

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 376 889**

51 Int. Cl.:
H04S 3/00

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **06777145 .1**
96 Fecha de presentación: **01.09.2006**
97 Número de publicación de la solicitud: **1999999**
97 Fecha de publicación de la solicitud: **10.12.2008**

54 Título: **Generación de mezclas descendentes espaciales a partir de representaciones paramétricas de señales multicanal**

30 Prioridad:
24.03.2006 SE 0600674
10.04.2006 US 744555 P

45 Fecha de publicación de la mención BOPI:
20.03.2012

45 Fecha de la publicación del folleto de la patente:
20.03.2012

73 Titular/es:
Dolby International AB
Apollo Building, 3E Herikerbergweg 1-35
1101 CN Amsterdam Zuid-oost, NL y
Koninklijke Philips Electronics N.V.

72 Inventor/es:
VILLEMOES, Lars;
KJÖRLING, Kristofer y
BREEBAART, Jeroen

74 Agente/Representante:
Arizti Acha, Monica

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).

ES 2 376 889 T3

DESCRIPCIÓN

Generación de mezclas descendentes espaciales a partir de representaciones paramétricas de señales multicanal.

Campo de la Invención

- 5 La presente invención se refiere a la decodificación de señales de audio multicanal codificadas basadas en una representación multicanal paramétrica y en particular a la generación de mezclas descendentes de 2 canales que proporcionan una experiencia de escucha espacial, tal como por ejemplo una mezcla descendente compatible con auriculares o una mezcla descendente espacial para instalaciones de dos altavoces.

Antecedentes de la Invención

- 10 El desarrollo reciente en la codificación de audio ha hecho disponible la capacidad de recrear una representación multicanal de una señal de audio basada en una señal estéreo (o mono) y datos de control correspondientes. Estos métodos difieren sustancialmente de las soluciones basadas en matriz más antiguas tales como Dolby Prologic, puesto que se transmiten los datos de control adicionales para controlar la recreación, también denominada como mezcla ascendente, de los canales envolventes basados en los canales mono o estéreo transmitidos.

- 15 De ahí, un decodificador de audio multicanal paramétrico de este tipo, por ejemplo MPEG Surround, reconstruye N canales basándose en M canales transmitidos, donde $N > M$, y los datos de control adicionales. Los datos de control adicionales representan una velocidad de datos más baja significativa que la transmisión de todos los N canales, haciendo que la codificación sea muy eficiente a la vez que se asegura la compatibilidad tanto con los dispositivos de M canales como dispositivos de N canales.

- 20 Estos métodos de codificación envolventes paramétricos comprenden normalmente una parametrización de la señal envolvente basada en IID (*Inter channel Intensity Difference*- Diferencia de Intensidad Intercanal) o CLD (*Channel Level Difference*- Diferencia de Nivel de Canal) e ICC (*Inter Channel Coherence*- Coherencia Intercanal). Estos parámetros describen relaciones y correlaciones de potencia entre pares de canales en el proceso de mezcla ascendente. Parámetros adicionales también utilizados en la técnica anterior comprenden parámetros de predicción utilizados para predecir canales intermedios o de salida durante el procedimiento de mezcla ascendente.

- 25 Otros desarrollos en la reproducción de contenido de audio multicanal han proporcionado medios para obtener una impresión de escucha espacial utilizando auriculares estereofónicos. Para obtener una experiencia de escucha espacial utilizando solamente los dos altavoces de los auriculares, se mezclan de manera descendente las señales multicanal a señales estéreo utilizando HRTF (*Head-related transfer functions*- funciones de transferencia relacionados con la cabeza), destinadas a tener en cuenta las características de transmisión extremadamente complejas de la cabeza humana para proporcionar la experiencia de escucha espacial.

- 30 Otro enfoque relacionado es usar un ambiente de reproducción de 2 canales convencional y filtrar los canales de una señal de audio multicanal con filtros apropiados para obtener la experiencia de escucha cercana a aquella de la reproducción con el número original de altavoces. El procesamiento de las señales es similar al caso de la reproducción de auriculares para crear una "mezcla descendente estéreo espacial" apropiada que tiene las propiedades deseadas. En contraste con el caso de auriculares, la señal de ambos altavoces llega directamente a ambos oídos de un oyente, provocando "efectos de diafonía" indeseables. Ya que esto se tiene que tener en cuenta para la calidad de reproducción óptima, los filtros utilizados para el tratamiento de señales comúnmente se denominan filtros de cancelación de diafonía. 35 En general, el objetivo de esta técnica es extender el intervalo posible de fuentes de sonido al exterior de la base de altavoz estéreo mediante la cancelación de diafonía inherente utilizando filtros de cancelación de diafonía complejos.

- 40 Debido a la filtración compleja, los filtros HRTF son muy largos, es decir, pueden comprender varios cientos de derivaciones de filtro cada uno. Por la misma razón, es difícilmente posible encontrar una parametrización de los filtros que trabaje lo suficientemente bien para no degradar la calidad de percepción cuando se utilizan en lugar del filtro real.

- Así, por una parte, representaciones paramétricas de ahorro de bits de señales multicanal existen que permiten un transporte eficiente de una señal multicanal codificada. Por otra parte, se conocen maneras elegantes de crear una experiencia de escucha espacial para una señal multicanal cuando se usan auriculares estereofónicos o altavoces estereofónicos solamente. Sin embargo, estos requieren el número completo de canales de la señal multicanal como entrada para la aplicación de las funciones de transferencia relacionadas con la cabeza que crean la señal de mezcla descendente de auriculares. Así, o bien se tiene que transmitir el conjunto completo de señales multicanal o bien se tiene que reconstruir por completo una representación paramétrica antes de aplicar las funciones de transferencia relacionadas con la cabeza o los filtros de cancelación de diafonía, y así o bien el ancho de banda de transmisión o bien 45 la complejidad de cálculo es inaceptablemente alta.

50 La solicitud estadounidense US 2006/004274 se refiere a la generación de una señal de sonido mediante la aplicación de dos funciones de transferencia relacionadas con la cabeza a una señal de sonido monofónica transmitida. Se deriva cada una de las funciones de transferencia relacionadas con la cabeza añadiendo dos funciones de transferencia relacionadas con la cabeza adicionales.

La solicitud internacional WO2006/008683 describe un método y un dispositivo para procesar una señal estéreo obtenida de un codificador que codifica señales de audio de n canales en parámetros espaciales y una mezcla descendente estéreo.

- 5 En Faller C et al.: "Binaural Cue Coding – Part II: Schemes and Applications", se introduce un esquema de codificación destinado a transmitir múltiples canales de una manera eficiente con respecto a la velocidad de bits.

- 10 La solicitud estadounidense US 2003/0035553 se refiere a la codificación perceptiva compatible inversa de indicaciones espaciales para convertir dos señales de audio o más en una señal de audio combinada que se incrusta con dos conjuntos o más de un parámetro o más de escenas auditivas, en la que cada conjunto de parámetros de escena auditiva (por ejemplo, una o más indicaciones espaciales tal como ILD, ITD o funciones de transferencia relacionadas con la cabeza) corresponde a una banda de frecuencia diferente en la señal de audio combinada.

Sumario de la Invención

Es un objeto de la presente invención proporcionar un concepto que permita una reconstrucción más eficiente de una señal de dos canales que proporcione una experiencia de escucha espacial utilizando representaciones paramétricas de señales multicanal.

- 15 De acuerdo con un primer aspecto de la presente invención, se logra este objeto por un descodificador según la reivindicación 1 ó 19.

De acuerdo con un segundo aspecto de la presente invención, se logra este objeto por un descodificador binaural según la reivindicación 18.

- 20 De acuerdo con un tercer aspecto de la presente invención, se logra este objeto mediante un método de derivar una señal de mezcla descendente de auriculares según la reivindicación 20.

De acuerdo con un cuarto aspecto de la presente invención, se logra este objeto por un receptor o reproductor de audio según la reivindicación 21.

De acuerdo con un quinto aspecto de la presente invención, se logra este objeto por un método de recibir o reproducir audio según la reivindicación 22..

- 25 De acuerdo con un sexto aspecto de la presente invención, se logra este objeto por un programa informático según la reivindicación 23.

- 30 La presente invención se basa en el descubrimiento de que una señal de mezcla descendente de auriculares puede derivarse de una mezcla descendente paramétrica de una señal multicanal cuando se usa un calculador de filtro para derivar HRTF (funciones de transferencia relacionadas con la cabeza) modificadas a partir de HRTF originales de la señal multicanal y cuando el convertidor de filtro utiliza un parámetro de nivel que tiene información sobre una relación de nivel entre dos canales de la señal multicanal, de tal manera que las HRTF modificadas se influyen más por la HRTF de un canal que tiene un nivel más alto que por la HRTF de un canal que tiene un nivel más bajo. Se derivan las HRTF modificadas durante el proceso de descodificación teniendo en cuenta la intensidad relativa de los canales asociados con las HRTF. Las HRTF originales se modifican de tal manera que una señal de mezcla descendente de una representación paramétrica de una señal multicanal puede utilizarse directamente para sintetizar la señal de mezcla descendente de auriculares sin la necesidad de una reconstrucción multicanal paramétrica completa de la señal de mezcla descendente paramétrica.

