

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 571 742**

51 Int. Cl.:

G10L 19/008 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.04.2012 E 12713720 (6)**

97 Fecha y número de publicación de la concesión europea: **02.03.2016 EP 2834814**

54 Título: **Método de determinación de un parámetro de codificación para una señal de audio multicanal y un codificador de audio multicanal**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
26.05.2016

73 Titular/es:

HUAWEI TECHNOLOGIES CO., LTD. (100.0%)
Huawei Administration Building, Bantian,
Longgang District
Shenzhen, Guangdong 518129, CN

72 Inventor/es:

VIRETTE, DAVID;
LANG, YUE y
XU, JIANFENG

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 571 742 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

Método de determinación de un parámetro de codificación para una señal de audio multicanal y un codificador de audio multicanal

5

CAMPO TÉCNICO

La presente invención se refiere a una codificación de audio y en particular, a una codificación de audio estéreo o multicanal paramétrica también conocida como codificación de audio espacial paramétrica.

10

ANTECEDENTES DE LA INVENCION

La codificación de audio multicanal o estéreo paramétrica se describe, a título de ejemplo, en C. Faller y F. Baumgarte, "Representación eficiente de una señal de audio espacial utilizando la parametrización perceptual" en Proc. IEEE Workshop sobre Aplicación de Proceso de Señal para Audio y Acústica, octubre 2001, páginas 199-202, en donde utiliza pistas espaciales para sintetizar señales de audio multicanal a partir de una mezcla descendente de normalmente señales de audio mono o estéreo, teniendo las señales de audio multicanal más canales que las señales de audio de mezcla descendente. En condiciones normales, las señales de audio de mezcla descendente resultan de una superposición de una pluralidad de señales de canales de audio de una señal de audio multicanal p.ej., de una señal de audio estéreo. Estos canales están codificados en formas de onda e información complementaria, esto es, las pistas espaciales, relacionadas con las relaciones de canales de señales originales se añade como parámetros de codificación a los canales de audio codificados. El decodificador utiliza esta información complementaria, para regenerar el número original de canales de audio sobre la base de los canales de audio codificados en forma de onda decodificados.

15

20

25

Un codificador estéreo paramétrico básico puede utilizar diferentes de nivel entre canales (ILD o CLD) como una pista necesaria para generar la señal estero a partir de la señal de audio de mezcla descendente monoaural. Codificadores más sofisticados pueden utilizar también la coherencia entre canales (ICC), que puede representar un grado de similitud entre las señales de canales de audio, esto es, canales de audio. Además, cuando se codifican las señales estéreo binaurales p.ej., para señales de audio en 3D o la entrega circundante basada en auriculares utilizando el filtrado de la función de transferencia relacionada con los auriculares (HRTF), una diferencia de tiempo inter-aural (ITD) puede desempeñar una función para reproducir diferencias de retardo entre los canales.

30

La diferencia de tiempo inter-aural (ITD) es la diferencia en el tiempo de llegada de un sonido 801 entre dos oídos 803, 805 que puede observarse en la ilustración de la Figura 8. Es importante para la localización de sonidos, puesto que proporciona una pista para identificar la dirección 807 o ángulo θ de incidencia de la fuente del sonido 801 (relativa al auricular 809). Si una señal llega a los oídos 803, 805 desde un solo lado, la señal tiene una ruta más larga 811 para alcanzar el oído lejano 803 (contralateral) y una ruta más corta 813 para alcanzar el oído cercano 805 (ipsilateral). Esta diferencia de longitudes de rutas da lugar a una diferencia temporal 815 entre la llegada de sonido a los oídos 803, 805, que se detecta y ayuda al proceso de identificación de la dirección 807 de la fuente de sonido 801.

35

40

La Figura 8 ilustra una realización, a modo de ejemplo, de ITD (indicada como Δt o diferencia temporal 815). Las diferencias en el tiempo de llegada en los dos oídos 803, 805 se indican por un retardo de la forma de onda del sonido. Si una forma de onda llega primero al oído izquierdo 803, el valor de ITD 815 es positivo y en caso contrario, es negativo. Si la fuente de sonido 801 está directamente en frente del oyente, la forma de onda llega al mismo tiempo a ambos oídos 803, 805 y por ello, el valor de ITD 815 es cero.

45

Las pistas de ITD son importantes para la mayor parte de la grabación en estéreo. A modo de ejemplo, la señal de audio binaural, que puede obtenerse a partir del registro real utilizando, a modo de ejemplo, un auricular ficticio o la síntesis binaural sobre la base del procesamiento de la Función de Transferencia Relacionada con los Auriculares (HRTF), se utiliza para el registro musical o las conferencias de audio. Por lo tanto, es un parámetro muy importante para un codificador-decodificador, códec, estéreo paramétrico de baja tasa binaria y en particular, para la aplicación conversacional es objetivo del codificador-decodificador. Un algoritmo de estimación de ITD de baja complejidad y estable se necesita para el codificador-decodificador estéreo paramétrico de baja tasa binaria. Además, el uso de parámetros de ITD, p.ej., en adición a otros parámetros tales como diferencias de nivel entre canales (CLDs o ILDs) y la coherencia entre canales (ICC), pueden aumentar la carga de tasa binaria. Para este escenario operativo de muy baja tasa binaria específico, solamente se puede transmitir un parámetro de banda completa ITD. Cuando se estima solamente un parámetro ITD de banda completa, la restricción sobre la estabilidad se hace todavía más difícil de conseguir.

50

55

60

Cuando un parámetro se estima utilizando una correlación cruzada, un espectro cruzado o una energía, el cambio rápido de la función de estimación puede dar lugar a una estimación inestable del parámetro. El parámetro estimado podría cambiar con demasiada rapidez y con demasiada frecuencia de una trama a otra, lo que no suele ser deseable. Este último puede ser el caso de si la magnitud de la trama es pequeña lo que puede dar lugar a un estimador no fiable de la correlación cruzada. El problema de la inestabilidad se percibirá como una fuente que

65

parece estar operativamente saltando desde el lado izquierdo al derecho y/o viceversa, aunque la fuente real no cambia su posición. El problema de la inestabilidad puede detectarse también por un oyente aun cuando la posición de la fuente no salte operativamente desde el lado izquierdo al lado derecho. Los cambios de posición de fuentes de pequeña magnitud, en el transcurso del tiempo, se perciben fácilmente por un oyente y entonces, deben evitarse cuando la fuente real está fijada.

A modo de ejemplo, la diferencia de tiempo inter-aural (ITD) es un parámetro importante para el codificador-decodificador estéreo paramétrico. Si la diferencia ITD se estima en el dominio de la frecuencia sobre la base del cálculo de una función de correlación cruzada, el valor de ITD estimado no suele ser estable durante tramas consecutivas, aun cuando la posición de la fuente de sonido esté fijada y el valor de ITD real sea estable. Los problemas de la estabilidad pueden resolverse aplicando una función de suavizado a la correlación cruzada antes de su uso para la estimación del parámetro ITD. Sin embargo, cuando se realiza el suavizado de la correlación cruzada, no se pueden seguir los cambios rápidos del valor de ITD real. Además, un suavizado estable reduce el comportamiento de seguimiento de cambios de ITD siguientes, con rapidez, cuando la fuente de sonido o la posición del oyente se desplazan con respecto entre sí.

El documento de F. Baumgarte y C. Faller titulado "Estimación de pistas espaciales auditoras para codificación de pistas binaurales", IEEE ICASSP 2002, 13 de mayo de 2002, da a conocer el ajuste empírico de una constante de tiempo fija $w=0.998$ para el suavizado exponencial para conseguir una solución de compromiso entre los cambios siguientes en la correlación de entrada con rapidez suficiente mientras se mantiene un resultado estable razonable para un parámetro ICLD o ICTD estacionario.

Otra realización, a modo de ejemplo, es la estimación de las diferencias de nivel de canales (CLD). CLD es un parámetro importante para el codificador-decodificador estéreo paramétrico. Si CLD se estima en el dominio de la frecuencia sobre la base del cálculo de la energía de cada contenedor o sub-banda, el valor de CLD estimado no suele ser estable durante tramas consecutivas, aun cuando la posición de la fuente de sonido esté fijada y la diferencia de niveles reales sea estable. Los problemas de la estabilidad pueden resolverse aplicando una función de suavizado a la energía antes de utilizarla para la estimación de CLD. Sin embargo, cuando se realiza la función de suavizado de la energía, no se pueden seguir los cambios rápidos del valor del CLD real con lo que se reduce el comportamiento de seguimiento de los cambios de CLD siguientes, con rapidez, cuando la fuente de sonido o la posición del oyente se desplazan entre sí.

Encontrar los coeficientes de suavizado correctos que permitan seguir, con rapidez, los cambios de parámetros ITD o CLD, mientras se mantiene el valor de ITD o CLD estable, ha resultado ser imposible, en particular cuando la función de correlación tiene una resolución deficiente, a modo de ejemplo, la resolución de frecuencia de una FFT.

SUMARIO DE LA INVENCIÓN

Es el objetivo de la invención dar a conocer un concepto para un codificador de audio multicanal que proporcione una estimación de parámetros estable y rápida a la vez.

Este objetivo se consigue por las características de las reivindicaciones independientes. Otras formas de puesta en práctica son evidentes a partir de las reivindicaciones subordinadas, de la descripción y de las Figuras adjuntas.

