y_i \\ z_i \end{bmatrix} i \quad (1)$$

$$x_1^2 + y_1^2 + z_1^2 = w_1^2 \quad (2)$$

$$x_2^2 + y_2^2 + z_2^2 = w_2^2 \quad (3)$$

La solución a estas ecuaciones es

$$\begin{bmatrix} w_1 & x_1 & y_1 & z_1 \\ w_2 & x_2 & y_2 & z_2 \end{bmatrix} = \begin{bmatrix} \cos \phi_1 & \cos \phi_1 \\ \sin \phi_2 & \sin \phi_2 \end{bmatrix}^{-1} \begin{bmatrix} w_r & x_r & y_r & z_r \\ w_i & x_i & y_i & z_i \end{bmatrix} \quad (4)$$

donde

$$\cos^2 \phi_n = \frac{2a^2 - bc + b^2 \pm 2a\sqrt{a^2 - bc}}{(c - b)^2 + 4a^2} \quad (5)$$

$$a = -w_r w_i + x_r x_i + y_r y_i + z_r z_i \quad (6)$$

$$b = -w_r^2 + x_r^2 + y_r^2 + z_r^2 \quad (7)$$

$$c = -w_i^2 + x_i^2 + y_i^2 + z_i^2 \quad (8)$$

Los dos posibles signos en la ecuación 5 dan los valores de $\cos^2 \phi_1$ y $\cos^2 \phi_2$, respectivamente, siempre que $a^2 - bc$ no sea negativo. Cada valor para $\cos^2 \phi_n$ corresponde a varios valores posibles de ϕ_n , uno en cada cuadrante, o los valores 0 y π , o los valores $\pi/2$ y $3\pi/2$. Solo uno de estos es correcto. El cuadrante correcto se puede determinar a partir de la ecuación 9 y el requisito de que w_1 y w_2 sean positivos.

$$\sin \phi_n \cos \phi_n = \frac{(c - b) \cos^2 \phi_n + b}{2a} \quad (9)$$

Cuando la ecuación 5 da soluciones reales, más de dos ondas planas son necesarias para reconstruir el campo de sonido local. También puede ser ventajoso usar un método alternativo cuando la matriz a invertir en la ecuación 4 es singular o casi singular. Al permitir más de dos ondas planas, existe una cantidad infinita de posibles soluciones. Dado que este método alternativo es necesario solo para una pequeña parte de la mayoría de las señales, la elección de la solución no es crítica. Una opción posible es la de dos ondas planas que viajan en las direcciones de los ejes principales de la elipse, que se describe mediante el vector de velocidad dependiente del tiempo asociado

con cada banda de frecuencia. Además de estas dos ondas planas, una onda esférica es necesaria para reconstruir el componente W de la señal entrante:

$$\begin{bmatrix} w_0 \\ 0 \\ 0 \\ 0 \end{bmatrix} e^{i\phi_0} + \begin{bmatrix} w_1 \\ x_1 \\ y_1 \\ z_1 \end{bmatrix} e^{i\phi_1} + \begin{bmatrix} w_2 \\ x_2 \\ y_2 \\ z_2 \end{bmatrix} e^{i\phi_2} = \begin{bmatrix} w_r \\ x_r \\ y_r \\ z_r \end{bmatrix} + \begin{bmatrix} w_i \\ x_i \\ y_i \\ z_i \end{bmatrix} i \quad (10)$$

5

$$x_1^2 + y_1^2 + z_1^2 = w_1^2 \quad (11)$$

$$x_2^2 + y_2^2 + z_2^2 = w_2^2 \quad (12)$$

10 La solución elegida es

$$\begin{bmatrix} w'_1 & x_1 & y_1 & z_1 \\ w'_2 & x_2 & y_2 & z_2 \end{bmatrix} = \begin{bmatrix} \cos \phi_1 & \cos \phi_1 \\ \sin \phi_2 & \sin \phi_2 \end{bmatrix}^{-1} \begin{bmatrix} w_r & x_r & y_r & z_r \\ w_i & x_i & y_i & z_i \end{bmatrix} \quad (13)$$

donde

15

$$\cos^2 \phi_n = \frac{1}{2} \pm \frac{b-c}{2\sqrt{4a^2 + (b-c)^2}} \quad (14)$$

$$a = x_r x_i + y_r y_i + z_r z_i \quad (15)$$

20

$$b = x_r^2 + y_r^2 + z_r^2 \quad (16)$$

$$c = x_i^2 + y_i^2 + z_i^2 \quad (17)$$

25 Como antes, el cuadrante de ϕ se puede determinar sobre la base de otra ecuación (18) y el requisito de que w'_1 y w'_2 debe ser positivo.

$$\sin \phi_n \cos \phi_n = \frac{2a \cos^2 \phi_n - a}{b-c} \quad (18)$$

30 Los valores de w_0 y ϕ_0 no se utilizan en las etapas posteriores.

35 La salida de (5) se compone de los dos vectores $\langle x_1, y_1, z_1 \rangle$ y $\langle x_2, y_2, z_2 \rangle$. Esta salida está conectada a un elemento (6) que ordena estos dos vectores de acuerdo con sus longitudes o el valor de su elemento y . En una realización alternativa de la invención, solo uno de los dos vectores se pasa desde el elemento (6). La elección puede ser la del vector más largo o el que tenga el mayor grado de similitud con los vectores vecinos. La salida de (6) está conectada a un elemento de suavizado (7) que suprime los cambios rápidos en las estimaciones de dirección. La salida de (7) está conectada a un elemento (8) que genera funciones de transferencia adecuadas desde cada una de las señales de entrada a cada una de las señales de salida, un total de ocho funciones de transferencia. Cada una de estas funciones de transferencia se pasa a través de un elemento de suavizado (9). Este elemento suprime las grandes diferencias de fase y de amplitud entre las bandas de frecuencia vecinas y también suprime los cambios temporales rápidos de fase y de amplitud. La salida de (9) se pasa a un multiplicador matricial (10) que aplica las funciones de transferencia a las señales de entrada y crea dos señales de salida. Los elementos (11) y (12) suman cada una de las señales de salida de (10) a través de todas las bandas de filtro para producir una señal binaural. Por lo general,

40

no es necesario aplicar suavizado antes y después de la generación de la matriz de transferencia, por lo que normalmente se puede eliminar el elemento (7) o el elemento (9). Es preferible en ese caso eliminar el elemento (7).

5 Con referencia a la figura 3, se ilustra esquemáticamente la realización preferida del generador de matriz de transferencia al que se hace referencia en la figura 2. Un elemento (1) genera dos vectores nuevos cuyas direcciones se eligen para distribuir los altavoces virtuales sobre la esfera de la unidad. En una realización alternativa de la invención, solo se pasa un vector al generador de la matriz de transferencia. En este caso, el elemento (1) debe generar tres vectores nuevos, preferiblemente de manera tal que los cuatro vectores resultantes apuntan hacia los vértices de un tetraedro regular. Este enfoque alternativo también es beneficioso en los casos en que los dos vectores de entrada son colineales o casi colineales.