- 35 En una realización de la presente invención, se utiliza un descodificador de la invención que implementa una reconstrucción multicanal paramétrica así como una reconstrucción binaural de la invención de una mezcla descendente paramétrica transmitida de una señal multicanal original. De acuerdo con la presente invención, no se requiere una reconstrucción completa de la señal multicanal antes de la mezcla descendente binaural, lo cual tiene la mayor ventaja obvia de una complejidad de cálculo fuertemente reducida. Esto permite, por ejemplo, que dispositivos móviles que sólo tienen reservas de energía limitadas prolonguen la duración de reproducción significativamente. Una ventaja adicional es que el mismo dispositivo puede servir como proveedor para señales multicanal completas (por ejemplo, señales 5,1, 7,1, 7,2) así como para la mezcla descendente binaural de la señal que tiene una experiencia de escucha espacial aún cuando se utilizan auriculares de sólo dos altavoces. Esto podría, por ejemplo, ser extremadamente ventajoso en configuraciones de entretenimiento en casa.

- 40 En una realización adicional de la presente invención, se utiliza un calculador de filtro para derivar las HRTF modificadas no sólo operativas para combinar las HRTF de los dos canales al aplicar factores de ponderación individuales a la HRTF, sino al introducir factores de fase adicionales para cada HTRF a combinar. La introducción del factor de fase tiene la ventaja de obtener una compensación de retardo de los dos filtros antes de su superposición o combinación. Esto conduce a una respuesta combinada que modela un tiempo de retardo principal correspondiente a una posición intermedia entre los altavoces frontal y trasero.

Una segunda ventaja es que un factor de ganancia, que tiene que aplicarse durante la combinación de los filtros para

- 5 asegurar la conservación de energía, es mucho más estable con respecto a su comportamiento con frecuencia que sin la introducción del factor de fase. Esto es particularmente relevante para el concepto de la invención, ya que de acuerdo con una realización de la presente invención se procesa una representación de una mezcla descendente de una señal multicanal dentro de un dominio de banco de filtros para derivar la señal de mezcla descendente de auriculares. Como tal, hay que procesar individualmente diferentes bandas de frecuencia de la representación de la señal de mezcla descendente y por lo tanto, un comportamiento uniforme de las funciones de ganancia aplicadas individualmente es vital.
- 10 En una realización adicional de la presente invención, se convierten las funciones de transferencia relacionadas con la cabeza a filtros de subbanda para los dominios de subbanda de tal manera que el número total de las HRTF modificadas utilizadas en el dominio de subbanda es menor que el número total de HRTF originales. Esto tiene la ventaja evidente de que la complejidad de cálculo para derivar las señales de mezcla descendente de auriculares aún se disminuye en comparación con la mezcla descendente utilizando filtros de HRTF estándar.
- 15 La implementación del concepto de la invención permite el uso de HTRF extremadamente largas y así permite la reconstrucción de señales de mezcla descendente de auriculares en base a una representación de una mezcla descendente paramétrica de una señal multicanal con una calidad de percepción excelente.
- Además, el uso del concepto de la invención en filtros de cancelación de diafonía permite la generación de una mezcla descendente estéreo espacial a utilizarse con una instalación de dos altavoces estándar en base a una representación de una mezcla descendente paramétrica de una señal multicanal con una calidad de percepción excelente.
- 20 Una gran ventaja adicional del concepto de descodificación de la invención es que puede utilizarse un solo descodificador binaural de la invención que implementa el concepto de la invención para derivar una mezcla descendente binaural así como una reconstrucción multicanal de una mezcla descendente transmitida teniendo en cuenta los parámetros espaciales transmitidos adicionalmente.
- 25 En una realización de la presente invención, un descodificador binaural de la invención tiene un banco de filtros de análisis para derivar la representación de la mezcla descendente de la señal multicanal en un dominio de subbanda y un descodificador de la invención que implementa el cálculo de las HRTF modificadas. El descodificador comprende además un banco de filtros de síntesis para finalmente derivar una representación de dominio de tiempo de una señal de mezcla descendente de auriculares, que está preparada para reproducirse mediante cualquier equipo de reproducción de audio convencional.
- 30 Se explican en más detalle en los siguientes párrafos unos esquemas de descodificación multicanal paramétricos y unos esquemas de descodificación binaurales con referencia a las figuras adjuntas para bosquejar más claramente las mayores ventajas del concepto de la invención.
- 35 La mayoría de las realizaciones de la presente invención detalladas a continuación describen el concepto de la invención utilizando las HRTF. Tal como se ha indicado anteriormente, el procesamiento de HTRF es similar al uso de filtros de cancelación de diafonía. Por lo tanto, debe entenderse que todas las realizaciones se refieren al procesamiento de HTRF, así como a filtros de cancelación de diafonía. Es decir, podrían sustituirse todos los filtros de HRTF por filtros de cancelación de diafonía a continuación para aplicar el concepto de la invención al uso de filtros de cancelación de diafonía.

Breve Descripción de las Figuras

- 40 A continuación se describen unas realizaciones preferidas de la presente invención con referencia a las figuras adjuntas, en las que:
- la figura 1 muestra una síntesis binaural convencional utilizando unas HRTF;
 - la figura 1b muestra el uso convencional de filtros de cancelación de diafonía;
 - la figura 2 muestra un ejemplo de un codificador espacial multicanal;
 - la figura 3 muestra un ejemplo de descodificadores espaciales/binaurales de la técnica anterior;
 - 45 la figura 4 muestra un ejemplo de un codificador multicanal paramétrico;
 - la figura 5 muestra un ejemplo de un descodificador multicanal paramétrico;
 - la figura 6 muestra un ejemplo de un descodificador de la invención;
 - la figura 7 muestra un diagrama de bloques que ilustra el concepto de transformar filtros al dominio de subbanda;
 - 50 la figura 8 muestra un ejemplo de un descodificador de la invención;

la figura 9 muestra un ejemplo adicional de un descodificador de la invención; y

la figura 10 muestra un ejemplo de un receptor o reproductor de audio de la invención.

Descripción Detallada de Realizaciones Preferidas

- 5 Las realizaciones que se describen a continuación son solamente ilustrativas para los principios de la presente invención para la descodificación binaural de señales multicanal mediante el filtrado de HRTF modificada. Se comprenderá que las modificaciones y variaciones de las disposiciones y los detalles descritos en la presente serán evidentes para otros expertos en la materia. Por lo tanto, se propone limitarse únicamente por el alcance de las reivindicaciones de patente pendientes y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones de la presente.
- 10 Con el fin de mejor describir los elementos y ventajas de la presente invención, a continuación se proporciona una descripción más elaborada de la técnica anterior.
- 15 Un algoritmo de síntesis binaural convencional se describe en la figura 1. Se filtra un conjunto de canales 10a, 10b, 10c, 10d y 10e de entrada (frontal izquierdo (LF), frontal derecho (RF), envolvente izquierdo (LS), envolvente derecho (RS) y central (C)) por un conjunto de HRTF 12a a 12j. Cada señal de entrada se divide en dos señales (un componente izquierdo "L" y un componente derecho "R"), donde posteriormente se filtra cada uno de estos componentes de señal por una HRTF correspondiente a la posición de sonido deseada. Finalmente, se suman todas las señales de oído izquierdo por un sumador 14a para generar la señal L de salida binaural izquierda y se suman las señales de oído derecho por un sumador 14b para generar la señal R de salida binaural derecha. Puede observarse que la convolución de HRTF puede realizarse principalmente en el dominio de tiempo, pero frecuentemente se prefiere que se realice el filtrado en el dominio de frecuencia debido a la eficiencia de cálculo incrementada. Esto significa que la suma mostrada en la figura 1 se realiza también en el dominio de frecuencia y además se requiere una transformación posterior a un dominio de tiempo.
- 20 La figura 1b ilustra el procesamiento de cancelación de diafonía destinado a obtener una impresión de escucha espacial utilizando sólo dos altavoces de un ambiente de reproducción estéreo estándar.
- 25 El objetivo es la reproducción de una señal multicanal mediante un sistema de reproducción estéreo que tiene sólo dos altavoces 16a y 16b, de tal manera que el oyente 18 tiene una experiencia de escucha espacial. Una diferencia principal con respecto a la reproducción con auriculares es que las señales de ambos altavoces 16a y 16b llegan directamente a ambos oídos del oyente 18. Por lo tanto, es necesario adicionalmente tener en cuenta las señales indicadas por las líneas discontinuas (diafonía).
- 30 Para facilitar la explicación, se ilustra en la figura 1b una señal de una entrada de sólo 3 canales que tiene 3 fuentes 20a a 20c. Resulta evidente que el escenario puede en principio extenderse a un número arbitrario de canales.
- 35 Para derivar la señal estéreo a reproducirse, se procesa cada fuente de entrada por 2 de los filtros 21a a 21f de cancelación de diafonía, un filtro para cada canal de la señal de reproducción. Finalmente, se suman todas las señales filtradas para el canal 16a de reproducción izquierdo y el canal 16b de reproducción derecho para la reproducción. Es evidente que los filtros de cancelación de diafonía en general serán diferentes para cada fuente 20a y 20b (dependiendo de su posición percibida deseada) y que podrían además hasta depender del oyente.
- 40 Debido a la alta flexibilidad del concepto de la invención, uno se beneficia de la alta flexibilidad en el diseño y aplicación de los filtros de cancelación de diafonía de tal manera que los filtros pueden optimizarse para cada aplicación o dispositivo de reproducción individualmente. Una ventaja adicional es que el método es extremadamente eficiente desde el punto de vista de cálculo, puesto que sólo se requieren dos bancos de filtros de síntesis.
- Se muestra en la figura 2 un esquema principal de un codificador de audio espacial. En tal escenario de codificación básico, un descodificador 40 de audio espacial comprende un codificador 42 espacial, un codificador 44 de mezcla descendente y un multiplexor 46.
- 45 Se analiza una señal 50 de entrada multicanal por el codificador 42 espacial, extrayendo los parámetros espaciales que describen las propiedades espaciales de la señal de entrada multicanal que tienen que transmitirse al lado del descodificador. La señal mezclada de forma descendente generada por el codificador 42 espacial puede ser, por ejemplo, una señal monofónica o estéreo, dependiendo de diferentes escenarios de codificación. A continuación el codificador 44 de mezcla descendente puede codificar la señal de mezcla descendente monofónica o estéreo utilizando cualquier esquema de codificación de audio mono o estéreo convencional. El multiplexor 46 crea un flujo de bits de salida al combinar los parámetros espaciales y la señal de mezcla descendente codificada al flujo de bits de salida.
- 50 La figura 3 muestra una posible combinación directa de un descodificador multicanal correspondiente al codificador de la figura 2 y un método de síntesis binaural tal como se muestra, por ejemplo, en la figura 1. Como se puede observar, el enfoque de la técnica anterior de combinar las características es simple y directo. La instalación comprende un desmultiplexor 60, un descodificador 62 de mezcla descendente, un descodificador 64 espacial y un sintetizador 66 binaural. Se desmultiplexa un flujo 68 de bits de entrada, dando como resultado unos parámetros 70 espaciales y un