Esta invención está basada en el hallazgo de que la aplicación, a la vez, de un fuerte suavizado y de un suavizado débil, también referido como suavizado bajo, a la correlación cruzada en caso de ITD o a la energía en el caso de CLD da lugar a dos parámetros de codificación diferentes en donde uno de ellos sigue rápidamente los cambios de ITD o de CLD mientras que el otro proporciona un valor de parámetro estable durante tramas consecutivas. Utilizando un procedimiento de detección inteligente que depende de un criterio de calidad, tal como un criterio de estabilidad, el parámetro de codificación resultante es, a la vez, estable y con un seguimiento rápido de los cambios de los parámetros ITD o CLD.

Una evaluación única de la correlación no es suficiente para obtener la estabilidad, esto es, mantener una evaluación coherente del parámetro ITD en el transcurso del tiempo cuando al fuente real no se desplaza y la reactividad, esto es, para cambiar la función de evaluación de forma muy rápida cuando la fuente real se está desplazando o cuando una nueva fuente con una posición diferente aparece en la escena operativa de audio. El hecho de que se tengan dos funciones de evaluación diferentes del mismo parámetro con diferente efecto de memoria sobre la base de factores de suavizado diferentes permite centrar una evaluación sobre la estabilidad y la otra sobre la reactividad. Un algoritmo de selección se da a conocer para seleccionar la mejor evaluación, esto es, la evaluación más fiable. Aspectos de la idea inventiva están basados en dos versiones de la misma función de evaluación con factores de suavizado distintos. Un criterio de calidad o de fiabilidad se introduce para la decisión para la conmutación desde una evaluación a largo plazo a una evaluación a corto plazo. Con el fin de beneficiarse de la evaluación a corto plazo y de la evaluación a largo plazo, el estatus a largo plazo se actualiza por el estatus a corto plazo con el fin de cancelar el efecto de memoria.

Para poder describir la invención en detalle, se utilizarán los términos, abreviaturas y notaciones siguientes:

ES 2 571 742 T3

5	BCC:	significa codificación de pistas binaurales, que es una codificación de señales multicanal o estéreo que utiliza una mezcla descendente y pistas binaurales (o parámetros espaciales) para describir las relaciones entre canales.
	Pistas binaurales:	pistas entre canales entre las señales de entrada de los oídos izquierdo y derecho (véase también ITD, ILD e IC).
10	CLD:	Diferencia de niveles de canales, lo mismo que ILD.
	FFT:	Implantación rápida de la DFT, denominada Transformada de Fourier Rápida.
15	HRTF:	Función de transferencia relacionada con auriculares, modelado de la transducción de sonido desde una fuente a las entradas de los oídos izquierdo y derecho en un campo libre.
	IC:	Coherencia inter-aural, esto es, grado de similitud entre las señales de entrada en los oídos izquierdo y derecho. A veces, también se refiere como una IAC o correlación cruzada inter-aural (IACC).
20	ICC:	Coherencia intercanal, correlación intercanal. Lo mismo que IC, pero con una definición más general entre cualquier par de señales (p.ej., par de señales de altavoces, par de señales de entrada a los oídos, etc.).
	ICPD:	Diferencia de fase intercanal. Diferencia de fase media entre un par de señales.
25	ICLD:	Diferencia de niveles intercanal. Lo mismo que ILD, pero con una definición más general entre cualquier par de señales (p.ej., par de señales de altavoces, par de señales de entrada a los oídos, etc.).
30	ICTD:	Diferencia de tiempo entre canales. Lo mismo que ITD, pero con una definición más general entre cualquier par de señales (p.ej., par de señales de altavoces, par de señales de entrada a los oídos, etc.).
	ILD:	Diferencia de nivel inter-aural, esto es, diferencia de niveles entre las señales de entrada a los oídos izquierdo y derecho. A veces, se refiere también como una diferencia de intensidad inter-aural (IID).
35	IPD:	Diferencia de fase inter-aural, esto es, diferencia de fase entre las señales de entrada a los oídos izquierdo y derecho.
40	ITD:	Diferencia de tiempo inter-aural, esto es, diferencia de tiempo entre las señales de entrada a los oídos izquierdo y derecho. A veces, se refiere también como un retardo inter-aural.
	ICD:	Diferencia entre canales. El término general para una diferencia entre dos canales, p.ej., una diferencia temporal, una diferencia de fase, una diferencia de nivel o una coherencia entre los dos canales.
45		
	Mezcla:	Dadas varias señales de fuente (p.ej., instrumentos registrados por separado, registro multipista), el proceso de generación de señales de audio multicanal o estéreo previstas para la reproducción de audio espacial se denomina mezcla.
50	OCPD:	Diferencia de fase de canal global. Una modificación de fase común de dos o más canales de audio.
	Audio espacial:	Señales de audio que, cuando se reproducen a través de un sistema de reproducción adecuado, evocan una imagen espacial auditiva.
55		
	Pistas espaciales:	Pistas pertinentes para la percepción espacial. Este término se utiliza para las pistas entre pares de canales de una señal de audio multicanal o estéreo (véase también ICTD, ICLD e ICC). También se indica como parámetros espaciales o pistas binaurales.
60		

En conformidad con un primer aspecto de la idea inventiva, la invención se refiere a un método para la determinación de un parámetro de codificación para una señal de canal de audio de una pluralidad de señales de canales de audio de una señal de audio multicanal, teniendo cada señal de canal de audio valores de señales de canales de audio, cuyo método comprende: la determinación para la señal de canal de audio de un conjunto de funciones a partir de los valores de señal de canal de audio de la señal de canal de audio y de los valores de señal

de audio de referencia de una señal de audio de referencia, en donde la señal de audio de referencia es otra señal de canal de audio de la pluralidad de señales de canales de audio; la determinación de un primer conjunto de parámetros de codificación sobre la base de un suavizado operativo del conjunto de funciones con respecto a una secuencia de tramas de la señal de audio multicanal, estando el suavizado basado en un primer coeficiente de suavizado, con la determinación de un segundo conjunto de parámetros de codificación sobre la base de un suavizado del conjunto de funciones con respecto a la secuencia de tramas de la señal de audio multicanal, estando el suavizado basado en un segundo coeficiente de suavizado; la determinación del parámetro de codificación sobre la base de un criterio de calidad con respecto al primer conjunto de parámetros de codificaciones y/o el segundo conjunto de parámetros de codificación.

En conformidad con un segundo aspecto de la idea inventiva, la invención se refiere a un método para la determinación de un parámetro de codificación para una señal de canal de audio de una pluralidad de señales de canal de audio de una señal de audio multicanal, teniendo cada señal de canal de audio valores de señales de canales de audio, cuyo método comprende: la determinación para la señal de canal de audio de un conjunto de funciones a partir de los valores de señales de canal de audio de la señal de canal de audio y valores de la señal de audio de referencia de una señal de audio de referencia, en donde la señal de audio de referencia es una señal de audio de mezcla descendente derivada de al menos dos señales de canal de audio de la pluralidad de señales de audio multicanal; la determinación de un primer conjunto de parámetros de codificación sobre la base de un suavizado operativo del conjunto de funciones con respecto a una secuencia de tramas de la señal de audio multicanal, estando el suavizado basado en un primer coeficiente de suavizado; la determinación de un segundo conjunto de parámetros de codificación sobre la base de un suavizado del conjunto de funciones con respecto a la secuencia de tramas de la señal de audio multicanal, estando el suavizado basado en un segundo coeficiente de suavizado; la determinación del parámetro de codificación sobre la base de un criterio de calidad con respecto al primer conjunto de parámetros de codificación y/o el segundo conjunto de parámetros de codificación.

La versión fuertemente suavizada del conjunto de funciones, p.ej., el suavizado basado en el primer parámetro de suavizado hace estable la estimación. La versión débilmente suavizada del conjunto de funciones p.ej., el suavizado basado en el segundo parámetro de suavizado que se determina al mismo tiempo, hace que la estimación siga los cambios rápidos reales del parámetro de estimación, esto es, el parámetro ITD o el parámetro CLD. La memoria de la versión fuertemente suavizada del conjunto de funciones se actualiza por la versión débilmente suavizada del conjunto de funciones con lo que se proporciona el resultado óptimo con respecto al seguimiento de la velocidad y estabilidad. La decisión de que se utilice la versión suavizada está basada en una métrica de calidad del primer conjunto y/o el segundo conjunto de parámetros de codificación. Por lo tanto, se proporciona, a la vez, una estimación de parámetros estable y rápida.

En una primera posible forma de puesta en práctica del método según el primer aspecto o según el segundo aspecto de la idea inventiva, la determinación del conjunto de funciones comprende: la determinación de una transformada de frecuencia de los valores de señal de canal de audio de la señal de canal de audio; la determinación de una transformada de frecuencia de los valores de la señal de audio de referencia de la señal de audio de referencia; la determinación del conjunto de funciones como un espectro cruzado o una correlación cruzada para al menos cada sub-banda de frecuencias de un subconjunto de sub-bandas de frecuencias, siendo cada función del conjunto de funciones calculada entre una parte de señal de banda limitada de la señal de canal de audio y una parte de señal de banda limitada de la señal de audio de referencia en la respectiva sub-banda de frecuencias a la que está asociada la función del conjunto de funciones.

Cuando se estima el parámetro de codificación en el dominio de la frecuencia sobre la base de una correlación cruzada, se aumenta la estabilidad de la estimación del parámetro de codificación. El conjunto de funciones puede procesarse para sub-bandas de frecuencias, con lo que se mejora la flexibilidad en la elección del parámetro de codificación y se mejora la robustez operativa con respecto al ruido puesto que una sub-banda de frecuencias es menos sensible al ruido que la banda de frecuencia completa.