10 Los cuatro vectores se utilizan para representar las instrucciones para cuatro altavoces virtuales que serán utilizados para reproducir las señales de entrada. Un elemento (6) calcula una matriz de decodificación invirtiendo la siguiente matriz:

15

$$\mathbf{G} = \begin{bmatrix} 1 & x'_1 & y'_1 & z'_1 \\ 1 & x'_2 & y'_2 & z'_2 \\ 1 & x'_3 & y'_3 & z'_3 \\ 1 & x'_4 & y'_4 & z'_4 \end{bmatrix} \quad (19)$$

donde

$$\begin{bmatrix} x'_n & y'_n & z'_n \end{bmatrix} = \frac{\begin{bmatrix} x_n & y_n & z_n \end{bmatrix}}{\sqrt{x_n^2 + y_n^2 + z_n^2}} \quad (20)$$

20

Un elemento (5) almacena un conjunto de funciones de transferencia relacionadas con la cabeza.

25 El elemento (2) usa las direcciones del altavoz virtual para seleccionar e interpolar entre las funciones de transferencia relacionadas con la cabeza más cercanas a la dirección de cada altavoz virtual. Para cada altavoz virtual, hay dos funciones de transferencia relacionadas con la cabeza; una para cada oído, que proporciona un total de ocho funciones de transferencia que se pasan al elemento (7). Las salidas de los elementos (2) y (6) se multiplican en una multiplicación de matriz (7) para producir la matriz de transferencia adecuada.

30 El diseño ilustrado en la figura 2 puede modificarse de las siguientes maneras para producir una salida multicanal adecuada para la alimentación de una matriz de altavoces de n altavoces:

- El generador de matriz de transferencia (8) se modifica para producir n x 4 funciones de transferencia en lugar de 2 x 4.
- El elemento de suavizado (9) se modifica para suavizar las funciones de transferencia n x 4.
- 35 - El multiplicador matricial (10) se modifica para multiplicar el vector de señal de entrada con una matriz n x 4 y para producir un vector de salida con n elementos.
- Se agregan unidades de suma adicionales para procesar las salidas adicionales de (10).

40 El diseño ilustrado en la figura 3 puede modificarse de las siguientes maneras de producir n x 4 funciones de transferencia adecuadas para producir una salida multicanal:

- Las funciones de transferencia relacionadas con la cabeza en el elemento (5) se reemplazan por funciones de panoramización por pares, funciones de panoramización de amplitud de base vectorial, características de micrófono virtual u otras funciones adecuadas para producir la ilusión de sonido que emana de las direcciones de los altavoces virtuales.
- 45 - El elemento (2) se modifica para seleccionar n x 4 funciones de transferencia en lugar de 2 x 4.
- El elemento (7) se modifica para producir n x 4 funciones de transferencia en lugar de 2 x 4.

50 El diseño ilustrado en la figura 2 puede ser modificado en las siguientes formas de procesar tres señales de entrada de audio que constituyen una única horizontal de la señal de formato B:

- Se retira el banco de filtros Z (3)
- El elemento de descomposición de onda plana (5) se modifica eliminando z_r , z_i , z_1 y z_2 de las ecuaciones 1-17.

- El multiplicador matricial (10) se modifica para recibir tres entradas en lugar de cuatro.
- El elemento de suavizado (9) se modifica para suavizar las funciones de transferencia de 2 x 3 en lugar de 2 x 4.
- El generador de matriz de transferencia (8) se modifica para producir 2 x 3 funciones de transferencia en lugar de 2 x 4.

5 El diseño ilustrado en la figura 3 puede modificarse de las siguientes maneras para producir funciones de transferencia 2 x 3 adecuadas para el procesamiento de tres señales de entrada de audio que constituyen una única horizontal de la señal de formato B:

- 10 - El elemento (1) genera un nuevo vector cuya dirección se elige para maximizar los ángulos entre los tres vectores resultantes. En una realización alternativa de la invención, solo se pasa un vector a veces al generador de la matriz de transferencia. En este caso, el elemento (1) debe generar dos vectores nuevos, preferiblemente de manera que los tres vectores resultantes apuntan hacia los vértices de un triángulo equilátero.
- 15 - El elemento (6) calcula una matriz de decodificación invirtiendo la siguiente matriz:

$$\mathbf{G} \equiv \begin{bmatrix} 1 & x'_1 & y'_1 \\ 1 & x'_2 & y'_2 \\ 1 & x'_3 & y'_3 \end{bmatrix} \quad (21)$$

donde

$$\begin{bmatrix} x'_n & y'_n \end{bmatrix} = \frac{\begin{bmatrix} x_n & y_n \end{bmatrix}}{\sqrt{x_n^2 + y_n^2}} \quad (22)$$

- 20 - El elemento (2) se modifica para seleccionar 2 x 3 funciones de transferencia en lugar de 2 x 4.
- El elemento (4) se modifica para integrar la fase de 2 x 3 funciones de transferencia en lugar de 2 x 4.
- 25 - El elemento (7) se modifica para producir 2 x 3 funciones de transferencia en lugar de 2 x 4.

En los casos donde se encuentran que una serie de altavoces virtuales diferentes a partir del número de canales de entrada es ventajoso, el diseño de la figura 3 se puede modificar de la siguiente manera:

- 30 - El elemento de vértices opuestos (1) se modifica para generar un número de direcciones más pequeño o más grande.
- El elemento (6) se altera para calcular el pseudo-inverso de Moore-Penrose de la matriz G, que en este caso no es una matriz cuadrada.
- El elemento (2) se modifica para seleccionar el número requerido de funciones de transferencia.
- 35 - El elemento (7) se altera para multiplicar las matrices de entrada de diferentes tamaños.

Estos cambios no alteran la forma de la matriz de transferencia resultante.

40 Otra de las mejoras para el diseño ilustrado en la figura 3 pertenece a transferir funciones que contienen un retardo de tiempo, tales como funciones de transferencia relacionadas con la cabeza. La diferencia en el tiempo de propagación para cada una de las dos orejas conduce a un retardo de tiempo interaural que depende de la ubicación de la fuente. Este retraso se manifiesta en las funciones de transferencia relacionadas con la cabeza como un cambio de fase interaural que es aproximadamente proporcional a la frecuencia y depende de la ubicación de la fuente. En el contexto de esta invención, solo se conoce una estimación de la ubicación de la fuente, y cualquier incertidumbre en esta estimación se traduce en una incertidumbre en el desplazamiento de fase interaural que es proporcional a la frecuencia. Esto puede conducir a una reproducción pobre de sonidos transitorios.

45 La capacidad humana para percibir desplazamiento de fase interaural se limita a frecuencias por debajo de aproximadamente 1200-1600 Hz. Aunque el cambio de fase interaural en sí mismo no contribuye a la localización a frecuencias más altas, la demora del grupo interaural sí lo hace. El retraso del grupo interaural se define como la derivada parcial negativa del cambio de fase interaural con respecto a la frecuencia. A diferencia del cambio de fase interaural, la demora del grupo interaural permanece aproximadamente constante a través de todas las frecuencias para cualquier ubicación de fuente dada. Para reducir el ruido de fase, es por lo tanto ventajoso calcular el retraso del grupo interaural por diferenciación numérica de los HRTF antes de que el elemento (2) seleccione HRTF dependiendo de las direcciones de los altavoces virtuales. Después de la selección, pero antes de que las funciones de transferencia resultantes pasen al elemento (7), es necesario calcular el desplazamiento de fase de las funciones

de transferencia resultantes por integración numérica.