- flujo de bits de señal de mezcla descendente. Este último flujo de bits de señal de mezcla descendente se descodifica por el descodificador 62 de mezcla descendente utilizando un descodificador mono o estéreo convencional. Se introduce la mezcla descendente descodificada, junto con los parámetros 70 ,espaciales, al descodificador 64 espacial que genera una señal 72 de salida multicanal que tiene las propiedades espaciales indicadas por los parámetros 70 espaciales. El tener una señal 72 multicanal completamente reconstruida, el enfoque de simplemente agregar un sintetizador 66 binaural para implementar el concepto de síntesis binaural de la figura 1 es directo. Por lo tanto, se utiliza la señal 72 de salida multicanal como entrada para el sintetizador 66 binaural que procesa la señal de salida multicanal para derivar la señal 74 de salida binaural resultante. El enfoque que se muestra en la figura 3 tiene por lo menos tres desventajas: es necesario calcular una representación de señal multicanal completa como una etapa intermedia, seguida por la convolución de HRTF y mezcla descendente en la síntesis binaural. Aunque la convolución de HRTF debe realizarse por cada canal dado que cada canal de audio puede tener una posición espacial diferente, esta es una situación indeseable desde el punto de vista de la complejidad. Así, la complejidad de cálculo es alta y se desperdicia energía.
- El descodificador espacial opera en un dominio de banco de filtros (QMF). Por otra parte, normalmente se aplica la convolución de HRTF en el dominio de FFT. Por lo tanto, es necesaria una cascada de un banco de filtros de síntesis de QMF multicanal, una transformada de DFT multicanal y una transformada de DFT inversa estéreo, dando como resultado un sistema con altas demandas de cálculo.
- Serán audibles y posiblemente realizados en la salida binaural (estéreo) unos artefactos de codificación creados por el descodificador espacial para crear una reconstrucción multicanal.
- Se proporciona una descripción aún más detallada de la codificación y descodificación multicanal en las figuras 4 y 5.
- El codificador 100 espacial mostrado en la figura 4 comprende un primer OTT (codificador 1 a 2) 102a, un segundo OTT 102b y un bloque 104 de TTT (codificador de 3 a 2). Se procesa una señal 106 de entrada multicanal que consiste en canales LF, LS, C, RF, RS (frontal izquierdo, envolvente izquierdo, central, frontal derecho y envolvente derecho) por el codificador 100 espacial. Los bloques de OTT reciben dos canales de audio de entrada cada uno y derivan un canal de salida de audio monofónico individual y parámetros espaciales asociados, teniendo los parámetros información sobre las propiedades espaciales de los canales originales entre sí o con respecto al canal de salida (por ejemplo parámetros CLD, ICC). Se procesan en el codificador 100 los canales de LF y LS por el codificador 102a de OTT y se procesan los canales de RF y RS por el codificador 102b de OTT. Se generan dos señales, L y R, una que sólo tiene información sobre el lateral izquierdo y la otra que sólo tiene información sobre el lateral derecho. Se procesan adicionalmente las señales L, R y C por el codificador 104 de TTT, generando una mezcla descendente estéreo y parámetros adicionales.
- Los parámetros resultantes del codificador de TTT normalmente consisten en un par de coeficientes de predicción para cada banda de parámetros, o un par de diferencia de nivel para describir las relaciones de energía de las tres señales de entrada. Los parámetros de los codificadores de "OTT" consisten en diferencias de nivel y valores de coherencia o de correlación cruzada entre las señales de entrada para cada banda de frecuencia.
- Puede observarse que aunque el dibujo esquemático del codificador 100 espacial apunta a un procesamiento secuencial de los canales individuales de la señal de mezcla descendente durante la codificación, también es posible implementar el proceso de mezcla descendente completo del codificador 100 dentro de una sola operación de matriz.
- La figura 5 muestra un descodificador espacial correspondiente, que recibe como entrada las señales de mezcla descendente, tal como se proporcionan por el codificador de la figura 4 y los parámetros espaciales correspondientes.
- El descodificador 120 espacial comprende un descodificador 122 de 2 a 3 y descodificadores 124a a 124c de 1 a 2. Se introducen las señales L_0 y R_0 de mezcla descendente al descodificador 122 de 2 a 3 que recrea un central C canal, un canal R derecho y un canal L izquierdo. Estos tres canales se procesan adicionalmente por los descodificadores 124a a 124c de OTT produciendo seis canales de salida. Puede observarse que la derivación de un canal de mejora de baja frecuencia LFE no es determinante y puede omitirse de tal manera que un solo codificador de OTT puede conservarse dentro del descodificador 120 envolvente mostrado en la figura 5.
- De acuerdo con una realización de la presente invención, se aplica el concepto de la invención en un descodificador tal como se muestra en la figura 6. El descodificador 200 de la invención comprende un descodificador 104 de 2 a 3 y seis filtros 106a a 106f de HRTF. Se procesa una señal (L_0 , R_0) de entrada estéreo por el descodificador 104 de TTT, derivando tres señales L, C y R. Puede observarse que se supone que se alimenta la señal de entrada estéreo dentro de un dominio de subbanda, puesto que codificador de TTT puede ser el mismo codificador que se muestra en la figura 5 y de ahí adaptado para ser operativo en señales de subbanda. Se someten las señales L, R y C a un procesamiento de parámetros de HRTF por los filtros 106a a 106f de HRTF.

Se suman los 6 canales resultantes para generar el par (L_b , R_b) de salida binaural estéreo.

Puede describirse el descodificador 106 de TTT como la siguiente operación de matriz:

$$\begin{bmatrix} L \\ R \\ C \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{bmatrix} \begin{bmatrix} L_0 \\ R_0 \end{bmatrix},$$

5 con entradas m_{xy} de matriz dependientes de los parámetros espaciales. La relación de los parámetros espaciales y las entradas de matriz es idéntica a aquellas relaciones, como en el decodificador envolvente de MPEG de multicanal de 5.1. Se divide cada una de las tres señales L , R y C resultantes en dos y se procesan con parámetros de HRTF correspondientes a la posición deseada (percibida) de estas fuentes de sonido. Para el canal central (C), pueden aplicarse directamente los parámetros espaciales de la posición de la fuente de sonido, dando como resultado dos señales de salida para el centro, $L_B(C)$ y $R_B(C)$:

$$10 \quad \begin{bmatrix} L_B(C) \\ R_B(C) \end{bmatrix} = \begin{bmatrix} H_L(C) \\ H_R(C) \end{bmatrix} C.$$

15 Para el canal izquierdo (L), se combinan los parámetros de HRTF de los canales frontal izquierdo y envolvente izquierdo en un solo conjunto de parámetros de HRTF, utilizando los pesos w_{lf} y w_{lf} . Los parámetros de HTRF “combinados” resultantes simulan el efecto de los canales tanto frontales como envolventes en un sentido estadístico. Se utilizan las siguientes ecuaciones para generar el par (L_B , R_B) de salida binaural para el canal izquierdo:

$$\begin{bmatrix} L_B(L) \\ R_B(L) \end{bmatrix} = \begin{bmatrix} H_L(L) \\ H_R(L) \end{bmatrix} L,$$

20 Se obtiene de manera similar la salida binaural para el canal derecho de acuerdo con:

$$\begin{bmatrix} L_B(R) \\ R_B(R) \end{bmatrix} = \begin{bmatrix} H_L(R) \\ H_R(R) \end{bmatrix} R,$$

25 Dadas las definiciones anteriores de $L_B(C)$, $R_B(C)$, $L_B(L)$, $R_B(L)$, $L_B(R)$ y $R_B(R)$, pueden derivarse las señales de L_B y R_B completas a partir de una sola matriz de 2×2 dada la señal de entrada estéreo:

$$\begin{bmatrix} L_B \\ R_B \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} L_0 \\ R_0 \end{bmatrix},$$

con

$$30 \quad \begin{aligned} h_{11} &= m_{11}H_L(L) + m_{21}H_L(R) + m_{31}H_L(C), \\ h_{12} &= m_{12}H_L(L) + m_{22}H_L(R) + m_{32}H_L(C), \\ h_{21} &= m_{11}H_R(L) + m_{21}H_R(R) + m_{31}H_R(C), \\ h_{22} &= m_{12}H_R(L) + m_{22}H_R(R) + m_{32}H_R(C). \end{aligned}$$

35 En lo anterior, se supuso que los elementos $H_Y(X)$ para $Y = L_0, R_0$ y $X = L, R, C$, eran escalares complejos. Sin embargo, la presente invención enseña cómo extender el enfoque de un decodificador binaural de matriz de 2 por 2 para manejar filtros de HRTF de longitud arbitraria. Para conseguir esto, la presente invención comprende las siguientes etapas:

- Transformar las respuestas de filtro de HRTF a un dominio de banco de filtros;
- Extraer la diferencia de retardo global o la diferencia de fase de pares de filtros de HRTF;
- 40 • Modificar las respuestas del par de filtros de HRTF como función de los parámetros CLD
- Ajustar la ganancia.