En una segunda posible forma de realización del método según la primera forma de realización del primer aspecto de la idea inventiva o en conformidad con la primera forma de realización del segundo aspecto, una sub-banda de frecuencias comprende una o una pluralidad de contenedores de frecuencia.

La magnitud de la sub-banda de frecuencia puede ajustarse flexiblemente permitiendo, de este modo, utilizar diferentes parámetros de codificación para la sub-banda de frecuencias.

En una tercera posible forma de realización del método según el primer aspecto de la idea inventiva como tal o en conformidad con el segundo aspecto o en conformidad con cualquiera de las formas de realización precedentes del primer aspecto o en conformidad con cualquiera de las formas de realización precedentes del segundo aspecto, los primero y segundo conjuntos de parámetros de codificación comprenden diferencias entre canales, en donde las diferencias entre canales comprenden diferencias de tiempo entre canales y/o diferencias de niveles entre canales.

Las diferencias entre canales pueden utilizarse como parámetros espaciales para detectar una diferencia entre un primero y un segundo canal de audio de una señal de audio multicanal. La diferencia puede ser, a modo de ejemplo,

una diferencia en el tiempo de llegada tal como una diferencia de tiempo inter-aural o una diferencia de tiempo entre canales o una diferencia en el nivel de ambos canales de audio. Ambas diferencias son adecuadas para utilizarse como parámetro de codificación.

5 En una cuarta posible forma de realización del método según el primer aspecto de la idea inventiva como tal o en conformidad con el segundo aspecto como tal o en conformidad con cualquiera de las formas de realización precedentes del primer aspecto o en conformidad con cualquiera de las formas de realización precedentes del segundo aspecto, la determinación del parámetro de codificación sobre la base de un criterio de calidad comprende la determinación de un parámetro de estabilidad, utilizándose el parámetro de estabilidad mediante el criterio de
10 calidad.

El criterio de calidad puede basarse, a modo de ejemplo, en un parámetro de estabilidad con lo que se aumenta la estabilidad de la estimación del parámetro de codificación. De forma adicional o alternativa, el criterio de calidad puede basarse en un criterio de calidad de experiencia (QoE) para aumentar la calidad QoE para el usuario. El
15 criterio de calidad puede basarse en un criterio de ancho de banda para una utilización eficiente del ancho de banda cuando se realiza la codificación de audio.

En una posible quinta forma de realización del método en conformidad con la cuarta forma de realización del primer aspecto de la idea inventiva o en conformidad con la cuarta forma de realización del segundo aspecto, la determinación del parámetro de codificación comprende: la determinación de un parámetro de estabilidad del segundo conjunto de parámetros de codificación sobre la base de una comparación entre valores consecutivos del segundo conjunto de parámetros de codificación con respecto a la secuencia de tramas, y la determinación del parámetro de codificación dependiendo del parámetro de estabilidad.

25 Utilizando un parámetro de estabilidad, se mejora la estabilidad de la estimación. Además, se aumenta la velocidad de estimación puesto que el suavizado de la correlación cruzada o de la energía puede reducirse hasta que el parámetro de estabilidad indique una pérdida de estabilidad.

En una sexta posible forma de realización del método en conformidad con la cuarta forma de realización del primer aspecto de la idea inventiva o en conformidad con la cuarta forma de realización del segundo aspecto, el parámetro de estabilidad está basado al menos en una desviación estándar del segundo conjunto de parámetros de codificación.
30

La desviación estándar es fácil de calcular y proporciona una medida exacta de la estabilidad. Cuando la desviación estándar es de pequeña magnitud, la estimación es estable o fiable, mientras que cuando la desviación estándar es grande, la estimación es inestable y no fiable.
35

En una séptima posible forma de realización del método según la cuarta forma de realización del primer aspecto de la idea inventiva o en conformidad con la cuarta forma de realización del segundo aspecto o en conformidad con la quinta forma de realización del primer aspecto o en conformidad con la quinta forma de realización del segundo aspecto, el parámetro de estabilidad se determina durante una sola trama o durante múltiples tramas de la señal de audio multicanal.
40

La determinación del parámetro de estabilidad durante una sola trama de la señal de audio multicanal es fácil de poner en práctica y tiene una baja complejidad de cálculo mientras que la determinación del parámetro de estabilidad durante múltiples tramas proporciona una estimación exacta para la estabilidad.
45

En una octava posible forma de realización del método según cualquiera de la cuarta a la séptima formas de realización del primer aspecto de la idea inventiva o en conformidad con cualesquiera de la cuarta a la séptima formas de realización del segundo aspecto de la idea inventiva, la determinación del parámetro de codificación se determina sobre la base de un cruce de umbral del parámetro de estabilidad.
50

Cuando el parámetro de estabilidad es inferior al valor umbral, la estimación es estable o fiable, mientras que cuando un parámetro de estabilidad es superior al valor umbral indica una estimación inestable o no fiable.
55

En una novena posible forma de realización del método en conformidad con la octava forma de realización del primer aspecto de la idea inventiva o en conformidad con la octava forma de realización del segundo aspecto, el método comprende, además: la actualización del primer conjunto de parámetros de codificación con el segundo conjunto de parámetros de codificación si el parámetro de estabilidad cruza el valor umbral.
60

Mediante la actualización, la estimación del primer conjunto de parámetros de codificación se puede mejorar. Cuando el parámetro de estabilidad es superior al valor umbral lo que indica una estimación estable, se puede actualizar el suavizado a largo plazo o sustituirse por un suavizado a corto plazo con lo que se aumenta la velocidad de estimación al mismo tiempo que se mantiene la estabilidad.
65

En una décima posible forma de realización del método según el primer aspecto de la idea inventiva o en

5 conformidad con el segundo aspecto como tal o en conformidad con cualquiera de las formas de realización precedentes del primer aspecto o en conformidad con cualquiera de las formas de realización precedentes del segundo aspecto, el suavizado del conjunto de funciones sobre la base de un primero y un segundo coeficiente de suavizado se calcula como una adición de un estado de memoria de la primera y segunda versiones suavizadas del conjunto de funciones multiplicado por un primer coeficiente sobre la base del primero y del segundo coeficiente de suavizado y el conjunto de funciones multiplicado por un segundo coeficiente sobre la base del primero y del segundo coeficiente de suavizado.

10 Dicho cálculo recursivo utiliza una memoria para memorizar valores anteriores de la primera y de la segunda versión suavizada del conjunto de funciones. La función de suavizado recursivo es eficiente en el cálculo puesto que el número de adiciones y multiplicaciones es bajo. El suavizado recursivo es eficiente en memoria puesto que solamente se requiere un estado de memoria para memorizar el conjunto de funciones de suavizado anterior, siendo actualizado el estado de memoria en cada etapa de cálculo.

15 En una undécima posible forma de realización del método según la décima forma de realización del primer aspecto de la idea inventiva o en conformidad con la décima forma de realización del segundo aspecto, el método comprende, además: la actualización del estado de memoria de la primera versión suavizada del conjunto de funciones con el estado de memoria de la segunda versión suavizada del conjunto de funciones si el parámetro de estabilidad cruza el valor umbral.

20 Mediante la actualización del estado de memoria de la primera versión suavizada del conjunto de funciones con el estado de memoria de la segunda versión suavizada del conjunto de funciones dependiendo del parámetro de estabilidad, se mejora la estabilidad y la velocidad de estimación. Cuando el parámetro de estabilidad es superior al valor umbral lo que indica una estimación estable, el suavizado a largo plazo, esto es, la primera versión suavizada del conjunto de funciones, puede actualizarse o sustituirse por un suavizado a corto plazo, esto es, la segunda versión suavizada del conjunto de funciones, con lo que se aumenta la velocidad de estimación al mismo tiempo que se mantiene la estabilidad.

30 En una duodécima posible forma de realización del método en conformidad con el primer aspecto de la idea inventiva o en conformidad con el segundo aspecto como tal o en conformidad con cualquiera de las formas de realización precedentes del primer aspecto o en conformidad con cualquiera de las formas de realización precedentes del segundo aspecto, el primer coeficiente de suavizado es superior al segundo coeficiente de suavizado.

35 El primer coeficiente de suavizado permite una estimación a largo plazo mientras que el segundo coeficiente de suavizado permite una estimación a corto plazo, con lo que es posible discriminar entre resultados de suavizado distintos.

40 En una décimo tercera posible forma de realización del método según el primer aspecto de la idea inventiva o en conformidad con el segundo aspecto como tal, o en conformidad con cualquiera de las formas de realización precedentes del primer aspecto o en conformidad con cualquiera de las formas de realización precedentes del segundo aspecto, el suavizado del conjunto de funciones es con respecto a al menos dos tramas consecutivas de la señal de audio multicanal.

45 El suavizado es más exacto si se utilizan dos o más tramas consecutivas de la señal de audio multicanal.

50 En una decimocuarta forma de realización posible del método según el primer aspecto de la idea inventiva o en conformidad con el segundo aspecto como tal o en conformidad con cualquiera de las formas de realización precedentes del primer aspecto o en conformidad con cualquiera de las formas de realización precedentes del segundo aspecto, el suavizado del conjunto de funciones discrimina entre valores positivos del segundo conjunto de parámetros de codificación y valores negativos del segundo conjunto de parámetros de codificación.

55 Mediante la discriminación entre valores positivos y negativos del segundo conjunto de parámetros de codificación, la estimación tiene un más alto grado de precisión.