Este proceso de reducción de ruido de fase se ilustra en la figura 4. El elemento (1) almacena un conjunto de HRTF para diferentes direcciones de incidencia. El elemento (2) descompone estas funciones de transferencia en una parte de amplitud y una parte de fase. El elemento (3) diferencia la parte de fase para calcular un retraso de grupo. El elemento (4) selecciona y (opcionalmente) interpola un retardo de amplitud, fase y grupo en función de una dirección de llegada. El elemento (5) diferencia el cambio de fase resultante después de la selección. El elemento (6) calcula una combinación lineal de las estimaciones de retardo de dos grupos de modo que su entrada izquierda se utiliza a bajas frecuencias, y transita sin problemas a la entrada correcta para frecuencias superiores a 1600 Hz. El elemento (7) recupera un desplazamiento de fase desde el retardo de grupo y el elemento (8) recupera una función de transferencia en componentes cartesianos (reales/imaginarios), adecuada para un procesamiento posterior.

Este proceso puede sustituir ventajosamente al elemento (2) en la figura 3, donde se requiere una instancia del proceso para cada altavoz virtual. Como el proceso conecta indirectamente las estimaciones de dirección de las bandas de frecuencia vecinas, es preferible que cada fuente de sonido se envíe al mismo altavoz virtual para todas las bandas de frecuencia vecinas donde esté presente. Este es el propósito del elemento de clasificación (6) en la figura 2.

El mismo proceso es aplicable a otras funciones de panorámica que HRTF que contienen un retardo entre canales también. Algunos ejemplos son las características de respuesta de micrófono virtual de una configuración de micrófono ORTF o Decca Tree o cualquier otra configuración de micrófono virtual espaciado.

En la disposición mostrada en la figura 3, la matriz de decodificación se multiplica por la matriz de funciones de transferencia antes de que su producto se multiplique con las señales de entrada. En una realización alternativa de la invención, las señales de entrada se multiplican primero con la matriz de decodificación y su producto se multiplica posteriormente con la matriz de función de transferencia. Sin embargo, esto excluiría la posibilidad de suavizar las funciones generales de transferencia. Este suavizado es ventajoso para la reproducción de sonidos transitorios.

El efecto global de la disposición mostrada en las figuras 2 y 3 es descomponer el espectro completo del campo de sonido local en un gran número de ondas planas y pasar estas ondas planas a través de las correspondientes funciones de transferencia relacionadas con la cabeza para producir una señal binaural adecuada para la reproducción de auriculares.

La figura 5 ilustra un diagrama de bloques de un dispositivo de audio con un procesador de audio de acuerdo con la invención, por ejemplo, el ilustrado en las figuras 2 y 3. El dispositivo puede ser una unidad de auriculares dedicada, un dispositivo de audio general que ofrece la conversión de una señal de entrada multicanal a otro formato de salida como opción, o el dispositivo puede ser un ordenador general con una tarjeta de sonido provista de un software adecuado para realizar el método de conversión de acuerdo con la invención.

El dispositivo puede ser capaz de realizar la conversión en línea de la señal de entrada, por ejemplo, mediante la recepción de la señal de audio de entrada de múltiples canales en forma de un flujo de bits digital. Alternativamente, por ejemplo, si el dispositivo es un ordenador, el dispositivo puede generar la señal de salida en forma de un archivo de salida de audio basado en un archivo de audio como entrada.

La figura 6 ilustra un diagrama de bloques de un dispositivo de audio con un procesador de audio de acuerdo con la invención, por ejemplo, el ilustrado en las figuras 2 y 3, modificado para salida multicanal. El dispositivo puede ser una unidad de decodificador dedicada, un dispositivo de audio general que ofrece la conversión de una señal de entrada multicanal a otro formato de salida como opción, o el dispositivo puede ser un ordenador general con una tarjeta de sonido provista de un software adecuado para realizar el método de conversión de acuerdo con la invención.

A continuación, se define un conjunto de aspectos:

E1. Un procesador de audio dispuesto para convertir una señal de entrada de audio multicanal (X, Y, Z, W) que comprende al menos dos canales, como una señal de campo de sonido de formato B, en un conjunto de señales de salida de audio (L, R), como un conjunto de dos señales de salida de audio (L, R) dispuestas para la reproducción de auriculares, comprendiendo el procesador de audio

- un banco de filtros dispuesto para separar la señal de entrada (X, Y, Z, W) en una pluralidad de bandas de frecuencia, tal como bandas de frecuencia parcialmente solapadas,
- una unidad de separación de fuente de sonido dispuesta, por lo menos para una parte de la pluralidad de bandas de frecuencia, para
- realizar un cálculo de expansión de onda plana en la señal de entrada de audio multicanal (X, Y, Z, W) para determinar al menos una dirección dominante correspondiente a una dirección de una fuente de sonido dominante en la señal de entrada de audio (X, Y, Z, W),