Esto se consigue al reemplazar las seis ganancias $H_Y(X)$ complejas para $Y = L_0, R_0$ y $X = L, R, C$ con seis filtros. Se derivan estos filtros de los diez filtros $H_Y(X)$ para $Y = L_0, R_0$ y $X = Lf, Ls, Rf, Rs, C$, que describen las respuestas de filtro de HRTF dadas en el dominio de QMF. Pueden obtenerse estas representaciones de QMF de acuerdo con el método descrito en uno de los párrafos a continuación.

- 5 En otras palabras, la presente invención enseña un concepto para derivar HRTF modificadas al modificar los filtros de canal envolvente del extremo frontal utilizando una combinación lineal compleja de acuerdo con

$$H_Y(X) = gw_f \exp(-j\phi_{XY}w_s^2)H_Y(Xf) + gw_s \exp(j\phi_{XY}w_f^2)H_Y(Xs) .$$

- 10 Como se puede ver de la fórmula anterior, la derivación de las HRTF modificadas es una superposición ponderada de las HRTF originales, aplicando adicionalmente factores de fase. Los pesos w_s , w_f dependen de los parámetros CLD destinados a utilizarse por los descodificadores 124a y 124b de OTT de la figura 5.

Los pesos w_f y w_s dependen del parámetro CLD del bloque de "OTT" para Lf y Ls :

$$w_{lf}^2 = \frac{10^{CLD_1/10}}{1 + 10^{CLD_1/10}} ,$$

15

$$w_{ls}^2 = \frac{1}{1 + 10^{CLD_1/10}} .$$

Los pesos w_{rf} y w_{rs} dependen del parámetro CLD del bloque de "OTT" para Rf y Rs :

$$w_{rf}^2 = \frac{10^{CLD_1/10}}{1 + 10^{CLD_1/10}} ,$$

20

$$w_{rs}^2 = \frac{1}{1 + 10^{CLD_1/10}} .$$

Puede derivarse el parámetro de fase ϕ_{XY} de la diferencia de tiempo de retardo principal τ_{XY} entre los filtros de HRTF frontal y trasero y el índice de subbanda n del banco QMF:

25

$$\phi_{XY} = \frac{\pi(n + \frac{1}{2})}{64} \tau_{XY} .$$

- 30 El papel de este parámetro de fase en la modificación de filtros es doble. En primer lugar, realiza una compensación de retardo de los dos filtros antes de la superposición, lo que conduce a una respuesta combinada que modela un tiempo de retardo principal correspondiente a una posición fuente entre los altavoces frontal y trasero. En segundo lugar, hace que el factor de compensación de ganancia g necesario sea mucho más estable y que varíe lentamente sobre la frecuencia que en el caso de una superposición simple con $\phi_{XY} = 0$.

Se determina el factor de ganancia g por la regla de potencia de adición incoherente,

$$P_Y(X)^2 = w_f^2 P_Y(Xf)^2 + w_s^2 P_Y(Xs)^2 ,$$

donde

$$P_Y(X)^2 = g^2 \left(w_f^2 P_Y(Xf)^2 + w_s^2 P_Y(Xs)^2 + 2w_f w_s P_Y(Xf) P_Y(Xs) \rho_{XY} \right)$$

y ρ_{XY} es el valor real de la correlación cruzada compleja normalizada entre los filtros

$$\exp(-j\phi_{XY}) H_Y(Xf) \text{ y } H_Y(Xs).$$

- 5 Para las ecuaciones anteriores, P denota un parámetro que describe un nivel promedio por banda de frecuencia para la respuesta de impulso del filtro especificado por los índices. Esta intensidad media es por supuesto se deriva una vez conocida la función de respuesta de filtro.

10 En el caso de una superposición simple con $\phi_{XY} = 0$, el valor de ρ_{XY} varía de manera errática y oscilatoria como función de la frecuencia, lo que conduce a la necesidad de un ajuste de ganancia extenso. En una implementación práctica, es necesario limitar el valor de la ganancia g y no puede evitarse una coloración espectral restante de la señal.

15 En contraste, el uso de modificación con una compensación de fase a base de retardo tal como se enseña por la presente invención conduce a un comportamiento uniforme de ρ_{XY} como función de la frecuencia. Este valor es frecuentemente aún cercano a uno para pares de filtro derivados de HRTF naturales puesto que difieren principalmente en retardo y amplitud y el propósito del parámetro de fase es tener en cuenta la diferencia de fase en el dominio de banco de filtros de QMF.

Una elección ventajosa alternativa del parámetro de fase ϕ_{XY} enseñado por la presente invención se da por el ángulo de fase de la correlación cruzada compleja normalizada entre los filtros

$$20 \quad H_Y(Xf) \text{ y } H_Y(Xs),$$

y el desenvolvimiento de los valores de fase con técnicas de desenvoltura estándar como función del índice de subbanda n del banco de QMF. Esta elección tiene la consecuencia de que ρ_{XY} nunca es negativo y de ahí la ganancia de compensación g satisface $1/\sqrt{2} \leq g \leq 1$ para todas las subbandas. Además esta elección de parámetro de fase permite la modificación de los filtros de canal frontal y envolvente en situaciones en las que una

25 diferencia de tiempo de retardo principal τ_{XY} no está disponible.

Para la realización de la presente invención tal como se describe anteriormente, se enseña transformar exactamente las HRTF en una representación eficiente de los filtros de HRTF dentro del dominio de QMF.

30 La figura 7 ilustra un dibujo principal del concepto para transformar exactamente filtros de tiempo-dominio en filtros dentro del dominio de subbanda que tienen el mismo efecto neto sobre una señal reconstruida. La figura 7 muestra un banco 300 de análisis complejo, un banco 302 de síntesis correspondientes al banco 300 de análisis, un convertidor 304 de filtro y un filtro 306 de subbanda.

35 Se proporciona una señal 310 de entrada para la cual se conoce un filtro 312 que tiene propiedades deseadas. El objetivo de la implementación del convertidor 304 de filtro es que la señal 314 de salida tenga las mismas características después del análisis por el banco 300 de filtros de análisis, el filtrado 306 y síntesis 302 de subbanda posterior como hubiese tenido cuando se filtra por el filtro 312 en el dominio de tiempo. La tarea de proporcionar un número de filtros de subbanda correspondientes a número de subbandas utilizadas se satisface por el convertidor 304 de filtro.

La siguiente descripción resume un método para implementar un filtro $h(v)$ de FIR dado en el dominio de subbanda de QMF complejo. Se muestra el principio de funcionamiento en la figura 7.

40 Aquí, el filtrado de subbanda es simplemente la aplicación de un filtro de FIR de valor complejo para cada subbanda, $n = 0, 1, \dots, L-1$ para transformar los índices c_n originales a sus contrapartes d_n filtradas de acuerdo con la siguiente fórmula:

$$d_n(k) = \sum_l g_n(l) c_n(k-l).$$

Obsérvese que esto se distingue de métodos bien conocidos desarrollados para tomar muestras críticamente de bancos de filtros puesto que aquellos métodos requieren el filtrado en multibanda con respuestas más largas. El componente clave es el convertidor de filtro, que convierte cualquier filtro de FIR de dominio de tiempo en los filtros de dominio de subbanda complejos. Puesto que se sobremuestra el dominio de subbanda de QMF complejo, no hay ningún conjunto canónico de filtros de subbanda para un filtro de dominio de tiempo dado. Diferentes filtros de subbanda pueden tener el mismo efecto neto de la señal de dominio de tiempo. Lo que se describirá en la presente es una solución aproximada particularmente atractiva que se obtiene al restringir el convertidor de filtro para ser un banco de análisis complejo similar al QMF.

Suponiendo que el prototipo de convertidor de filtro es de longitud $64K_H$, se transforma un filtro de FIR de derivación de $64K_H$ real en un conjunto de 64 filtros de subbanda de derivación $K_H + K_Q - 1$ complejos. Para $K_Q = 3$, se transforma un filtro de FIR de 1024 derivaciones en un filtrado en subbandas de 18 derivaciones con una calidad de aproximación de 50 dB.

Se calculan las derivaciones de filtro de subbanda a partir de la fórmula

$$g_n(k) = \sum_{v=-\infty}^{\infty} h(v + kL)q(v) \exp\left(-i \frac{\pi}{L} \left(n + \frac{1}{2}\right)v\right),$$

donde $q(v)$ es el filtro de prototipo de FIR derivado del filtro de prototipo de QMF. Como se puede observar, este es solo un análisis de banco de filtros complejo del filtro $h(v)$ dado.

En lo siguiente, se resumirá el concepto de la invención para una realización adicional de la presente invención, en la que está disponible una representación paramétrica multicanal para una señal multicanal que tiene cinco canales. Nótese que en esta realización particular de la presente invención, se modifican los diez filtros $V_{Y,X}$ de HRTF originales (tal como, por ejemplo, dado por una representación de QMF de los filtros 12a a 12j de la figura 1), a seis filtros $h_{v,x}$ para $Y = L, R$ y $X = L, R, C$.