60 En una decimoquinta posible forma de realización del método según la decimocuarta forma de realización del primer aspecto o en conformidad con la decimocuarta forma de realización del segundo aspecto, el suavizado del conjunto de funciones comprende: el conteo de un primer número de valores positivos del segundo conjunto de parámetros de codificación y un segundo número de valores negativos del segundo conjunto de parámetros de codificación durante varios contenedores de frecuencia o sub-bandas de frecuencia.

El conteo de los valores positivos y negativos permite discriminar el segundo conjunto de parámetros de codificación dependiendo de su signo. La velocidad de estimación se aumenta mediante esa discriminación.

65 En conformidad un tercer aspecto de la idea inventiva, la invención se refiere a un codificador de audio multicanal para determinar un parámetro de codificación para una señal de canal de audio de una pluralidad de señales de

canal de audio de una señal de audio multicanal, teniendo cada señal de canal de audio valores de señales de canal de audio, y comprendiendo el codificador de audio multicanal: un primer dispositivo de determinación para determinar para la señal de canal de audio un conjunto de funciones a partir de los valores de la señal de canal de audio de la señal de canal de audio y de los valores de señal de audio de referencia de una señal de audio de referencia, en donde la señal de audio de referencia es otra señal de canal de audio de la pluralidad de señales de canal de audio; un segundo dispositivo de determinación para determinar un primer conjunto de parámetros de codificación sobre la base de un suavizado del conjunto de funciones con respecto a una secuencia de tramas de la señal de audio multicanal, estando el suavizado basado en un primer coeficiente de suavizado; un tercer dispositivo de determinación para determinar un segundo conjunto de parámetros de codificación sobre la base de un suavizado del conjunto de funciones con respecto a la secuencia de tramas de la señal de audio multicanal, estando el suavizado basado en un segundo coeficiente de suavizado; un dispositivo de determinación de parámetro de codificaciones para determinar el parámetro de codificación sobre la base de un criterio de calidad con respecto al primer conjunto de parámetros de codificación y/o segundo conjunto de parámetros de codificación.

En conformidad con un cuarto aspecto de la idea inventiva, la invención se refiere a un codificador de audio multicanal para determinar un parámetro de codificación para una señal de canal de audio de entre una pluralidad de señales de canal de audio de una señal de audio multicanal, teniendo cada señal de canal de audio valores de señales de canal de audio, comprendiendo el codificador de audio multicanal: un primer dispositivo de determinación para determinar para la señal de canal de audio un conjunto de funciones a partir de los valores de señales de canal de audio de la señal de canal de audio y de los valores de señal de audio de referencia de una señal de audio de referencia, en donde la señal de audio de referencia es una señal de audio de mezcla descendente derivada de al menos dos señales de canal de audio de entre la pluralidad de las señales de audio multicanal; un segundo dispositivo de determinación para determinar un primer conjunto de parámetros de codificación sobre la base de un suavizado del conjunto de funciones con respecto a una secuencia de tramas de la señal de audio multicanal, estando el suavizado basado en un primer coeficiente de suavizado; un tercer dispositivo de determinación para determinar un segundo conjunto de parámetros de codificación sobre la base de un suavizado del conjunto de funciones con respecto a la secuencia de tramas de la señal de audio multicanal, estando el suavizado basado en un segundo coeficiente de suavizado; un dispositivo de determinación de parámetros de codificación para determinar el parámetro de codificación sobre la base de un criterio de calidad con respecto al primer conjunto de parámetros de codificación y/o el segundo conjunto de parámetros de codificación.

Dicho codificador de audio multicanal proporciona una codificación óptima con respecto a velocidad y estabilidad. La versión fuertemente suavizada del conjunto de funciones, p.ej., el suavizado basado en el primer parámetro de suavizado hace estable la estimación. La versión débilmente suavizada del conjunto de funciones p.ej., el suavizado basado en el segundo parámetro de suavizado que se determina al mismo tiempo, hace que la estimación siga los cambios rápidos reales del parámetro de estimación, esto es, el parámetro ITD o el parámetro CLD. La memoria de la versión fuertemente suavizada del conjunto de funciones se actualiza por la versión débilmente suavizada del conjunto de funciones con lo que se proporciona el resultado óptimo con respecto a la velocidad de seguimiento y la estabilidad. La decisión de qué versión suavizada utilizar está basado en una métrica de calidad del primer conjunto y/o el segundo conjunto de parámetros de codificación. De este modo, se proporciona, a la vez, una estimación de parámetros estable y rápida.

En conformidad con un quinto aspecto, la invención se refiere a un programa informático con un código de programa para realizar el método en conformidad con el primer aspecto de la idea inventiva o en conformidad con el segundo aspecto como tal o en conformidad con cualquiera de las formas de realización precedentes del primer aspecto o en conformidad con cualquiera de las formas de realización precedentes del segundo aspecto cuando se ejecuta en un ordenador.

En conformidad con un sexto aspecto, la invención se refiere a un soporte legible por máquina tal como una memoria de almacenamiento, en particular, un disco compacto, con un programa informático que comprende un código de programa para realizar el método según el primer aspecto como tal o en conformidad con el segundo aspecto como tal o en conformidad con cualquiera de las reivindicaciones precedentes del primer aspecto o en conformidad con cualquiera de las reivindicaciones precedentes del segundo aspecto cuando se ejecuta en un ordenador.

Aspectos de la invención anteriormente descritos pueden utilizarse para la estimación del parámetro ITD en un codificador de audio espacial paramétrico. En un codificador de audio espacial paramétrico o codificador de audio multicanal, los parámetros espaciales se extraen y cuantifican antes de multiplexarse en el flujo de bits. El parámetro (a modo de ejemplo ITD) puede estimarse en el dominio de la frecuencia sobre la base de una correlación cruzada. Con el fin de hacer la estimación más estable, la correlación cruzada del dominio de la frecuencia está fuertemente suavizada para la estimación del parámetro (ITD). Con el fin de seguir los cambios rápidos del parámetro, una versión débilmente suavizada de la correlación cruzada del dominio de la frecuencia se calcula también al mismo tiempo sobre la base de una estimación casi instantánea de la correlación cruzada reduciendo el efecto de memoria.

La versión débilmente suavizada de la función de estimación se utiliza para estimar el parámetro (ITD) y para actualizar la memoria de correlación cruzada de la versión fuertemente suavizada de la correlación cruzada en caso

de cambios en el estatus del parámetro. La decisión de utilizar la versión débilmente suavizada está basada en una métrica de calidad de los parámetros estimados. El parámetro se estima sobre la base de las dos versiones de la función de estimación. La mejor estimación se mantiene y si se selecciona la función débilmente suavizada, se utiliza también para actualizar la versión fuertemente suavizada.

5 A modo de ejemplo, en el caso de estimación del parámetro ITD, se calcula ITD_inst (una versión débilmente suavizada de ITD) sobre la base de la versión débilmente suavizada de la correlación cruzada del dominio de la frecuencia. Si la desviación estándar de ITD_inst durante varios contenedores/sub-bandas de frecuencia es inferior a un valor umbral predeterminado, la memoria de la correlación cruzada fuertemente suavizada será actualizada por una de entre la versión débilmente suavizada y el parámetro ITD estimado con la función débilmente suavizada es objeto de selección.

15 Una métrica de calidad simple está basada en la desviación estándar de la estimación del parámetro ITD de la versión débilmente suavizada. Por supuesto, se puede utilizar similarmente otras métricas de calidad. A modo de ejemplo, una posibilidad de cambio de posición puede calcularse sobre la base de toda la información espacial disponible (CLD, ITD, ICC). A modo de ejemplo, la correlación entre un cambio rápido de ITD y un cambio rápido de CLD representará una alta probabilidad de modificación de la imagen espacial.

20 Los métodos aquí descritos pueden ponerse en práctica como software en un Procesador de Señal Digital (DSP), en un microcontrolador o en cualquier otro procesador lateral o como circuitos de hardware dentro de un circuito integrado específico de la aplicación (ASIC).

25 La invención puede ponerse en práctica en circuitos electrónicos digitales, o en hardware informático, firmware, software o en una de sus combinaciones.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

30 Formas de realización adicionales de la invención se describirán haciendo referencia a las Figuras siguientes, en donde:

La Figura 1a ilustra un diagrama esquemático de un método para determinar un parámetro de codificación para una señal de canal de audio en conformidad con una forma de realización;

35 La Figura 1b ilustra un diagrama esquemático de un método para determinar un parámetro de codificación para una señal de canal de audio en conformidad con una forma de realización;

La Figura 2 ilustra un diagrama esquemático de un algoritmo de estimación de ITD en conformidad con una forma de realización;

40 La Figura 3 ilustra un diagrama esquemático de un algoritmo de estimación de CLD en conformidad con una forma de realización;

45 La Figura 4 ilustra un diagrama de bloques de un codificador de audio paramétrico en conformidad con una forma de realización;

La Figura 5 ilustra un diagrama de bloques de un decodificador de audio paramétrico en conformidad con una forma de realización;

50 La Figura 6 ilustra un diagrama de bloques de un codificador y decodificador de audio estéreo paramétrico en conformidad con una forma de realización;

La Figura 7 ilustra un diagrama de bloques de un algoritmo de selección de ITD en conformidad con una forma de realización; y

55 La Figura 8 ilustra un diagrama esquemático que representa los principios de las diferencias de tiempo inter-aurales.

DESCRIPCIÓN DETALLADA DE LAS FORMAS DE REALIZACIÓN DE LA INVENCION

60 La Figura 1a ilustra un diagrama esquemático de un método 100a para determinar un parámetro de codificación para una señal de canal de audio en conformidad con una forma de realización.