ES 2 690 164 T3

- determinar una matriz de al menos dos, tal como cuatro, posiciones de altavoz virtual seleccionadas de modo que una o más de las posiciones de altavoz virtuales coincidan al menos sustancialmente, tal como coincide exactamente, con al menos una dirección dominante, y
 - decodificar la señal de entrada de audio (X, Y, Z, W) en señales de altavoces virtuales correspondientes a cada una de las posiciones del altavoz virtual, y
- 5
- una unidad sumadora dispuesta para sumar las señales del altavoz virtual para que la al menos parte de la pluralidad de bandas de frecuencia llegue al conjunto de señales de salida de audio (L, R).
- 10
- E2. Procesador de audio según E1, en el que el banco de filtros comprende al menos 500, tal como de 1000 a 5000, filtros parcialmente solapados que cubren un intervalo de frecuencia de 0 Hz a 22 kHz.
- E3. Procesador de audio de acuerdo con E1 o E2, en el que las posiciones del altavoz virtual se seleccionan mediante una rotación de un conjunto de al menos tres posiciones en una interrelación espacial fija.
- 15
- E4. Procesador de audio según E3, en el que el conjunto de posiciones en una interrelación espacial fija comprende cuatro posiciones, tal como cuatro posiciones dispuestas en un tetraedro.
- E5. Procesador de audio según cualquiera de E1-E4, en el que la expansión de onda determina dos direcciones dominantes, y en el que la matriz de al menos dos posiciones de altavoz virtual se selecciona de tal manera que dos de las posiciones de altavoz virtual coinciden al menos sustancialmente, tal como coincide exactamente, con las dos direcciones dominantes.
- 20
- E6. Procesador de audio según E1-E5, que comprende una unidad de sintetizador binaural dispuesta para generar una primera y segunda señales de salida de audio (L, R) aplicando funciones de transferencia relacionadas con la cabeza (HRTF) a cada una de las señales de altavoces virtuales.
- 25
- E7. Procesador de audio según E6, en el que una matriz de decodificación correspondiente a las posiciones de altavoz virtuales determinadas y una matriz de función de transferencia correspondiente a las funciones de transferencia relacionadas con la cabeza (HRTF) se combinan en una matriz de transferencia de salida antes de aplicarse a las señales de entrada de audio (X, Y, Z, W).
- 30
- E8. Procesador de audio según E7, en el que se realiza un suavizado en las funciones de transferencia de la matriz de transferencia de salida antes de aplicarse a las señales de entrada (X, Y, Z, W).
- 35
- E9. Procesador de audio según cualquiera de E6-E8, en el que la fase de las funciones de transferencia relacionadas con la cabeza (HRTF) se diferencia con respecto a la frecuencia y después de combinar componentes de funciones de transferencia relacionadas con la cabeza (HRTF) correspondientes a diferentes direcciones, la fase de las funciones de transferencia combinadas está integrada con respecto a la frecuencia.
- 40
- E10. Procesador de audio según cualquiera de E1-E9, en el que la fase de las funciones de transferencia relacionadas con la cabeza (HRTF) se mantiene inalterada por debajo de un primer límite de frecuencia, tal como por debajo de 1,6 kHz, y diferenciada con respecto a frecuencia en frecuencias superiores a una segunda frecuencia límite con una frecuencia más alta que el primer límite de frecuencia, tal como 2,0 kHz, y con una transición gradual entre, y después de combinar componentes de funciones de transferencia relacionadas con la cabeza (HRTF) correspondientes a diferentes direcciones, la operación inversa se aplica a la función combinada.
- 45
- E11. Procesador de audio según cualquiera de E1-E10, en el que la señal de entrada de audio es una señal de audio multicanal dispuesta para la descomposición en componentes de onda plana, tal como uno de: una señal de campo de sonido de formato B, una grabación ambisónica de orden superior, una grabación estéreo y una grabación de sonido envolvente.
- 50
- E12. Procesador de audio según cualquiera de E1-E12, en el que la unidad de separación de fuente de sonido determina la al menos una dirección dominante en cada banda de frecuencia para cada marco de tiempo, en el que un marco de tiempo tiene un tamaño de 2.000 a 10.000 muestras.
- 55
- E13. Procesador de audio según cualquiera de E1-E12, en el que el conjunto de señales de salida de audio (L, R) están dispuestas para la reproducción a través de auriculares.
- 60
- E14. Dispositivo que comprende un procesador de audio según E1-E13, tal como el dispositivo que es uno de: un dispositivo para grabar señales de sonido o video, un dispositivo para reproducir señales de sonido o video, un dispositivo portátil, un dispositivo informático, un dispositivo de videojuegos, un dispositivo de alta fidelidad, un dispositivo convertidor de audio y una unidad de auriculares.
- 65
- E15. Método para convertir una señal de entrada de audio multicanal (X, Y, Z, W) que comprende al menos dos canales, tal como una señal de campo de sonido de formato B, en un conjunto de señales de salida de audio (L,

R), tal como un conjunto de dos señales de salida de audio (L, R) dispuestas para la reproducción de auriculares, comprendiendo el método

- 5 - separar la señal de entrada (X, Y, Z, W) en una pluralidad de bandas de frecuencia, tal como bandas de frecuencia parcialmente solapadas,
- realizar una separación de fuente de sonido para al menos una parte de la pluralidad de bandas de frecuencia, que comprende
- 10 - realizar un cálculo de expansión de onda plana en la señal de entrada de audio multicanal (X, Y, Z, W) para determinar al menos una dirección dominante correspondiente a una dirección de una fuente de sonido dominante en la señal de entrada de audio (X, Y, Z, W),
- determinar una matriz de al menos dos, tal como cuatro, posiciones de altavoz virtual seleccionadas de modo que una o más de las posiciones de altavoz virtuales coincidan al menos sustancialmente, tal como coincide exactamente, con al menos una dirección dominante, y
- 15 - decodificar la señal de entrada de audio (X, Y, Z, W) en señales de altavoces virtuales correspondientes a cada una de las posiciones del altavoz virtual, y
- sumar las señales del altavoz virtual para que la al menos parte de la pluralidad de bandas de frecuencia llegue al conjunto de señales de salida de audio (L, R).

A continuación, se define otro conjunto de aspectos:

20 EE1. Un procesador de audio dispuesto para convertir una señal de entrada de audio multicanal que comprende al menos dos canales, tal como una señal estéreo o una señal de campo de sonido en formato B de tres o cuatro canales, en un conjunto de señales de salida de audio, tal como conjunto de dos señales de salida de audio dispuestas para auriculares o dos o más señales de salida de audio dispuestas para su reproducción a través de una matriz de altavoces, comprendiendo el procesador de audio

- 25 - un banco de filtros dispuesto para separar la señal de entrada en una pluralidad de bandas de frecuencia, tal como bandas de frecuencia parcialmente solapadas,
- 30 - una unidad de separación de fuente de sonido dispuesta, por lo menos para una parte de la pluralidad de bandas de frecuencia, para
- realizar un cálculo de expansión de onda plana en la señal de entrada de audio multicanal para determinar al menos una dirección dominante correspondiente a una dirección de una fuente de sonido dominante en la señal de entrada de audio,
- 35 - realizar una decodificación de la señal de entrada de audio en una serie de canales de salida, en el que dicha decodificación se controla según dicha al menos una dirección dominante, y
- 40 - una unidad sumadora dispuesta para sumar las señales resultantes de los canales de salida respectivos para que la al menos parte de la pluralidad de bandas de frecuencia llegue al conjunto de señales de salida de audio.

EE2. Procesador de audio según EE1, en el que dicha decodificación de la señal de entrada en el número de canales de salida representa

- 45 - determinar una matriz de al menos dos, tal como cuatro, posiciones de altavoz virtual seleccionadas de modo que una o más de las posiciones de altavoz virtuales coincidan al menos sustancialmente, tal como coincide exactamente, con la al menos una dirección dominante,
- decodificar la señal de entrada de audio en señales de altavoces virtuales correspondientes a cada una de las posiciones de altavoces virtuales, y
- 50 - aplicar una función de transferencia adecuada a las señales del altavoz virtual para mapear espacialmente las posiciones del altavoz virtual en el número de canales de salida que representan direcciones espaciales fijas.

55 EE3. Procesador de audio según EE1 o EE2, en el que la señal de entrada de audio multicanal comprende dos, tres o cuatro canales, en el que el banco de filtros está dispuesto para separar cada uno de los canales de entrada de audio en una pluralidad de bandas de frecuencia, tal como bandas de frecuencia parcialmente solapadas, en el que una unidad de expansión de ondas planas está dispuesta para expandir un campo de sonido local representado en los canales ondas planas o al menos determina una o dos direcciones estimadas de llegada,

60 en el que una unidad de vértices opuestos está dispuesta para complementar las direcciones estimadas con direcciones fantasmas, en el que un calculador de matriz de decodificación está dispuesto para calcular una matriz decodificadora adecuada para descomponer la señal de entrada de audio en alimentadores para altavoces virtuales, donde las direcciones de dichos altavoces virtuales están determinadas por las salidas combinadas de la unidad de expansión de onda plana y la unidad de vértices opuestos

65 en el que un selector de función de transferencia está dispuesto para calcular una matriz de funciones de