Los diez filtros $V_{Y,X}$ para $Y = L, R$ y $X = FL, BL, FR, BR, C$ describen las respuestas de filtro de HRTF dadas en un dominio de QMF híbrido.

La combinación de los filtros de canales frontal y envolvente se realiza con una combinación lineal compleja de acuerdo con

$$\begin{aligned} \mathbf{h}_{L,C} &= \mathbf{v}_{L,C} \\ \mathbf{h}_{R,C} &= \mathbf{v}_{R,C} \\ \mathbf{h}_{L,L} &= g_{L,L} \sigma_{FL} \exp(-j\phi_{FL,BL}^L \sigma_{BL}^2) \mathbf{v}_{L,FL} + g_{L,L} \sigma_{BL} \exp(j\phi_{FL,BL}^L \sigma_{FL}^2) \mathbf{v}_{L,BL} \\ \mathbf{h}_{L,R} &= g_{L,R} \sigma_{FR} \exp(-j\phi_{FR,BR}^L \sigma_{BR}^2) \mathbf{v}_{L,FR} + g_{L,R} \sigma_{BR} \exp(j\phi_{FR,BR}^L \sigma_{FR}^2) \mathbf{v}_{L,BR} \\ \mathbf{h}_{R,L} &= g_{R,L} \sigma_{FL} \exp(-j\phi_{FL,BL}^R \sigma_{BL}^2) \mathbf{v}_{R,FL} + g_{R,L} \sigma_{BL} \exp(j\phi_{FL,BL}^R \sigma_{FL}^2) \mathbf{v}_{R,BL} \\ \mathbf{h}_{R,R} &= g_{R,R} \sigma_{FR} \exp(-j\phi_{FR,BR}^R \sigma_{BR}^2) \mathbf{v}_{R,FR} + g_{R,R} \sigma_{BR} \exp(j\phi_{FR,BR}^R \sigma_{FR}^2) \mathbf{v}_{R,BR} \end{aligned}$$

Se determinan los factores de ganancia $g_{L,L}, g_{L,R}, g_{R,L}, g_{R,R}$ por

$$g_{Y,X} = \left(\frac{\sigma_{FX}^2 CFB_{Y,X}^2 + \sigma_{BX}^2}{\sigma_{FX}^2 CFB_{Y,X}^2 + \sigma_{BX}^2 + 2\sigma_{FX}\sigma_{BX} ICCFB_{Y,X}^\phi} \right)^{1/2}$$

Se definen los parámetros $CFB_{Y,X}, ICCFB_{Y,X}^\phi$ y los parámetros de fases ϕ de la siguiente manera:

Se define un cociente promedio de nivel frontal/trasero por banda híbrida para los filtros de HRTF para $Y = L, R$ y $X = L, R$ por

$$(CFB_{Y,X})_k = \frac{\left(\sum_{l=0}^{L_q-1} |(v_{Y,FX})_k(l)|^2 \right)^{1/2}}{\left(\sum_{l=0}^{L_q-1} |(v_{Y,BX})_k(l)|^2 \right)^{1/2}}.$$

5 Además, entonces se definen los parámetros de fase $\phi_{FL,BL}^L, \phi_{FR,BR}^L, \phi_{FL,BL}^R, \phi_{FR,BR}^R$ para $Y=L,R$ y $X=L,R$ por

$$(CIC_{Y,X})_k = \left| (CIC_{Y,X})_k \right| \exp \left(j(\phi_{FX,BX}^Y)_k \right),$$

donde se definen las correlaciones cruzadas complejas $(CIC_{Y,X})_k$ por

$$10 \quad (CIC_{Y,X})_k = \frac{\sum_{l=0}^{L_q-1} (v_{Y,FX})_k(l) (v_{Y,BX})_k^*(l)}{\left(\sum_{l=0}^{L_q-1} |(v_{Y,FX})_k(l)|^2 \right)^{1/2} \left(\sum_{l=0}^{L_q-1} |(v_{Y,BX})_k(l)|^2 \right)^{1/2}}.$$

Se aplica una desenvoltura de fase a los parámetros de fase a lo largo del índice de subbanda k , de tal manera que el valor absoluto del incremento de fase de la subbanda k a la subbanda $k+1$ es menor o igual a π para $k=0,1,\dots$.

15 En casos en los que hay dos elecciones, $\pm\pi$, para el incremento, se elige el signo de incremento para una medición de fase en el intervalo $[-\pi, \pi]$. Finalmente, se definen las correlaciones cruzadas compensadas en fase normalizadas para $Y=L,R$ y $X=L,R$ por

$$(ICCFB_{Y,X}^\phi)_k = \left| (CIC_{Y,X})_k \right|.$$

20 Nótese que en el caso en el que se realiza el procesamiento multicanal dentro de un dominio de subbanda híbrido, es decir, en un dominio en el que las subbandas se descomponen adicionalmente en diferentes bandas de frecuencia, podría, por ejemplo, realizarse un mapeo de las respuestas de HRTF a los filtros de banda híbridos de la siguiente manera:

25 Como en el caso sin un banco de filtros híbrido, se convierten las diez respuestas de impulso de HRTF dadas de la fuente $X=FL,BL,FR,BR,C$ al objetivo $Y=L,R$ en filtros de subbanda de QMF de acuerdo con el método resumido anteriormente. El resultado son diez filtros de subbanda $\hat{v}_{Y,X}$ con componentes

$$(\hat{v}_{Y,X})_m(l)$$

para la subbanda de QMF $m=0,1,\dots,63$ y la ranura de tiempo de QMF $l=0,1,\dots,L_q-1$. Sea el mapeo de índice a

30 partir de la banda híbrida k a la banda QMF m denotada por $m=Q(k)$.

Entonces se definen los filtros de HRTF $\mathbf{v}_{Y,X}$ en dominio de banda híbrido por

$$(v_{Y,X})_k(l) = (\hat{v}_{Y,X})_{Q(k)}(l).$$

Para la realización específica descrita en los párrafos anteriores, puede implementarse la conversión de filtro de los
35 filtros de HRTF a dominios de QMF de la siguiente manera, dado un filtro de FIR $h(\nu)$ de longitud N_h a transferirse al dominio de subbanda de QMF complejo:

El filtrado en subbanda consiste de la aplicación individual de un filtro de FIR de valor complejo $h_m(l)$ para cada subbanda de QMF, $m = 0, 1, \dots, 63$. El componente clave es el convertidor de filtro, que convierte el filtro de FIR de dominio de tiempo $h(v)$ dado en los filtros de dominio de subbanda complejos $h_m(l)$. El convertidor de filtro es un banco de análisis complejo similar al banco de análisis de QMF. Su filtro prototipo $q(v)$ es de longitud 192. Se define una extensión con ceros del filtro de FIR de dominio de tiempo por

$$\tilde{h}(v) = \begin{cases} h(v), & v = 0, 1, \dots, N_h - 1; \\ 0, & \text{de otra manera,} \end{cases}$$

Entonces se dan los filtros de dominio de subbanda de longitud $L_q = K_h + 2$ donde $K_h = \lceil N_h / 64 \rceil$ para $m = 0, 1, \dots, 63$ y $l = 0, 1, \dots, K_h + 1$ por

$$h_m(l) = \sum_{v=0}^{191} \tilde{h}(v + 64(l - 2)) q(v) \exp\left(-j \frac{\pi}{64} \left(m + \frac{1}{2}\right) (v - 95)\right).$$

Aunque se ha detallado el concepto de la invención con respecto a una señal de mezcla descendente que tiene dos canales, es decir, una señal estéreo transmitida, la aplicación del concepto de la invención no se limita de ninguna manera a un escenario que tenga una señal de mezcla descendente estéreo.

En resumen, la presente invención se refiere al problema de utilizar filtros de cancelación de diafonía o filtros de HTRF largos para la producción binaural de señales multicanal paramétricas. La invención enseña nuevas maneras para extender el enfoque de HTRF paramétricas a la longitud arbitraria de filtros de HTRF.

La presente invención comprende las siguientes características:

- Multiplicar la señal de mezcla descendente estéreo por una matriz de 2×2 en la que cada elemento de matriz es un filtro de FIR o un longitud arbitraria (tal como se da por el filtro de HRTF);
- Derivar los filtros en la matriz de 2×2 mediante la modificación de los filtros de HRTF originales en base a los parámetros multicanal transmitidos;
- Calcular la modificación de los filtros de HRTF de tal manera que se obtiene la envolvente espectral y la energía global correctas.

La figura 8 muestra un ejemplo para un descodificador 300 de la invención para derivar una señal de mezcla descendente de auriculares. El descodificador comprende un calculador 302 de filtro y un sintetizador 304. El calculador de filtro recibe como primera entrada parámetros 306 de nivel y como segunda entrada HTRF (funciones de transferencia relacionadas con la cabeza) 308 para derivar HRTF codificadas 310 que tienen el mismo efecto neto sobre una señal cuando se aplican a la señal en el dominio de subbanda que las funciones 308 de transferencia relacionadas con la cabeza aplicadas en el dominio de tiempo. Las HRTF 310 modificadas sirven como primera entrada al sintetizador 304 que recibe como segunda entrada una representación de una señal 312 de mezcla descendente dentro de un dominio de subbanda. Se deriva la representación de la señal 312 de mezcla descendente por un codificador multicanal parámetro y destinado a utilizarse como base para la reconstrucción de una señal multicanal completa por un descodificador multicanal. De esta forma el sintetizador 404 es capaz de derivar una señal 314 de mezcla descendente de auriculares utilizando las HRTF 310 modificadas y la representación de la señal 312 de mezcla descendente.