65 El método 100a es para determinar un parámetro de codificación ITD, p.ej., una diferencia de tiempo entre canales o una diferencia de tiempo inter-aural, para una señal de canal de audio x_1 de una pluralidad de señales de canal de audio x_1 , x_2 de una señal de canal de audio multicanal. Cada señal de canal de audio x_1 , x_2 comprende valores de señales de canal de audio $x_1[n]$, $x_2[n]$. El método 100a comprende:

- 5 la determinación 101 para la señal de canal de audio x_1 de un conjunto de funciones $c[b]$ a partir de los valores de la señal de canal de audio $x_1[n]$ de la señal de canal de audio x_1 y valores de señales de audio de referencia $x_2[n]$ de una señal de audio de referencia x_2 , en donde la señal de audio de referencia es otra señal de canal de audio x_2 de entre la pluralidad de señales de canal de audio o una señal de audio de mezcla descendente derivada de al menos dos señales de canal de audio x_1 , x_2 de la pluralidad de señales de audio multicanal;
- 10 la determinación 103a de un primer conjunto de parámetros de codificación ITD[b] sobre la base de un suavizado del conjunto de funciones $c[b]$ con respecto a una secuencia de tramas i de la señal de audio multicanal, basándose el suavizado en un primer coeficiente de suavizado SMW_1 .
- 15 la determinación 105a de un segundo conjunto de parámetros de codificación ITD_inst[b] sobre la base de un suavizado del conjunto de funciones $c[b]$ con respecto a la secuencia de tramas i de la señal de audio multicanal, basándose el suavizado en un segundo coeficiente de suavizado SMW_2 ; y
- 20 la determinación 107a del parámetro de codificación ITD sobre la base de un criterio de calidad con respecto al primer conjunto de parámetros de codificación ITD[b] y/o el segundo conjunto de parámetros de codificación ITD_inst[b].
- 25 En una forma de realización, la determinación 107a del parámetro de codificación ITD comprende la comprobación de la estabilidad del segundo conjunto de parámetros de codificación ITD_inst[b]. Si el segundo conjunto de parámetros de codificación ITD_inst[b] es estable durante todos los contenedores de frecuencia b , la selección del parámetro de codificación ITD sobre la base del segundo conjunto de parámetros de codificación ITD_inst[b] como la estimación final y la actualización de una memoria del suavizado del conjunto de funciones $c[b]$ sobre la base del primer coeficiente de suavizado SMW_1 mediante el suavizado del conjunto de funciones $c[b]$ sobre la base del segundo coeficiente de suavizado SMW_2 . Si el segundo conjunto de parámetros de codificación ITD_inst[b] no es estable durante todos los contenedores de frecuencia b , seleccionar el parámetro de codificación ITD sobre la base del primer conjunto de parámetros de codificación ITD[b] como la estimación final.
- 30 En una forma de realización, el método 100a comprende las etapas siguientes:
- 35 Calcular 101 una primera función $c[b]$ y calcular 103a la función suavizada asociada $C_{sm}[b]$ para la estimación del parámetro ITD a partir de la señal de entrada $x_1[n]$, $x_2[n]$ sobre la base de un primer coeficiente de suavizado.
- Calcular 105a una segunda función suavizada $C_{sm_inst}[b]$ para la estimación del parámetro ITD a partir de la señal de entrada $X_1[n]$, $x_2[n]$ sobre la base de un segundo coeficiente de suavizado.
- Calcular 107a la primera y la segunda estimación del parámetro ITD e ITD_inst sobre la base de las dos versiones suavizadas $C_{sm}[b]$ y $C_{sm_inst}[b]$ de la función de estimación.
- 40 Comprobar 107a la estabilidad de la segunda estimación del parámetro ITD_inst. Si la segunda estimación del parámetro es estable, seleccionar la segunda estimación del parámetro ITD_inst como la estimación final y actualizar la memoria de la primera función suavizada por la segunda función suavizada. Si la segunda estimación del parámetro no es estable, seleccionar la primera estimación del parámetro ITD como la estimación final.
- 45 En una forma de realización, el método 100a comprende las etapas siguientes:
1. Calcular la FFT de la primera $x_1[n]$ y segunda $x_2[n]$ señales de canal.
 2. Calcular la correlación cruzada $c[n]$ de esos dos canales en el dominio de la frecuencia.
 - 2.1. Efectuar un fuerte suavizado de la correlación cruzada $c[n]$ y calcular el parámetro ITD (estimación a largo plazo de la diferencia de tiempo entre canales) de cada contenedor de frecuencia (o banda de frecuencia) con respecto al primer coeficiente de suavizado, esto es, el coeficiente de suavizado a largo plazo.
 - 2.2. Efectuar un débil suavizado de la correlación cruzada $c[n]$ y calcular el parámetro ITD_inst (estimación a corto plazo de la diferencia de tiempo entre canales) de cada contenedor de frecuencias (o banda de frecuencias) con respecto al segundo coeficiente de suavizado, esto es, coeficiente de suavizado a corto plazo.
 3. Calcular el valor medio y la desviación estándar de ITD_inst.
 4. Si la desviación estándar de ITD_inst es menor que un valor umbral, actualizar la memoria de la correlación cruzada fuertemente suavizada con la de la versión débilmente suavizada,
- y proporcionar a la salida el valor medio de ITD_inst como el ITD final. Si la desviación estándar de ITD_inst es superior al valor umbral, proporcionar, a la salida, el valor medio de ITD como el ITD final.
- 65

La Figura 1b ilustra un diagrama esquemático de un método 100b para determinar un parámetro de codificación para una señal de canal de audio en conformidad con una forma de realización.

5 El método 100b es para determinar un parámetro de codificación CLD, p.ej., una diferencia de niveles entre canales, para una señal de canal de audio x_1 de entre una pluralidad de señales de canal de audio x_1 , x_2 de una señal de audio multicanal. Cada señal de canal de audio x_1 , x_2 comprende valores de señales de canal de audio $x_1[n]$, $x_2[n]$. El método 100b comprende:

10 la determinación 101 para la señal de canal de audio x_1 de un conjunto de funciones $c[b]$ a partir de los valores de la señal de canal de audio $x_1[n]$ de la señal de canal de audio x_1 y de los valores de la señal de audio de referencia $x_2[n]$ de una señal de audio de referencia x_2 , en donde la señal de audio de referencia es otra señal de canal de audio x_2 de entre la pluralidad de señales de canal de audio o una señal de audio de mezcla descendente derivada de al menos dos señales de canal de audio x_1 , x_2 de entre la pluralidad de señales de audio multicanal;

15 la determinación 103b de un primer conjunto de parámetros de codificación $CLD[b]$ sobre la base de un suavizado del conjunto de funciones $c[b]$ con respecto a una secuencia de trama i de la señal de audio multicanal, estando el suavizado basado en un primer coeficiente de suavizado SMW_1 ;

20 la determinación 105b de un segundo conjunto de parámetros de codificación $CLD_inst[b]$ sobre la base de un suavizado del conjunto de funciones $c[b]$ con respecto a la secuencia de tramas i de la señal de audio multicanal, estando el suavizado basado en un segundo coeficiente de suavizado SMW_2 ; y

25 la determinación 107b del parámetro de codificación CLD sobre la base de un criterio de calidad con respecto al primer conjunto de parámetros de codificación $CLD[b]$ y/o el segundo conjunto de parámetros de codificación $CLD_inst[b]$.

30 En una forma de realización, la determinación 107b del parámetro de codificación CLD comprende la comprobación de la estabilidad del segundo conjunto de parámetros de codificación $CLD_inst[b]$. Si el segundo conjunto de parámetros de codificación $CLD_inst[b]$ es estable durante todos los contenedores de frecuencia b , la selección del parámetro de codificación CLD basada en el segundo conjunto de parámetros de codificación $CLD_inst[b]$ como la estimación final y la actualización de una memoria del suavizado del conjunto de funciones $c[b]$ sobre la base del primer coeficiente de suavizado SMW_1 mediante el suavizado del conjunto de funciones $c[b]$ sobre la base del segundo coeficiente de suavizado SMW_2 . Si el segundo conjunto de parámetros de codificación $CLD_inst[b]$ no es estable durante todos los contenedores de frecuencia b , la selección del parámetro de codificación CLD basado en el primer conjunto de parámetros de codificación $CLD[b]$ como la estimación final.

35 En una forma de realización, el método 100b comprende las etapas siguientes:

40 Calcular 101 una primera función $c[b]$ y calcular 103b la función suavizada asociada $C_{sm}[b]$ para la estimación del parámetro CLD a partir de la señal de entrada $x_1[n]$, $x_2[n]$ sobre la base de un primer coeficiente de suavizado.

Calcular 105b una segunda función suavizada $C_{sm_inst}[b]$ para la estimación del parámetro CLD a partir de los valores de la señal de entrada $x_1[n]$, $x_2[n]$ sobre la base de un segundo coeficiente de suavizado.

45 Calcular 107b la primera y la segunda estimación del parámetro CLD y CLD_{inst} sobre la base de las dos versiones suavizadas $C_{sm}[b]$ y $C_{sm_inst}[b]$ de la función de estimación.