- transferencia adecuadas, tal como funciones de transferencia relacionadas con la cabeza, para producir una ilusión de sonido que emana de las direcciones de dichos altavoces virtuales, en el que una primera unidad de multiplicación de matriz está dispuesta para multiplicar las salidas del calculador de matriz de decodificación y el selector de función de transferencia,
- 5 en el que una segunda unidad de multiplicación de matriz está dispuesta para multiplicar una de la batería de filtros con una salida de la primera unidad de multiplicación de matriz, tal como una salida de una unidad de suavizado que opera en la salida de la primera unidad de multiplicación de matriz, y en el que una pluralidad de unidades sumadoras están dispuestas para sumar las señales respectivas en la pluralidad de bandas de frecuencia para producir el conjunto de señales de salida de audio.
- 10 EE4. Procesador de audio según EE1-EE3, en el que el banco de filtros comprende al menos 20, tal como al menos 100, tal como al menos 500, tal como de 1000 a 5000, filtros parcialmente solapados que cubren un intervalo de frecuencia de 0 Hz a 22 kHz.
- 15 EE5. Procesador de audio según EE1-EE4, en el que una unidad de suavizado está conectada entre la unidad de expansión de onda plana y al menos una unidad que recibe una salida de la unidad de expansión de onda plana, en el que la unidad de suavizado está dispuesta para suprimir grandes diferencias en las estimaciones de dirección entre bandas de frecuencia vecinas y rápidos cambios de dirección en el tiempo.
- 20 EE6. Procesador de audio según EE1-EE5, en el que la primera unidad de multiplicación de matriz está conectada para recibir una salida del banco de filtros y la calculadora de matriz de decodificación, y en el que la segunda unidad de multiplicación de matriz está conectada a la primera unidad de multiplicación de matriz y al selector de función de transferencia.
- 25 EE7. Procesador de audio según cualquiera de EE1-EE6, en el que una unidad de suavizado está conectada entre la primera y la segunda unidades de multiplicación de matrices, en el que la unidad de suavizado está dispuesta para suprimir grandes diferencias entre elementos de matriz correspondientes en bandas de frecuencia vecinas y cambios rápidos de elementos de matriz en el tiempo.
- 30 EE8. Procesador de audio según cualquiera de EE1-EE7, que comprende un selector de función de transferencia que selecciona funciones de transferencia de una base de datos de funciones de transferencia relacionadas con la cabeza (HRTF), produciendo así dos canales de salida adecuados para la reproducción a través de auriculares.
- 35 EE9. Procesador de audio según EE8, en el que un diferenciador de fase calcula la diferencia de fase de las funciones de transferencia relacionadas con la cabeza (HRTF) entre bandas de frecuencia vecinas, y en el que un integrador de fase acumula las diferencias de fase después de combinar componentes de funciones de transferencia relacionadas con la cabeza (HRTF) correspondientes a diferentes direcciones.
- 40 EE10. Procesador de audio según EE9, en el que el diferenciador de fase deja la fase inalterada por debajo de un primer límite de frecuencia, tal como por debajo de 1,6 kHz, y calcula la diferencia de fase entre bandas de frecuencia vecinas por encima de un segundo límite de frecuencia con una frecuencia mayor que el primer límite de frecuencia, tal como 2,0 kHz, y con una transición gradual entre, y donde el integrador de fase realiza la operación inversa.
- 45 EE11. Procesador de audio según cualquiera de EE1-EE10, que comprende un selector de función de transferencia que selecciona funciones de transferencia de acuerdo con una ley de panoramización por pares, produciendo así dos o más canales de salida adecuados para la reproducción en una matriz horizontal de altavoces.
- 50 EE12. Procesador de audio según EE1-EE11, que comprende un selector de función de transferencia que selecciona funciones de transferencia de acuerdo con la amplitud de base de vector, panorámica equivalente ambisónica, o síntesis de campo de onda, produciendo cuatro o más canales de salida adecuados para la reproducción en una matriz 3D de altavoces.
- 55 EE13. Procesador de audio según cualquiera de EE1-EE12, que comprende un selector de función de transferencia que selecciona transferencia evaluando funciones armónicas esféricas, produciendo así tres o más canales de salida adecuados para decodificación con un decodificador ambisónico de primer orden o un decodificador ambisónico de orden superior.
- 60 EE14. Procesador de audio según cualquiera de EE1-EE13, en el que la señal de entrada de audio es una señal de campo de sonido de formato B de tres o cuatro canales.
- 65 EE15. Procesador de audio según cualquiera de EE1-EE14, en el que una unidad de retardo está conectada a la salida del banco de filtros y la entrada de la unidad de expansión de ondas planas, y en el que se mantiene la conexión directa entre dichas dos unidades, y en el que la señal de entrada de audio es una señal estéreo, tal

como una mezcla estéreo de una pluralidad de fuentes de sonido, tal como una mezcla que usa una técnica de pan-pot.

5 EE16. Procesador de audio según EE15, en el que la señal de entrada de audio se origina a partir de una configuración de micrófono coincidente, tal como un par Blumlein, un par X/Y, una configuración Mid/Side con un micrófono medio cardioide, una configuración Mid/Side con un micrófono medio hipercardioide, una configuración Mid/Side con un micrófono medio subcardioide, una configuración Mid/Side con un micrófono medio omnidireccional.

10 EE17. Procesador de audio según EE16, en el que la sensibilidad medida de los micrófonos, en función del azimut y la frecuencia, se usa en la unidad de expansión de onda plana y en el calculador de la matriz de decodificación.

15 EE18. Procesador de audio según cualquiera de EE15-EE17, en el que se inserta una segunda unidad de retardo entre las salidas del banco de filtros y la segunda unidad de multiplicación de matrices.

20 EE19. Procesador de audio según cualquiera de EE1-EE18, en el que la unidad de separación de fuente de sonido opera en entradas con un marco de tiempo que tiene un tamaño de 1.000 a 20.000 muestras, tal como 2.000 a 10.000 muestras, tales como 3.000 a 7.000 muestras.

EE20. Procesador de audio según EE19, en el que la unidad de expansión de onda plana determina solo una dirección dominante en cada banda de frecuencia para cada marco de tiempo.

25 EE21. Dispositivo que comprende un procesador de audio según cualquiera de las reivindicaciones anteriores, tal como el dispositivo que es uno de: un dispositivo para grabar señales de sonido o vídeo, un dispositivo para reproducir señales de sonido o vídeo, un dispositivo portátil, un dispositivo informático, un dispositivo de videojuegos, un dispositivo de alta fidelidad, un dispositivo convertidor de audio y una unidad de auriculares.

30 EE22. Método para convertir una señal de entrada de audio multicanal que comprende al menos dos, tal como dos, tres o cuatro canales, tal como una señal estéreo o una señal de campo de sonido de formato B, en un conjunto de señales de salida de audio, tal como conjunto de dos señales de salida de audio (L, R) dispuestas para la reproducción de auriculares o dos o más señales de salida de audio dispuestas para su reproducción a través de una matriz de altavoces, comprendiendo el método

- 35 - separar la señal de entrada de audio en una pluralidad de bandas de frecuencia, tal como bandas de frecuencia parcialmente solapadas,
 - realizar una separación de fuente de sonido que comprende
- 40 - realizar un cálculo de expansión de onda plana en la señal de entrada de audio multicanal para determinar al menos una dirección dominante correspondiente a una dirección de una fuente de sonido dominante en la señal de entrada de audio,
 - decodificar la señal de entrada de audio en una serie de canales de salida, en el que dicha decodificación se controla según dicha al menos una dirección dominante, y
- 45 - sumar las señales resultantes de los canales de salida respectivos para que la al menos parte de la pluralidad de bandas de frecuencia llegue al conjunto de señales de salida de audio.