Puede notarse que podrían proporcionarse las HRTF en cualquier representación paramétrica posible, por ejemplo como la función de transferencia asociada al filtro, como la respuesta de impulso del filtro o como una serie de coeficientes de derivación para un filtro de FIR.

Los ejemplos anteriores suponen que la representación de la señal de mezcla descendente ya está suministrada como una representación de un banco de filtros, es decir, como muestras derivadas por un banco de filtros. Sin embargo, en aplicaciones prácticas, normalmente se suministra y se transmite una señal de mezcla descendente de dominio de tiempo para permitir también una reproducción directa de la señal presentada en entornos de reproducción simples. Por lo tanto, en la figura 9, en una realización adicional de la presente invención, en la que un descodificador 400 compatible binaural comprende un banco 402 de filtros de análisis y un banco 404 de filtros de síntesis y un descodificador de la invención que podría ser por ejemplo el descodificador 300 de la figura 8. Las funcionalidades del descodificador y sus descripciones son aplicables en la figura 9 así como en la figura 8 y se omitirá la descripción del descodificador 300 en el siguiente párrafo.

El banco 402 de filtros de análisis recibe una mezcla descendente de una señal 406 multicanal tal como se crea por un

- codificador multicanal paramétrico. El banco 402 de filtros de análisis deriva la representación del banco de filtros de la señal 406 de mezcla descendente recibida que entonces se introduce en el descodificador 300 que deriva una señal 408 de mezcla descendente de auriculares, todavía dentro del dominio de banco de filtros. Es decir, se representa la mezcla descendente por una multitud de muestras o coeficientes dentro de las bandas de frecuencia introducidas por el
- 5 banco 402 de filtros de análisis. Por lo tanto, para proporcionar una señal 410 de mezcla descendente de auriculares final en el dominio de tiempo, se introduce la señal 408 de mezcla descendente de auriculares en el banco 404 de filtros de síntesis que deriva la señal 410 de mezcla descendente de auriculares, que está preparada para reproducirse por el equipo de reproducción estéreo.
- 10 La figura 10 muestra un receptor o reproductor 500 de audio de la invención que tiene un descodificador 501 de audio de la invención, una entrada 502 de flujo de bits y una salida 504 de audio.
- Puede introducirse un flujo de bits en la entrada 502 del receptor/reproductor 500 de audio de la invención. Entonces se descodifica el flujo de bits por el descodificador 501 y se emite o se reproduce la señal decodificada en la salida 504 del receptor/reproductor 500 de audio de la invención.
- 15 Aunque se han derivado algunos ejemplos en los párrafos precedentes para implementar el concepto de la invención que depende de una mezcla descendente estéreo transmitida, puede aplicarse también el concepto de la invención en configuraciones basadas en un solo canal de mezcla descendente monofónico o en más de dos canales de mezcla descendentes.
- Se proporciona en la descripción de la presente invención una implementación particular de la transferencia de funciones de transferencia relacionadas con la cabeza en el dominio de subbanda. Sin embargo, podrían utilizarse también otras técnicas de derivar los filtros de subbanda sin limitar el concepto de la invención.
- 20 Pueden derivarse también los factores de fase introducidos en la derivación de las HRTF modificadas por otros cálculos distintos a los presentados anteriormente.
- Aunque se muestra el concepto de la invención en particular para los filtros de HRTF y de cancelación de diafonía, puede utilizarse para otros filtros definidos para uno o más canales individuales de una señal multicanal para permitir una generación eficiente desde el punto de vista de cálculo de una señal de reproducción estéreo de alta calidad. Además los filtros no sólo se limitan a filtros destinados para modelar un entorno de escucha. Pueden utilizarse incluso los filtros que agregan componentes "artificiales" a una señal, tal como por ejemplo reverberación u otros filtros de distorsión.
- 25 Dependiendo de ciertos requerimientos de implementación de los métodos de la invención, pueden implementarse los métodos de la invención en hardware o en software. Puede realizarse la implementación utilizando un medio de almacenamiento digital, en particular un disco, DVD o un CD que tiene almacenadas en el mismo señales de control que pueden leerse electrónicamente, que cooperan con un sistema informático programable de tal manera que se realizan los métodos de la invención. En general, por lo tanto, la presente invención es un producto de programa de un programa informático con un código de programa almacenado en un portador que puede leerse por la máquina, siendo operativo el código de programa para realizar los métodos de la invención cuando el producto de programa informático se ejecuta en un ordenador. En otras palabras, los métodos de la invención son por lo tanto un programa informático que tiene un código de programa para realizar al menos uno de los métodos de la invención cuando el programa informático se ejecuta en un ordenador.
- 30 Mientras que lo anterior se ha mostrado y descrito en particular con referencia a realizaciones particulares de la misma, los expertos en la técnica entenderán que pueden realizarse varios cambios de forma y detalle.
- 40

REIVINDICACIONES

1. Descodificador para derivar una señal (314) de mezcla descendente de auriculares utilizando una representación de una mezcla descendente de una señal (312) multicanal y utilizando un parámetro (306) de nivel que tiene información sobre una relación de nivel entre dos canales de la señal multicanal y utilizando funciones (308) de transferencia relacionadas con la cabeza relacionadas con los dos canales de la señal multicanal, en el que un primer canal de los dos canales es un canal frontal del lateral izquierdo o derecho de la señal multicanal y un segundo canal de los dos canales es un canal trasero del mismo lateral, que comprende:
 - 5 un calculador (302) de filtro para derivar una función (310) de transferencia relacionada con la cabeza modificada $H_Y(X)$ mediante la ponderación de la función de transferencia relacionada con la cabeza $H_Y(X_f)$ del canal frontal y la función de transferencia relacionada con la cabeza $H_Y(X_s)$ del canal trasero utilizando el parámetro (306) de nivel, de tal manera que la función (310) de transferencia relacionada con la cabeza modificada $H_Y(X)$ se influye más por la función (308) de transferencia relacionada con la cabeza de un canal que tiene un nivel más alto que por la función (308) de transferencia relacionada con la cabeza de un canal que tiene un nivel más bajo utilizando la siguiente combinación lineal compleja:

$$H_Y(X) = gw_f \exp(-j\phi_{XY}w_s^2)H_Y(X_f) + gw_s \exp(j\phi_{XY}w_f^2)H_Y(X_s),$$
 donde g es un parámetro de fase, w_f y w_s son factores de ponderación derivados utilizando el parámetro (306) de nivel y y es un factor de ganancia común derivado utilizando el parámetro (306) de nivel; y
 - 20 un sintetizador (304) para derivar la señal (314) de mezcla descendente de auriculares utilizando la función (310) de transferencia relacionada con la cabeza modificada y la representación de la señal (312) de mezcla descendente.
 2. Descodificador según la reivindicación 1, en el que el calculador (302) de filtro es operativo de tal manera que el número de funciones (310) de transferencia relacionadas con la cabeza modificadas derivadas es menor que el número de funciones (308) de transferencia relacionadas con la cabeza asociadas de los dos canales
 3. Descodificador según la reivindicación 1, en el que el calculador (302) de filtro es operativo para derivar una función (310) de transferencia relacionada con la cabeza modificada adaptada para aplicarse a una representación de banco de filtros de la señal de mezcla descendente.
 4. Descodificador según la reivindicación 1, adaptado para utilizar una representación de la señal de mezcla descendente derivada en un dominio de banco de filtros.
 5. Descodificador según la reivindicación 1, en el que el calculador (302) de filtro es operativo para derivar la función (310) de transferencia relacionada con la cabeza modificada utilizando funciones (308) de transferencia relacionadas con la cabeza caracterizadas por más de tres parámetros.
 6. Descodificador según la reivindicación 1, en el que el calculador (302) de filtro es operativo para derivar los factores de ponderación para las funciones (308) de transferencia relacionadas con la cabeza de los dos canales utilizando el mismo parámetro (306) de nivel.
 7. Descodificador según la reivindicación 6, en el que el calculador (302) de filtro es operativo para derivar un primer factor de ponderación w_f para un primer canal f y un segundo factor de ponderación w_s para un segundo canal s utilizando el parámetro de nivel CLD_1 de acuerdo con las siguientes formulas:

$$w_f^2 = \frac{10^{CLD_1/10}}{1 + 10^{CLD_1/10}},$$

$$w_s^2 = \frac{1}{1 + 10^{CLD_1/10}}.$$

9. Descodificador según la reivindicación 8, en el que el factor de ganancia común está dentro del intervalo de $[1/\sqrt{2}, 1]$.

5 10. Descodificador según la reivindicación 1, en el que el calculador (302) de filtro es operativo para derivar el parámetro de fase utilizando un tiempo de retardo entre respuestas de impulso de funciones (308) de transferencia relacionadas con la cabeza de los dos canales.

11. Descodificador según la reivindicación 10, en el que el calculador (302) de filtro es operativo en un dominio de banco de filtros que tiene n bandas de frecuencia y para derivar parámetros de fase individuales para cada banda de frecuencia utilizando el tiempo de retardo.