50 Comprobar 107b la estabilidad de la segunda estimación del parámetro CLD_{inst} . Si la segunda estimación del parámetro es estable, seleccionar la segunda estimación del parámetro CLD_{inst} como la estimación final y actualizar la memoria de la primera función suavizada por la segunda función suavizada. Si la segunda estimación del parámetro no es estable, seleccionar la primera estimación del parámetro CLD como la estimación final.

55 En una forma de realización, el método 100b comprende las etapas siguientes:

1. Calcular la FFT de la primera $x_1[n]$ y segunda $x_2[n]$ señales de canal.
2. Calcular la energía $en[n]$ de dichos dos canales en el dominio de la frecuencia.
 - 2.1. Efectuar un fuerte suavizado de la energía $en[n]$ y calcular el parámetro CLD (estimación a largo plazo de la diferencia de niveles entre canales) de cada contenedor de frecuencia (o banda de frecuencia) con respecto al primer coeficiente de suavizado, esto es, el coeficiente de suavizado a largo plazo.
 - 2.2. Realizar un suavizado débil de la energía $en[n]$ y calcular el parámetro CLD_inst (estimación a corto plazo de la diferencia de niveles entre canales) de cada contenedor de frecuencia (o banda de frecuencia) con respecto al segundo coeficiente de suavizado, esto es, el coeficiente de suavizado a corto plazo.

3. Comprobar la estabilidad de la imagen estéreo basada en el parámetro CLD_inst.
 4. Si la imagen estéreo no es estable, actualizar la memoria de la energía fuertemente suavizada con una de entre la versión débilmente suavizada y proporcionar el parámetro CLD_inst como el CLD final.
- 5 Si la imagen estéreo es estable, proporcionar a la salida el parámetro CLD como el CLD final.

La Figura 2 ilustra un diagrama esquemático de un algoritmo de estimación de ITD 200 en conformidad con una forma de realización.

En una primera etapa 209, se aplica una transformada de tiempo-frecuencia sobre las muestras del primer canal de entrada $x_1[n]$ obteniendo así una representación de frecuencia $X_1[k]$ del primer canal de entrada x_1 . En una segunda etapa 211, se aplica una transformada de tiempo-frecuencia sobre las muestras del segundo canal de entrada $x_2[n]$ con la obtención de una representación de frecuencia $X_2[k]$ del segundo canal de entrada x_2 . En la forma de realización de los canales de entrada estéreo, el primer canal de entrada x_1 puede ser el canal izquierdo y el canal de entrada x_2 puede ser el canal derecho. En una forma de realización preferida, la transformada de tiempo-frecuencia es una Transformada de Fourier Rápida (FFT) o una Transformada de Fourier a corto plazo (STFT). En una forma de realización alternativa, la transformada de tiempo-frecuencia es un banco de filtros con modulación cosenoidal o un banco de filtros complejo.

En una tercera etapa 213, un espectro cruzado $c[b]$ se calcula a partir de la representación de frecuencias $X_1[k]$ y $X_2[k]$ del primer y del segundo canales de entrada x_1 , x_2 por sub-banda como

$$c[b] = \sum_{k=k_b}^{k_{b+1}-1} X_1[k]X_2^*[k]$$

en donde $c[b]$ es el espectro cruzado de la sub-banda b . $X_1[k]$ y $X_2[k]$ son los coeficientes de FFT de los dos canales (a modo de ejemplo, canales izquierdo y derecho en el caso de estéreo). * indica la conjugación compleja. k_b es el contenedor inicial de la sub-banda b y k_{b+1} es el contenedor de inicio de la sub-banda adyacente $b+1$. De este modo, los contenedores de frecuencia $[k]$ de la FFT desde k_b a $k_{b+1}-1$ representan la sub-banda $[b]$. Como alternativa, el espectro cruzado se calcula para cada contenedor de frecuencia la FFT como

$$c[b] = X_1[b]X_2^*[b]$$

en donde $c[b]$ es el espectro cruzado del contenedor de frecuencia $[b]$ y $X_1[b]$ y $X_2[b]$ son los coeficientes de FFT de los dos canales. * indica una conjugación compleja. Para este caso, una sub-banda $[b]$ corresponde directamente a un solo contenedor de frecuencia $[k]$, un contenedor de frecuencia $[b]$ y $[k]$ representa exactamente el mismo contenedor de frecuencia. El espectro cruzado $c[b]$ en esta forma de realización corresponde al conjunto de funciones $c[b]$ descrito con respecto a las Figuras 1a y 1b.

En una cuarta 215 y una quinta etapa 219, dos versiones de espectros cruzados suavizados $C_{sm}[b,i]$ y $C_{sm_inst}[b,i]$ se calculan a partir del espectro cruzado $c[b]$ como

$$c_{sm}[b, i] = SMW_1 * c_{sm}[b, i - 1] + (1 - SMW_1) * c[b]$$

$$c_{sm_inst}[b, i] = SMW_2 * c_{sm_inst}[b, i - 1] + (1 - SMW_2) * c[b]$$

en donde SMW_1 y SMW_2 son los factores de suavizado respectivos y $SMW_1 > SMW_2$. i es el índice de tramas de los espectros cruzados respectivos sobre la base de la señal de audio multicanal. En una forma de realización a modo de ejemplo, pero preferida, $SMW_1 = 0.9844$ y $SMW_2 = 0.75$.

En una sexta 221 y séptima etapa 223 las dos versiones de la diferencia de tiempos entre canales ITD e ITD_inst se calculan por contenedor o por sub-banda sobre la base del espectro cruzado fuertemente suavizado $C_{sm}[b,i]$ y el espectro cruzado débilmente suavizado $C_{sm_inst}[b,i]$ respectivamente como

$$ITD[b] = \frac{\angle c_{sm}[b,i]*N}{\pi*b}$$

$$ITD_inst[b] = \frac{\angle c_{sm_inst}[b,i]*N}{\pi*b}$$

en donde la operación \angle es el operador del argumento para calcular el ángulo del espectro cruzado suavizado. N es el número del contenedor de FFT.

5 En una octava etapa 225, el valor medio de la versión fuertemente suavizada de la diferencia de tiempo entre canales ITD se calcula durante todos los contenedores de interés (o sub-bandas).

$$ITD_{mean} = \frac{\sum_{b=B_1}^{B_2} ITD[b]}{B_2 - B_1}$$

10 en donde B_1 y B_2 son los índices del primero y del último contenedor (o sub-bandas) dentro de la zona de frecuencias de interés.

15 En una novena 227 y una décima etapa 229, el valor medio $ITD_{inst_{mean}}$ y la desviación estándar $ITD_{inst_{std}}$ de la versión débilmente suavizada de la de diferencia de tiempo entre canales ITD_{inst} se calculan durante todos los contenedores de frecuencia de interés (o sub-bandas de frecuencia).

$$ITD_{inst_{mean}} = \frac{\sum_{b=B_1}^{B_2} ITD_{inst}[b]}{B_2 - B_1}$$

$$ITD_{inst_{std}} = \sqrt{\frac{\sum_{b=B_1}^{B_2} (ITD_{inst}[b] - ITD_{inst_{mean}})^2}{B_2 - B_1}}$$

20 En una undécima etapa 231, se comprueba mediante comparación, si la desviación estándar de la versión débilmente suavizada de la diferencia de tiempo entre canales $ITD_{inst_{std}}$ es menor que un valor umbral (thr): $ITD_{inst_{std}} < thr$. Si esto es verdadero (Y= sí), la primera función suavizada $C_{sm}[b, i]$ está actualizada en una duodécima etapa 217 en conformidad con $C_{sm}[b, i] = C_{sm_{inst}}[b, i]$ y el valor medio $ITD_{inst_{mean}}$ de la versión débilmente suavizada de la diferencia de tiempo entre canales ITD_{inst} se proporciona, a la salida, como el parámetro de codificación final ITD en una decimotercera etapa 233. Si esto último no es verdadero (N=no), el valor medio ITD_{mean} de la versión fuertemente suavizada de la diferencia de tiempo entre canales ITD es proporcionada a la salida como el parámetro de codificación ITD final en una decimocuarta etapa 235.

30 Las etapas 209, 211 y 213 anteriormente descritas pueden representarse como una etapa 201 que corresponde a la etapa 101 según se describe con respecto a la Figura 1a. Las etapas 215 y 221 anteriormente descritas pueden representarse como la etapa 203 que corresponde a la etapa 103a según se describe con respecto a la Figura 1a. Las etapas 217, 219 y 223 descritas con anterioridad pueden representarse como una etapa 205 que corresponde a la etapa 105a según se describe con respecto a la Figura 1a. Las etapas 225, 227, 229, 231, 233 y 235 descritas con anterioridad pueden representarse como una etapa 207 que corresponde a la etapa 107a según se describe con respecto a la Figura 1a.