50 EE23. Método según EE22, en el que dicha etapa de decodificación de la señal de entrada en el número de canales de salida representa

- determinar una matriz de al menos dos, tal como cuatro, posiciones de altavoz virtual seleccionadas de modo que una o más de las posiciones de altavoz virtuales coincidan al menos sustancialmente, tal como coincide exactamente, con la al menos una dirección dominante,
 - decodificar la señal de entrada de audio en señales de altavoces virtuales correspondientes a cada una de las posiciones de altavoces virtuales, y
 - aplicar una función de transferencia adecuada a las señales del altavoz virtual para mapear espacialmente las posiciones del altavoz virtual en el número de canales de salida que representan direcciones espaciales fijas.

60 EE24. Método según EE22 o EE23, que comprende

- calcular los parámetros necesarios para expandir el campo de sonido local en dos ondas planas o determinar al menos una o dos direcciones de llegada estimadas,
 - complementar las direcciones estimadas con direcciones fantasma de modo que un número total sea igual al número de canales de entrada,
 - calcular una matriz decodificadora adecuada para descomponer la señal de entrada en alimentadores

virtuales, colocando los altavoces virtuales en las direcciones calculadas por la expansión de ondas planas y en las direcciones fantasma,

- seleccionar una matriz de funciones de transferencia adecuadas para crear una ilusión de sonido que emana de las direcciones de dichos altavoces virtuales
- multiplicar la matriz de decodificación por la matriz de funciones de transferencia
- multiplicar la matriz resultante por el vector de señales de entrada
- sumar el vector resultante en todas las bandas de frecuencia para producir un conjunto de señales de audio de salida.

En resumen, la invención proporciona un procesador de audio para convertir una señal de entrada de audio multicanal, tal como una señal de campo de sonido de formato B, en un conjunto de señales de salida de audio (L, R), tal como un conjunto de dos o más señales de salida de audio dispuestas para la reproducción de auriculares o para la reproducción a través de una matriz de altavoces. Un banco de filtros divide cada uno de los canales de entrada en bandas de frecuencia. La señal de entrada se descompone en ondas planas para determinar una o dos direcciones de fuente de sonido dominante. Se utilizan para determinar un conjunto de posiciones de altavoces virtuales seleccionados de modo que una o dos de las posiciones de altavoces virtuales coincidan con una o ambas direcciones dominantes. La señal de entrada se decodifica en señales de altavoces virtuales correspondientes a cada una de las posiciones de altavoces virtuales, y las señales de altavoces virtuales se procesan con funciones de transferencia adecuadas para crear la ilusión de sonido que emana de las direcciones de los altavoces virtuales. Se obtiene una alta fidelidad espacial debido a la coincidencia de las posiciones del altavoz virtual y la(s) dirección(es) de la fuente de sonido dominante determinada.

En las reivindicaciones, el término "que comprende" no excluye la presencia de otros elementos o etapas. Adicionalmente, aunque las características individuales pueden incluirse en diferentes reivindicaciones, éstas pueden posiblemente combinarse ventajosamente, y la inclusión en diferentes reivindicaciones no implica que una combinación de características no sea factible y/o ventajosa. Además, las referencias singulares no excluyen una pluralidad. Por lo tanto, las referencias a "un", "una", "primero", "segundo", etc. no excluyen una pluralidad. Signos de referencia se incluyen en las reivindicaciones, sin embargo, la inclusión de los signos de referencia es solo por razones de claridad y no debe interpretarse como que limita el alcance de las reivindicaciones.

Las posiciones de altavoces virtuales, y las señales de altavoces virtuales se procesan con funciones de transferencia adecuadas para crear la ilusión de sonido que emana de las direcciones de los altavoces virtuales. Se obtiene una alta fidelidad espacial debido a la coincidencia de las posiciones del altavoz virtual y la(s) dirección(es) de la fuente de sonido dominante determinada.

En las reivindicaciones, el término "que comprende" no excluye la presencia de otros elementos o etapas. Adicionalmente, aunque las características individuales pueden incluirse en diferentes reivindicaciones, éstas pueden posiblemente combinarse ventajosamente, y la inclusión en diferentes reivindicaciones no implica que una combinación de características no sea factible y/o ventajosa. Además, las referencias singulares no excluyen una pluralidad. Por lo tanto, las referencias a "un", "una", "primero", "segundo", etc. no excluyen una pluralidad. Signos de referencia se incluyen en las reivindicaciones, sin embargo, la inclusión de los signos de referencia es solo por razones de claridad y no debe interpretarse como que limita el alcance de las reivindicaciones.

REIVINDICACIONES

- 5 1. Un procesador de audio dispuesto para convertir una señal de entrada de audio multicanal que comprende tres o cuatro canales, tal como una señal de campo de sonido de formato B, en un conjunto de señales de salida de audio, tal como un conjunto de dos señales de salida de audio dispuestas para auriculares o dos o más señales de salida de audio dispuestas para su reproducción a través de una matriz de altavoces, comprendiendo el procesador de audio
- 10 - un banco de filtros (FB) dispuesto para separar la señal de entrada en una pluralidad de bandas de frecuencia, tal como bandas de frecuencia parcialmente solapadas,
- 15 - una unidad de separación de fuente de sonido (SSS) que comprende, para al menos una parte de la pluralidad de bandas de frecuencia,
- una unidad de descomposición de onda plana paramétrica (PWD) para determinar al menos una dirección dominante correspondiente a una dirección de una fuente de sonido dominante en la señal de entrada de audio multicanal,
- 20 - una unidad de vértices opuestos (VLP) para determinar una matriz de dos o más, tal como dos, tres o cuatro posiciones de altavoces virtuales seleccionadas, de modo que una o más de las posiciones de altavoces virtuales coincidan al menos sustancialmente, tal como que coinciden exactamente con la al menos una dirección dominante,
- 25 - un decodificador para decodificar la señal de entrada de audio en señales de altavoces virtuales correspondientes a cada una de las posiciones de altavoces virtuales;
- un multiplicador para aplicar una función de transferencia adecuada a las señales de los altavoces virtuales para mapear espacialmente las posiciones de los altavoces virtuales en el número de canales de salida que representan direcciones espaciales fijas, y
- 30 - una unidad sumadora (SU) dispuesta para sumar las señales resultantes de los canales de salida respectivos para que la al menos parte de la pluralidad de bandas de frecuencia llegue al conjunto de señales de salida de audio.
- 30 2. Procesador de audio según la reivindicación 1, en el que el banco de filtros (FB, 1, 2, 3, 4) está dispuesto para separar cada uno de los canales de entrada de audio en una pluralidad de bandas de frecuencia, tal como bandas de frecuencia parcialmente solapadas,
- 35 en el que una unidad de descomposición de onda plana paramétrica (PWD, 5) está dispuesta para descomponer un campo de sonido local representado en los canales de entrada de audio en dos ondas planas o al menos determina una o dos direcciones de llegada estimadas,
- 40 en el que la unidad de vértices opuestos (VLP, 1) está dispuesta para complementar las direcciones estimadas con direcciones fantasma,
- 45 en el que un calculador de matriz decodificadora (6) está dispuesto para calcular una matriz decodificadora adecuada para descomponer la señal de entrada de audio en alimentaciones para altavoces virtuales, donde las direcciones de dichos altavoces virtuales están determinadas por las salidas combinadas de la unidad de descomposición de onda plana paramétrica y la unidad de vértices opuestos,
- 50 en el que un selector de función de transferencia (2) está dispuesto para calcular una matriz de funciones de transferencia panorámicas adecuadas, tal como funciones de transferencia relacionadas con la cabeza o funciones de panoramización por pares, para producir una ilusión de sonido que emana de las direcciones de dichos altavoces virtuales,
- 55 en el que una primera unidad de multiplicación de matriz (7) está dispuesta para multiplicar las salidas del calculador de matriz de decodificación y el selector de función de transferencia,
- en el que una segunda unidad de multiplicación de matriz (10) está dispuesta para multiplicar una salida de la batería de filtros por una salida de la primera unidad de multiplicación de matriz, tal como una salida de una unidad de suavizado que opera en la salida de la primera unidad de multiplicación de matriz, y
- 60 en el que una pluralidad de unidades de suma (11, 12) están dispuestas para sumar las señales respectivas en la pluralidad de bandas de frecuencia para producir el conjunto de señales de salida de audio.
- 65 3. Procesador de audio según las reivindicaciones 1 o 2, en el que el banco de filtros comprende al menos 20, tal como al menos 100, tal como al menos 500, tal como de 1000 a 5000, filtros parcialmente solapados que cubren un intervalo de frecuencia de 0 Hz a 22 kHz.
4. Procesador de audio según cualquiera de las reivindicaciones anteriores, en el que una unidad de suavizado está conectada entre la unidad de descomposición de onda plana paramétrica y al menos una unidad que recibe una salida de la unidad de descomposición de onda plana paramétrica, en el que la unidad de suavizado (7) está dispuesta para suprimir grandes diferencias en las estimaciones de dirección entre las bandas de frecuencia vecinas y rápidos cambios de dirección en el tiempo.
5. Procesador de audio según cualquiera de las reivindicaciones anteriores, en el que una primera unidad de multiplicación de matriz (10) está conectada para recibir una salida del banco de filtros y a una calculadora de matriz de decodificación (8), y en el que una segunda unidad de multiplicación de matriz (7) está conectada a la primera unidad de multiplicación de matriz y un selector de función de transferencia (2).