10 12. Descodificador según la reivindicación 10, en el que el calculador (302) de filtro es operativo en un dominio de banco de filtros que tiene más de dos bandas de frecuencia y para derivar parámetros de fase individuales para cada banda de frecuencia utilizando el tiempo de retardo de acuerdo con la siguiente fórmula:

$$\phi_{XY} = \frac{\pi(n + \frac{1}{2})}{64} \tau_{XY}.$$

15 13. Descodificador según la reivindicación 1, en el que el calculador (302) de filtro es operativo para derivar el parámetro de fase utilizando el ángulo de fase de la correlación cruzada compleja normalizada entre las respuestas de impulso de las funciones (308) de transferencia relacionadas con la cabeza del primer y del segundo canal.

20 14. Descodificador según la reivindicación 1, adaptado para utilizar una representación de una señal (312) de mezcla descendente que tiene un canal izquierdo y un canal derecho derivados de una señal multicanal que tiene un canal frontal izquierdo, un canal envolvente izquierdo, un canal frontal derecho, un canal envolvente derecho y un canal central.

15. Descodificador según la reivindicación 1, en el que el sintetizador es operativo para derivar canales de la señal (314) de mezcla descendente de auriculares aplicando una combinación lineal de las funciones (310) de transferencia relacionadas con la cabeza modificadas a la representación de la mezcla (312) descendente de la señal multicanal.

25 16. Descodificador según la reivindicación 15, en el que el sintetizador es operativo para utilizar coeficientes para la combinación lineal de las funciones (310) de transferencia relacionadas con la cabeza modificadas dependiendo del parámetro (306) de nivel.

17. Descodificador según la reivindicación 15, en el que el sintetizador (304) es operativo para utilizar coeficientes para la combinación lineal dependiendo de parámetros multicanal adicionales relacionados con propiedades espaciales adicionales de la señal multicanal.

30 18. Descodificador binaural, que comprende:

un decodificador de acuerdo con la reivindicación 1;

un banco (300) de filtros de análisis para derivar la representación de la mezcla descendente de la señal (312) multicanal mediante el filtrado en subbanda de la mezcla descendente de la señal multicanal; y

35 un banco (302) de filtros de síntesis para derivar una señal de auriculares de dominio de tiempo mediante la síntesis de la señal (314) de mezcla descendente de auriculares.

40 19. Descodificador para derivar una señal de mezcla descendente estéreo espacial utilizando una representación de una mezcla descendente de una señal (312) multicanal y utilizando un parámetro (306) de nivel que tiene información sobre una relación de nivel entre dos canales de la señal multicanal y utilizando filtros de cancelación de diafonía relacionados con los dos canales de la señal multicanal, en el que un primer canal de los dos canales es un canal frontal del lateral izquierdo o derecho de la señal multicanal y un segundo canal de los dos canales es un canal trasero del mismo lateral, que comprende:

un calculador (302) de filtro para derivar un filtro de cancelación de diafonía modificado $H_Y(X)$ mediante la ponderación de la función de transferencia relacionada con la cabeza $H_Y(Xf)$ del canal frontal y la función de transferencia relacionada con la cabeza $H_Y(Xs)$ del canal trasero de los dos canales utilizando el parámetro (306) de nivel, de tal manera que el filtro de cancelación de diafonía modificado $H_Y(X)$ se influye más por el filtro de cancelación de diafonía de un canal que tiene un nivel más alto que por el filtro de cancelación de diafonía de un canal que tiene un nivel más bajo utilizando la siguiente combinación lineal compleja:

$$H_Y(X) = gw_f \exp(-j\phi_{XY} w_s^2) H_Y(Xf) + gw_s \exp(j\phi_{XY} w_f^2) H_Y(Xs),$$

donde Φ_{xy} es un parámetro de fase, w_s y w_f son factores de ponderación derivados utilizando el parámetro (306) de nivel y g es un factor de ganancia común derivado utilizando el parámetro (306) de nivel; y

un sintetizador (304) para derivar la señal de mezcla descendente estéreo espacial utilizando los filtros de cancelación de diafonía modificados y la representación de la señal (312) de mezcla descendente.

- 5 20. Método de derivar una señal (314) de mezcla descendente de auriculares utilizando una representación de una mezcla descendente de una señal (312) multicanal y utilizando un parámetro (306) de nivel que tiene información sobre una relación de nivel entre dos canales de la señal multicanal y utilizando funciones (308) de transferencia relacionadas con la cabeza relacionadas con los dos canales de la señal multicanal, en el que un primer canal de los dos canales es un canal frontal del lateral izquierdo o derecho de la señal multicanal y un segundo canal de los dos canales es un canal trasero del mismo lateral, comprendiendo el método:
- 10

derivar, utilizando el parámetro (306) de nivel, una función (310) de transferencia relacionada con la cabeza modificada $H_Y(X)$ mediante la ponderación de la función de transferencia relacionada con la cabeza $H_Y(Xf)$ del canal frontal y la función de transferencia relacionada con la cabeza $H_Y(Xs)$ del canal trasero utilizando el parámetro (306) de nivel, de tal manera que la función de transferencia relacionada con la cabeza modificada $H_Y(X)$ se influye más por la función de transferencia relacionada con la cabeza de un canal que tiene un nivel más alto que por la función de transferencia relacionada con la cabeza de un canal que tiene un nivel más bajo utilizando la siguiente combinación lineal compleja:

15

$$H_Y(X) = gw_f \exp(-j\phi_{xy}w_s^2)H_Y(Xf) + gw_s \exp(j\phi_{xy}w_f^2)H_Y(Xs),$$

- 20 donde es un parámetro de fase, y son factores de ponderación derivados utilizando el parámetro (306) de nivel y es un factor de ganancia común derivado utilizando el parámetro (306) de nivel; y derivar la señal (314) de mezcla descendente de auriculares utilizando las funciones (310) de transferencia relacionadas con la cabeza modificadas y la representación de la señal de mezcla descendente.

21. Receptor o reproductor de audio que tiene un descodificador para derivar una señal (314) de mezcla descendente de auriculares según las reivindicaciones 1 a 17.
- 25

22. Método de recibir o reproducir audio, teniendo el método un método para derivar una señal (314) de mezcla descendente de auriculares según la reivindicación 20.

23. Programa informático que tiene un código de programa para realizar, cuando se ejecuta en un ordenador, uno de los métodos de las reivindicaciones 20 ó 22.