35 En una forma de realización preferida de la estimación de ITD, el parámetro de codificación ITD se calcula sobre la base de las dos versiones de suavizado para la diferencia de tiempo entre canales ITD e ITD_{inst} en donde cada una de las dos versiones de suavizado ITD e ITD_{inst} se determina sobre la base de un cálculo positivo y negativo de ITD e ITD_{inst} respectivamente, en conformidad con la puesta en práctica siguiente:

40 El conteo de los valores positivos y negativos de la versión fuertemente suavizada de la diferencia de tiempo entre canales ITD se realiza en esta etapa. El valor medio y la desviación estándar de los valores de ITD positivos y negativos están basados en el signo de ITD como sigue:

$$ITD_{mean_pos} = \frac{\sum_{i=0}^{i=M} ITD(i)}{Nb_{pos}} \text{ en donde } ITD(i) \geq 0$$

$$ITD_{mean_neg} = \frac{\sum_{i=0}^{i=M} ITD(i)}{Nb_{neg}} \text{ en donde } ITD(i) < 0$$

$$ITD_{std_pos} = \sqrt{\frac{\sum_{i=0}^{i=M} (ITD(i) - ITD_{mean_pos})^2}{Nb_{pos}}} \text{ en donde } ITD(i) \geq 0$$

$$ITD_{std_neg} = \sqrt{\frac{\sum_{i=0}^{i=M} (ITD(i) - ITD_{mean_neg})^2}{Nb_{neg}}} \text{ en donde } ITD(i) < 0$$

en donde $N_{b_{pos}}$ y $N_{b_{neg}}$ son el número de valores ITD positivos y negativos, respectivamente. M es el número total de parámetros ITDs que son objeto de extracción. Conviene señalar que, de forma alternativa, si ITD es igual a 0, puede ser objeto de conteo en ITD negativo o no contarse en ninguno de los valores medios.

5 ITD se selecciona de entre los valores ITD positivos y negativos basándose en la desviación estándar y en el valor medio en conformidad con el algoritmo de selección según se ilustra en la Figura 7.

10 El mismo cálculo se realiza para la versión débilmente suavizada de la diferencia de tiempo entre canales ITD_{inst} .

En una forma de realización en conformidad con una aplicación del método para el codificador-decodificador de audio paramétrico multicanal, el método 200 comprende las etapas siguientes:

15 En una primera y una segunda etapa 209 y 211, se aplica una transformada de tiempo-frecuencia en los canales de entrada. En una forma de realización preferida, la transformada de tiempo-frecuencia es una Transformada de Fourier Rápida (FFT) o una Transformada de Fourier a corto plazo (STFT). En una forma de realización alternativa, la transformada de tiempo-frecuencia puede ser un banco de filtros modulados en el modo cosenoidal o un banco de filtros complejos.

20 En una tercera etapa 213, se calcula un espectro cruzado del canal j por sub-bandas como

$$c_j[b] = \sum_{k=k_b}^{k_{b+1}-1} X_j[k]X_{ref}^*[k]$$

25 en donde $c_j[b]$ es el espectro cruzado del contenedor b o la sub-banda b . $X_j[b]$ y $X_{ref}[b]$ son los coeficientes de FFT del canal j y el canal de referencia. $*$ indica una conjugación compleja. k_b es el contenedor inicial de la banda b y k_{b+1} es el contenedor inicial de la sub-banda adyacente $b+1$. De este modo, los contenedores de frecuencias $[k]$ de la FFT desde k_b a $k_{b+1}-1$ representan la sub-banda $[b]$. En una forma de realización, el espectro de la señal de referencia X_{ref} se elige como uno del canal X_j (para j en $[1,M]$) y luego, se calculan $M-1$ pistas espaciales en el decodificador. En una forma de realización alternativa, X_{ref} es el espectro de una señal de mezcla descendente monoaural, que es la media de todos los canales M y luego, se calcula M pistas espaciales en el decodificador. La ventaja de utilizar una señal de mezcla descendente como una referencia para una señal de audio multicanal es evitar la utilización de una silenciosa como señal de referencia. En realidad, la señal de mezcla descendente representa una media de la energía de todos los canales y por lo tanto, está menos sujeta a la condición de silencio.

35 En una forma de realización alternativa, el espectro cruzado se calcula para cada contenedor de frecuencia de la FFT como:

$$c_j[b] = X_j[b]X_{ref}^*[b]$$

40 en donde $c_j[b]$ es el espectro cruzado del contenedor de frecuencia $[b]$. $X_{ref}[b]$ es el espectro de la señal de referencia y $X_j[b]$ (para j en $[1,M]$) son el espectro de cada canal de señal multicanal. $*$ indica una conjugación compleja. Para este caso, una sub-banda $[b]$ corresponde directamente a un contenedor de frecuencia $[k]$, los contenedores de frecuencia $[b]$ y $[k]$ representan exactamente el mismo contenedor de frecuencia.

45 En una cuarta etapa 215 y una quinta etapa 219, se calculan dos versiones de suavizado de espectro cruzado

$$c_{j,sm}[b,i] = SMW_1 * c_{j,sm}[b,i-1] + (1 - SMW_1) * c_j[b]$$

$$c_{j,sm_{inst}}[b,i] = SMW_2 * c_{j,sm_{inst}}[b,i-1] + (1 - SMW_2) * c_j[b]$$

50 en donde SMW_1 y SMW_2 son el factor de suavizado, y $SMW_1 > SMW_2$. i es el índice de tramas basado en la señal de audio multicanal. En una forma de realización preferida, $SMW_1 = 0.9844$ y $SMW_2 = 0.75$.

En una sexta etapa 221 y en una séptima etapa 223, los parámetros ITD e ITD_{inst} se calculan por contenedor o por sub-banda de frecuencias sobre la base del espectro cruzado fuertemente suavizado C_{sm} y el espectro cruzado débilmente suavizado $C_{sm_{inst}}$ respectivamente como:

55

$$ITD_j[b] = \frac{\angle c_{j,sm}[b,i] * N}{\pi * b}$$

$$ITD_inst_j[b] = \frac{\angle c_{j,sm_inst}[b,i] * N}{\pi * b}$$

en donde la operación \angle es el operador del argumento para calcular el ángulo del espectro cruzado suavizado. N es el número del contenedor de FFT.

5 En una octava etapa 225, el valor medio de ITD se calcula durante todos los contenedores de interés (o sub-bandas).

$$ITD_{mean,j} = \frac{\sum_{b=B_1}^{B_2} ITD_j[b]}{B_2 - B_1}$$

10 en donde B_1 y B_2 son los índices del primero y último contenedores (o sub-bandas) dentro de la zona de frecuencias de interés.

15 En una novena etapa 227 y una décima etapa 229, el valor medio y la derivación estándar de ITD_inst se calcula en todos los contenedores de interés (o sub-bandas) como sigue:

$$ITD_inst_{mean,j} = \frac{\sum_{b=B_1}^{B_2} ITD_inst_j[b]}{B_2 - B_1}$$

$$ITD_inst_{std,j} = \frac{\sqrt{\sum_{b=B_1}^{B_2} (ITD_inst_j[b] - ITD_inst_{mean,j})^2}}{B_2 - B_1}$$

20 En una undécima etapa 231, $ITD_inst_{std,j}$ se comprueba como siendo menor que un valor umbral thr en conformidad con $ITD_inst_{std,j} < \text{umbral}$. Si es más pequeño (ruta Y), la primera función suavizada se actualiza en una duodécima etapa 217 en conformidad con $C_{j,sm}[b, i] = C_{j,sm_inst}[b, i]$, y el valor medio de $ITD_inst_j(ITD_inst_{mean,j})$ se proporciona a la salida en una decimotercera etapa 233 como el parámetro ITD_j final. Si no es más pequeño (ruta N), el valor medio de ITD_j , ($ITD_{mean,j}$) se proporciona a la salida en una decimocuarta etapa 235 como el ITD_j final.

25 En una forma de realización preferida de la estimación de ITD, el parámetro de codificación ITD_j se calcula sobre la base de las dos versiones de suavizado para la diferencia de tiempo entre canales ITD_j e ITD_inst_j en donde cada una de las dos versiones de suavizado ITD_j e ITD_inst_j se determina sobre la base del cálculo positivo y negativo de ITD_j y de ITD_inst_j respectivamente, en conformidad con la forma de realización siguiente:

30 Se realiza el conteo de los valores positivos y negativos de la versión fuertemente suavizada de la diferencia de tiempo entre canales ITD. El valor medio y la desviación estándar de los valores de ITD positivos y negativos se basan en el signo de ITD como sigue:

$$ITD_{mean_pos} = \frac{\sum_{i=0}^{i=M} ITD(i)}{Nb_{pos}} \text{ en donde } ITD(i) \geq 0$$

$$ITD_{mean_neg} = \frac{\sum_{i=0}^{i=M} ITD(i)}{Nb_{neg}} \text{ en donde } ITD(i) < 0$$

$$ITD_{std_pos} = \frac{\sqrt{\sum_{i=0}^{i=M} (ITD(i) - ITD_{mean_pos})^2}}{Nb_{pos}} \text{ en donde } ITD(i) \geq 0$$

$$ITD_{std_neg} = \frac{\sqrt{\sum_{i=0}^{i=M} (ITD(i) - ITD_{mean_neg})^2}}{Nb_{neg}} \text{ en donde } ITD(i) < 0$$

35 en donde Nb_{pos} y Nb_{neg} son el número de valores de ITD positivos y negativos, respectivamente. M es el número total de ITDs que son objeto de extracción. Conviene señalar que, de forma alternativa, si ITD es igual 0, puede ser

objeto, o no, de conteo en ITD negativo o no contarse en ninguno de los valores medios.

ITD se selecciona de entre los valores de ITD positivos y negativos sobre la base del valor medio y de la desviación estándar en conformidad con el algoritmo de selección según se ilustra en la 7.

5 La Figura 3 ilustra un diagrama esquemático de un algoritmo de estimación de CLD en conformidad con una forma de realización.