- 5 6. Procesador de audio según la reivindicación 5, en el que una unidad de suavizado (9) está conectada entre la primera y la segunda unidades de multiplicación de matrices, en el que la unidad de suavizado está dispuesta para suprimir grandes diferencias de fase o de amplitud entre elementos de matriz correspondientes en bandas de frecuencia vecinas y cambios rápidos de fase o amplitud de los elementos de la matriz en el tiempo.
- 10 7. Procesador de audio según cualquiera de las reivindicaciones anteriores, que comprende un selector de función de transferencia (2) que selecciona funciones de transferencia de una base de datos de funciones de transferencia relacionadas con la cabeza (HRTF, 5), produciendo así dos canales de salida adecuados para reproducción a través de auriculares.
- 15 8. Procesador de audio según la reivindicación 2, en el que un diferenciador de fase (3) calcula el retardo de grupo de las funciones de transferencia de panoramización, y en el que un integrador de retardo de grupo (7) restablece un desplazamiento de fase después de combinar componentes de funciones de transferencia de panoramización correspondientes a direcciones diferentes.
- 20 9. Procesador de audio según la reivindicación 8, en el que un segundo diferenciador de fase (5) calcula el retardo de grupo de las funciones de transferencia resultantes de la combinación de componentes de funciones de transferencia de panoramización desde diferentes direcciones y donde un fundido cruzado (6) selecciona la salida de este segundo diferenciador de fase a bajas frecuencias, tal como por debajo de 1,6 kHz, y selecciona el retardo de grupo combinado procedente del diferenciador de primera fase a altas frecuencias, tal como por encima de 2,0 kHz, y con una transición gradual entre sí, y donde el integrador de retardo de grupo opera una salida de este fundido cruzado.
- 25 10. Procesador de audio según cualquiera de las reivindicaciones anteriores, que comprende un selector de función de transferencia que selecciona funciones de transferencia de acuerdo con al menos uno de:
- 30 1) una ley de panoramización por pares, produciendo así dos o más canales de salida adecuados para la reproducción a través de una matriz horizontal de altavoces,
 2) una panoramización de amplitud de base vectorial, panoramización equivalente ambisónica o síntesis de campo de onda, produciendo así cuatro o más canales de salida adecuados para la reproducción a través de una matriz 3D de altavoces, y
 3) evaluando las funciones armónicas esféricas, produciendo así cinco o más canales de salida adecuados para la decodificación con un decodificador ambisónico de orden superior.
- 35 11. Procesador de audio según cualquiera de las reivindicaciones anteriores, en el que la señal de entrada de audio es una señal de campo de sonido de formato B de tres o cuatro canales.
- 40 12. Procesador de audio según cualquiera de las reivindicaciones anteriores, en el que la unidad de separación de fuente de sonido opera en entradas con un marco de tiempo que tiene un tamaño de 1.000 a 20.000 muestras, tal como 2.000 a 10.000 muestras, tal como 3.000 a 7.000 muestras, y en el que la unidad de descomposición de onda plana paramétrica determina solo una dirección dominante en cada banda de frecuencia para cada marco de tiempo.
- 45 13. Dispositivo que comprende un procesador de audio según cualquiera de las reivindicaciones anteriores, tal como el dispositivo que es uno de: un dispositivo para grabar señales de sonido o de video, un dispositivo para reproducir señales de sonido o de video, un dispositivo portátil, un dispositivo informático, un dispositivo de videojuegos, un dispositivo de alta fidelidad, un dispositivo convertidor de audio y una unidad de auriculares.
- 50 14. Método para convertir una señal de entrada de audio multicanal que comprende tres o cuatro canales, tal como una señal de campo de sonido de formato B, en un conjunto de señales de salida de audio, tal como un conjunto de dos señales de salida de audio (L, R) dispuestas para la reproducción en auriculares o dos o más señales de salida de audio dispuestas para su reproducción a través de una matriz de altavoces, comprendiendo el método
- 55 - separar la señal de entrada de audio en una pluralidad de bandas de frecuencia, tal como bandas de frecuencia parcialmente solapadas,
 - realizar una separación de fuente de sonido que comprende
- 60 - realizar un cálculo de descomposición de onda plana paramétrica en la señal de entrada de audio multicanal para determinar al menos una dirección dominante correspondiente a una dirección de una fuente de sonido dominante en la señal de entrada de audio,
 - determinar una matriz de dos o más, tal como dos, tres o cuatro posiciones de altavoces virtuales seleccionadas de modo que una o más de las posiciones de altavoces virtuales coincidan al menos sustancialmente, tal como que coinciden exactamente con la al menos una dirección dominante,
 - decodificar la señal de entrada de audio en señales de altavoces virtuales correspondientes a cada una de
- 65 - las posiciones de altavoces virtuales,
 - aplicar una función de transferencia adecuada a las señales de altavoces virtuales para mapear

especialmente las posiciones de altavoces virtuales en el número de canales de salida que representan direcciones espaciales fijas, y

- 5 - sumar las señales resultantes de los canales de salida respectivos para que la al menos parte de la pluralidad de bandas de frecuencia llegue al conjunto de señales de salida de audio.