30

FIG 1

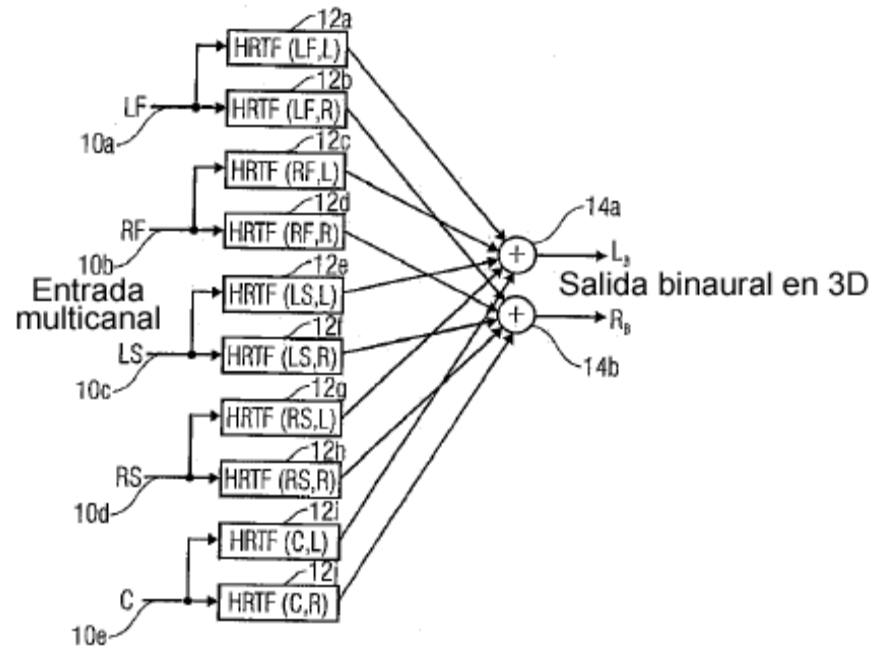


FIG 1b

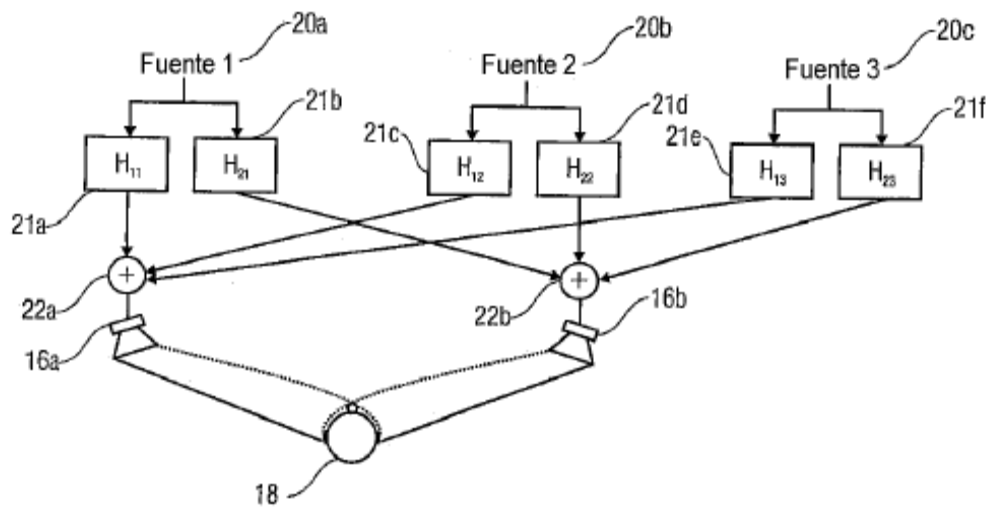


FIG 2

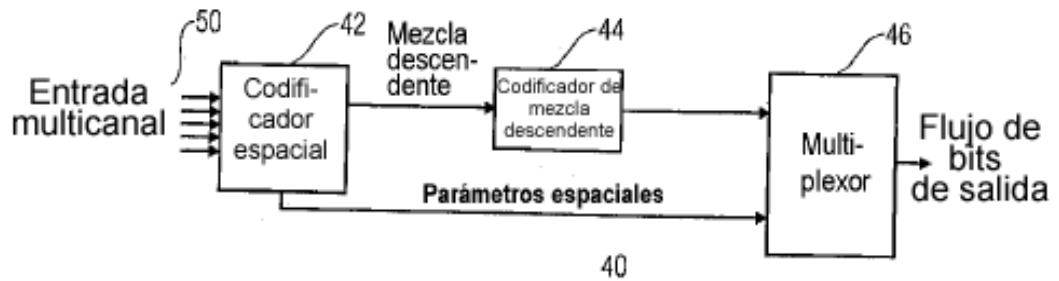


FIG 3

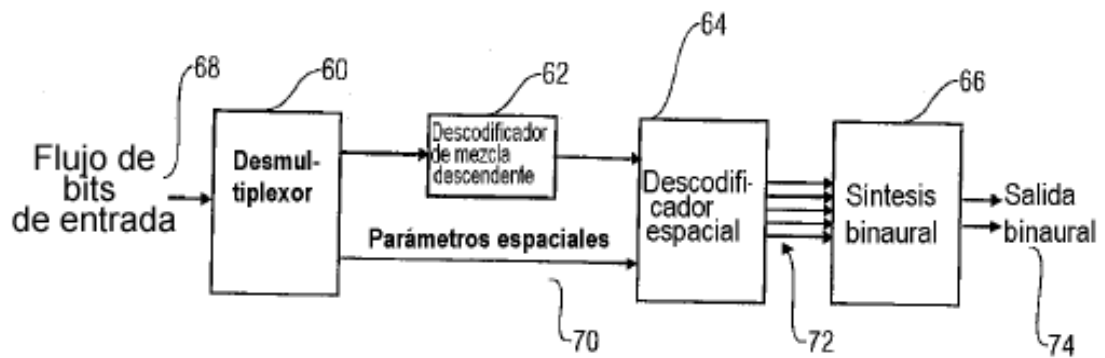


FIG 4

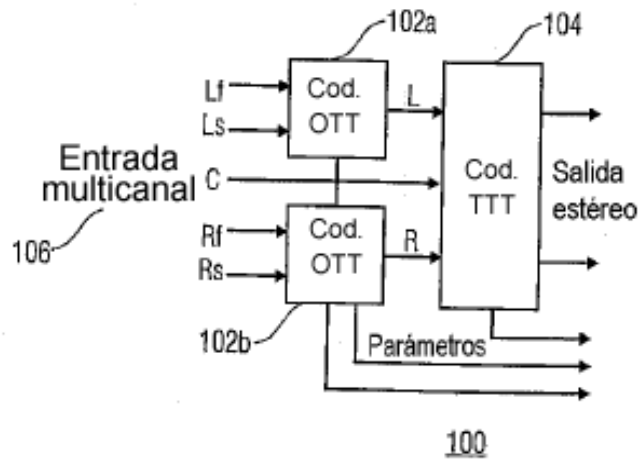


FIG 5

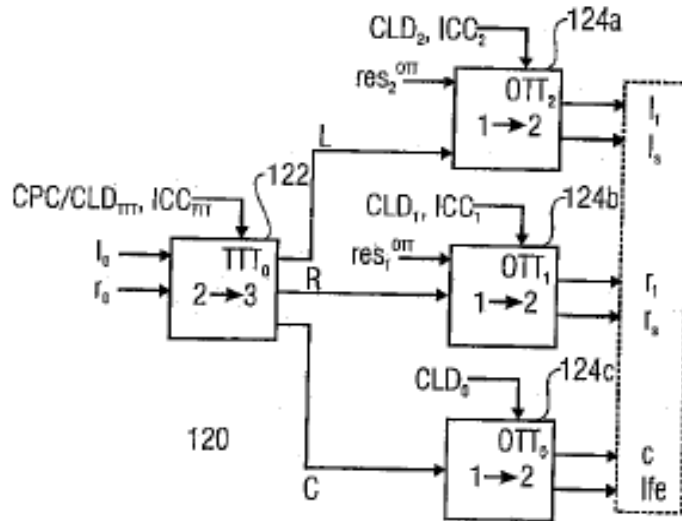


FIG 6

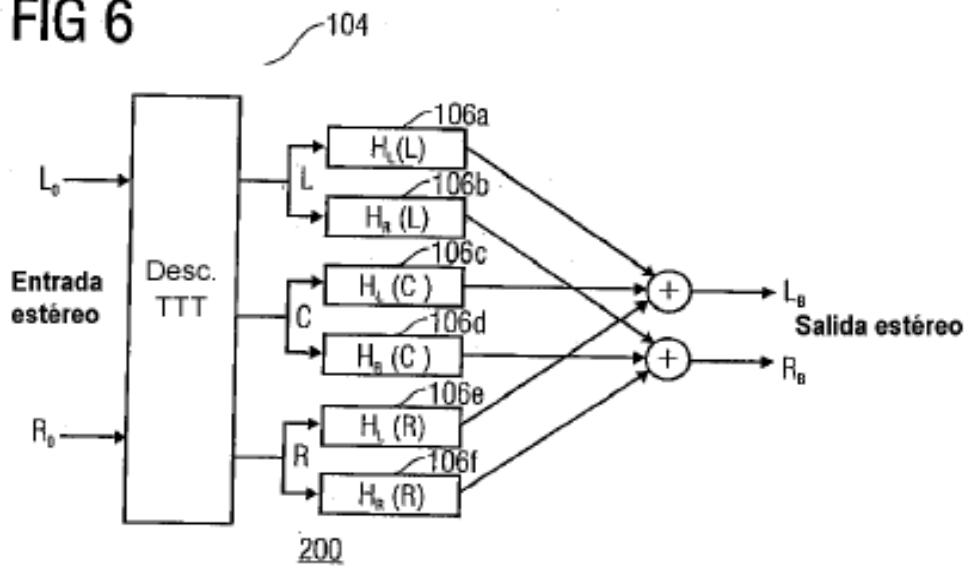


FIG 7

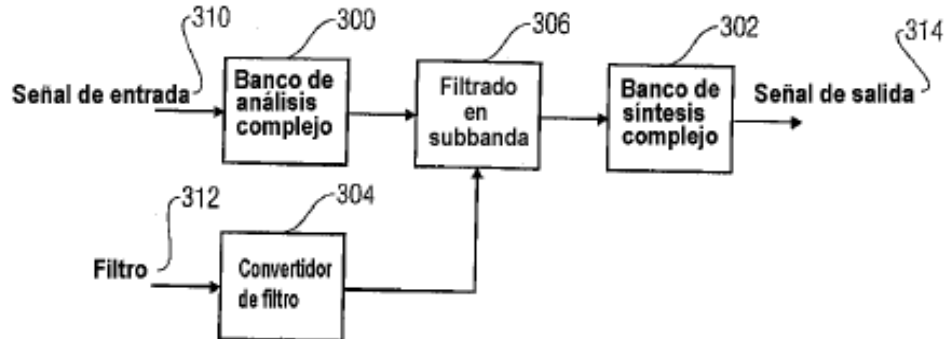


FIG 8

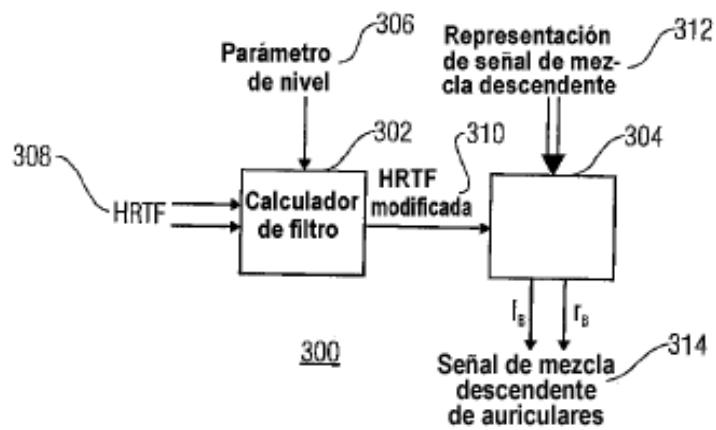


FIG 9

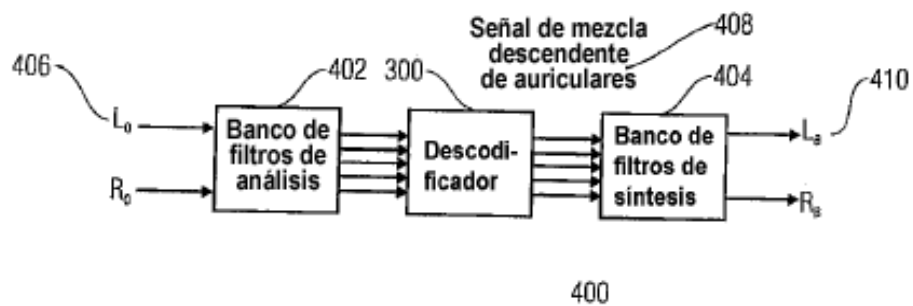


FIG 10