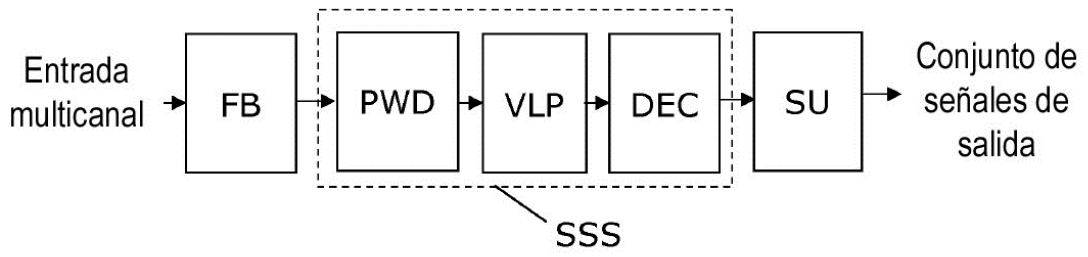


Fig. 1

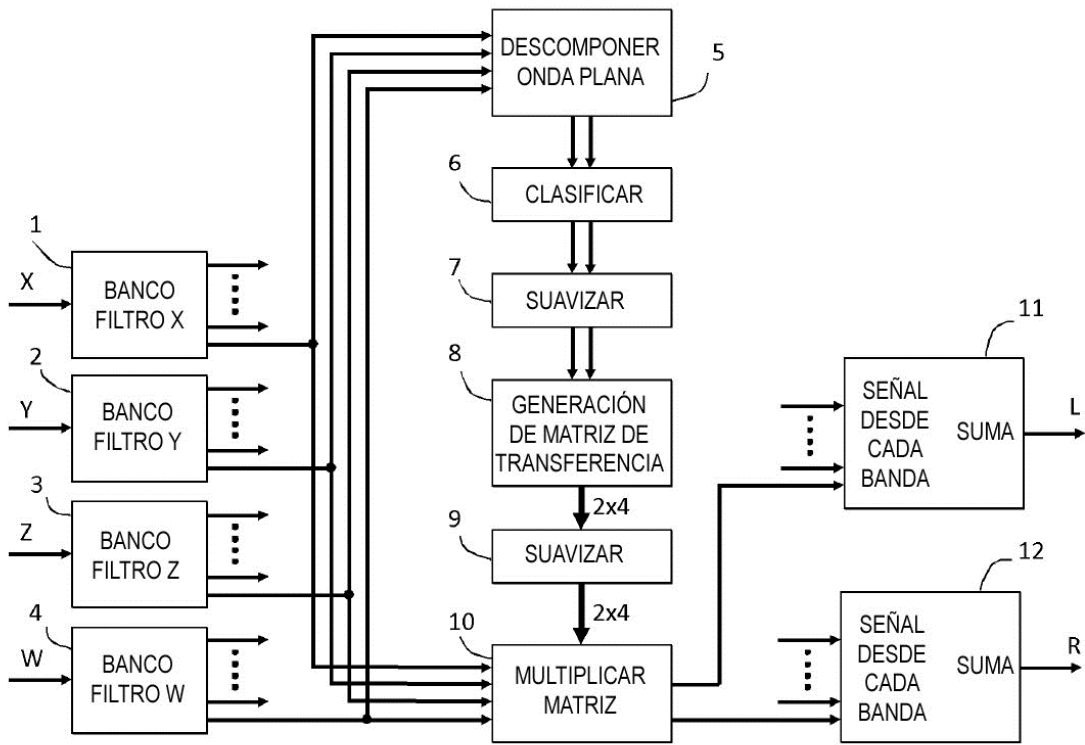


Fig. 2

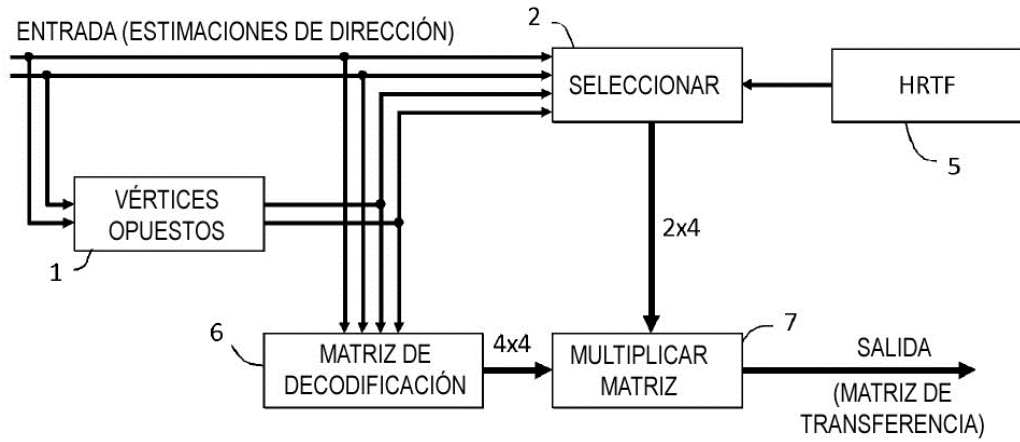


Fig. 3

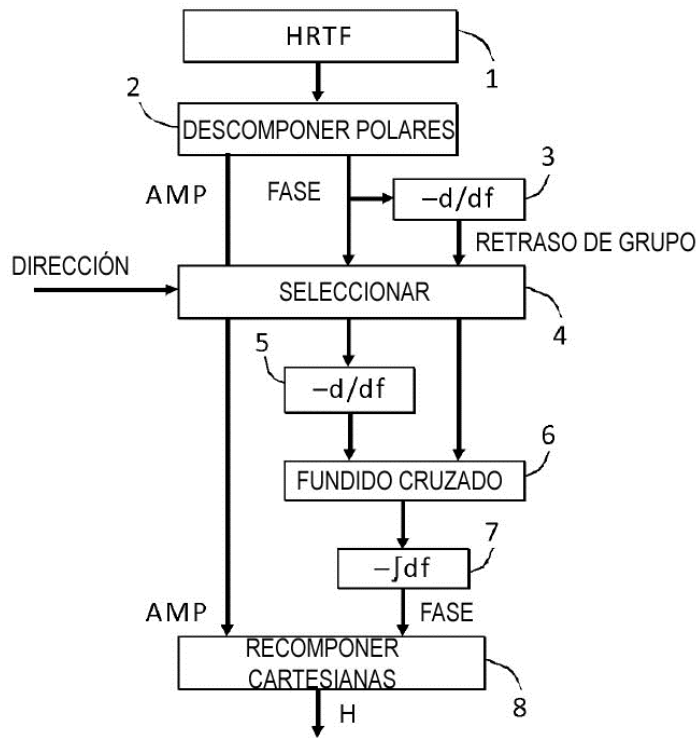


Fig. 4

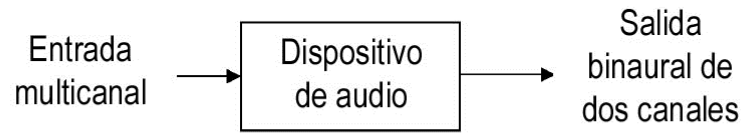


Fig. 5



Fig. 6