10 En una primera etapa 309, se aplica una transformada de tiempo-frecuencia en las muestras del primer canal de entrada $x_1[n]$ con la obtención de una representación de frecuencia $X_1[k]$ del primer canal de entrada x_1 . En una segunda etapa 311, se aplica una transformada de tiempo-frecuencia en las muestras del segundo canal de entrada $x_2[n]$ con la obtención de una representación de frecuencia $X_2[k]$ del segundo canal de entrada x_2 . En la forma de realización de canales de entrada estéreo, el primer canal de entrada x_1 puede ser el canal izquierdo y el segundo canal de entrada x_2 puede ser el canal derecho. En una forma de realización preferida la transformada de tiempo-frecuencia es una Transformada de Fourier Rápida (FFT) o una Transformada de Fourier a corto plazo (STFT). En una forma de realización alternativa, la transformada de tiempo-frecuencia es un banco de filtros modulados en el modo cosenoidal o un banco de filtros complejos.

20 En una tercera etapa 313, la energía $en_1[b]$ del primer canal x_1 y la energía $en_2[b]$ del segundo canal x_2 se calculan por sub-banda b como

$$en_1[b] = \sum_{k=k_b}^{k_{b+1}-1} X_1[k]X_1^*[k]$$

$$en_2[b] = \sum_{k=k_b}^{k_{b+1}-1} X_2[k]X_2^*[k]$$

25 en donde $en_1[b]$ y $en_2[b]$ son las energías de la sub-banda b . $X_1[k]$ y $X_2[k]$ son los coeficientes de FFT de los dos canales (a modo de ejemplo, los canales izquierdo y derecho en caso de estéreo). * indica una conjugación compleja. k_b es el contenedor inicial de la banda b y k_{b+1} es el contenedor inicial de la sub-banda adyacente $b+1$. De este modo, los contenedores de frecuencias $[k]$ de la FFT desde k_b a $k_{b+1}-1$ representan la sub-banda $[b]$.

30 De forma alternativa, las energías de los dos canales x_1 y x_2 para cada contenedor de frecuencia de la FFT se calculan en conformidad con:

$$en_1[b] = X_1[b]X_1^*[b]$$

$$en_2[b] = X_2[b]X_2^*[b]$$

35 en donde $en_1[b]$ y $en_2[b]$ son las energías del contenedor de frecuencia $[b]$ del primer y del segundo canal respectivamente, $X_1[b]$ y $X_2[b]$ son los coeficientes de FFT de los dos canales. * indica una conjugación compleja. Para este caso, una sub-banda $[b]$ corresponde directamente a un solo contenedor de frecuencia $[k]$. Los contenedores de frecuencia $[b]$ y $[k]$ representan exactamente el mismo contenedor de frecuencia.

40 En una cuarta etapa 315, una versión fuertemente suavizada $en_{1_sm}[b,i]$ de la energía del primer canal x_1 y una versión fuertemente suavizada $en_{2_sm}[b,i]$ de la energía del segundo canal x_2 se determina y en una quinta etapa 319 una versión débilmente suavizada $en_{1_sm_inst}[b,i]$ de la energía del primer canal x_1 y una versión débilmente suavizada $en_{2_sm_inst}[b,i]$ de la energía del segundo canal x_2 se determinan como:

$$en_{1_sm}[b,i] = SMW_1 * en_{1_sm}[b,i-1] + (1 - SMW_1) * en_1[b]$$

$$en_{1_sm_inst}[b,i] = SMW_2 * en_{1_sm_inst}[b,i-1] + (1 - SMW_2) * en_1[b]$$

$$en_{2_sm}[b,i] = SMW_1 * en_{2_sm}[b,i-1] + (1 - SMW_1) * en_2[b]$$

$$en_{2_sm_inst}[b,i] = SMW_2 * en_{2_sm_inst}[b,i-1] + (1 - SMW_2) * en_2[b]$$

45 en donde SMW_1 y SMW_2 son los factores de suavizado o coeficientes de suavizado y $SMW_1 > SMW_2$, esto es, SMW_1 es el factor de fuerte suavizado y SMW_2 es el factor de suavizado débil. i es el índice de tramas. En una forma de realización siguiente, que sigue la evolución exacta de CLD, SMW_2 se establece a cero.

50 En una sexta etapa 321 y en una séptima etapa 323, la versión fuertemente suavizada de la diferencia de niveles entre canales CLD y la versión débilmente suavizada de la diferencia de niveles entre canales CLD_inst se calculan

por contenedor o por sub-banda sobre la base de las energías fuertemente suavizadas en_{1_sm} y en_{2_sm} y sobre las energías débilmente suavizadas $en_{1_sm_inst}$ y $en_{2_sm_inst}$ respectivamente como sigue:

$$CLD[b] = 10\log\left(\frac{en_{1_sm}[b]}{en_{2_sm}[b]}\right)$$

$$CLD_inst[b] = 10\log\left(\frac{en_{1_sm_inst}[b]}{en_{2_sm_inst}[b]}\right)$$

5 En una octava etapa 329, la estabilidad de la imagen estéreo se calcula sobre la base de la versión débilmente suavizada de la diferencia de niveles entre canales CLD_inst. En una forma de realización, se determina un indicador de estabilidad en conformidad con el método descrito en la publicación de patente "WO 2010/079167 A1", esto es, se calcula una medida de la sensibilidad. La medida de la sensibilidad predice cómo de sensible es la trama actual a los errores en el estado del filtro de predicción a largo plazo (LTP) debido a las pérdidas de paquetes. La medida de la sensibilidad se calcula en conformidad con la fórmula:

$$s = 0.5 PG_{LTP} + 0.5 PG_{LTP,HP},$$

15 en donde PG_{LTP} es la ganancia de predicción a largo plazo, según se mide como la relación de la energía de la señal residual de LPC (Codificación Predictiva Lineal) r_{LPC} y la señal residual de LTP (Predicción a Largo Plazo) r_{LTP} , y $PG_{LTP,HP}$ es una señal obtenida ejecutando PG_{LTP} por intermedio de un filtro de paso alto de primer orden en conformidad con:

$$20 \quad PG_{LTP,HP}(n) = PG_{LTP}(n) - PG_{LTP}(n-1) + 0.5 PG_{LTP,HP}(n-1).$$

La medida de la sensibilidad es una combinación de la ganancia de predicción de LTP y una versión de paso alto de la misma medida. La ganancia de predicción de LTP se elige puesto que se refiere directamente al error de estado de LTP con el error de la señal de salida. La parte de paso alto se añade para poner énfasis sobre los cambios de las señales. Una señal cambiante tiene un alto riesgo de proporcionar una propagación con errores graves puesto que el estado de LTP en el codificador y en el decodificador probablemente será muy diferente, después de la pérdida de paquetes.

La medida de la sensibilidad proporcionará, a la salida, un indicador que muestra la estabilidad de la imagen estéreo. En una etapa de comparación 331, el indicador se comprueba si es uno o cero. Si el indicador es igual a cero (ruta N), la imagen estéreo es estable y la diferencia de niveles entre canales CLDs no tiene ningún cambio importante entre dos tramas consecutivas. Si el indicador es igual a uno (ruta Y), la imagen estéreo no es estable, lo que significa que la diferencia de niveles entre canales CLDs entre dos tramas consecutivas cambian de forma muy rápida.

35 En una novena etapa 331, el indicador de estabilidad se comprueba que se proporciona a la salida desde la etapa anterior 329. Si el indicador de estabilidad es igual a uno (ruta Y), la memoria se actualiza en una décima etapa 317, es decir, las energías fuertemente suavizadas se actualizan por las energías débilmente suavizadas como sigue: $en_{1_sm}[b,i] = en_{1_sm_inst}[b,i]$ y $en_{2_sm}[b,i] = en_{2_sm_inst}[b,i]$, y en una undécima etapa 333 la versión débilmente suavizada de la diferencia de niveles entre canales CLD_inst se proporciona, a la salida como el parámetro de codificación CLD final. Si el indicador de estabilidad es igual a cero (ruta N), en una duodécima etapa 335, la versión fuertemente suavizada de la diferencia de niveles entre canales CLD se proporciona a la salida, como el parámetro de codificación final CLD.

45 Las etapas 309, 311 y 313 descritas con anterioridad pueden representarse como una etapa 301 que corresponde a la etapa 101 según se describe con respecto a la Figura 1b. Las etapas 315 y 321 descritas con anterioridad pueden representarse como una etapa 303 que corresponde a la etapa 103b según se describe con respecto a la Figura 1b. Las etapas 317, 319 y 323 descritas con anterioridad pueden representarse como una etapa 305 que corresponde a la etapa 105b según se describe con respecto a la Figura 1b. Las etapas 329, 331, 333 y 335 descritas con anterioridad pueden representarse como una etapa 307 que corresponde a la etapa 107b según se describe con respecto a la Figura 1b.

La Figura 4 ilustra un diagrama de bloques de un codificador de audio paramétrico 400 en conformidad con una forma de realización. El codificador de audio paramétrico 400 recibe una señal de audio multicanal 401 como la señal de entrada y proporciona un flujo de bits como la señal de salida 403. El codificador de audio paramétrico 400 comprende un generador de parámetros 405 acoplado a la señal de audio multicanal 401 para generar un parámetro de codificación 415, un generador de señales de mezcla descendente 407 acoplado a la señal de audio multicanal 401 para generar una señal de mezcla descendente 411 o señal suma, un codificador de audio 409 acoplado al generador de señales de mezcla descendente 407 para codificar la señal de mezcla descendente 411 para proporcionar una señal de audio codificada 413 y un dispositivo combinador 417, p.ej., un dispositivo conformador

de flujo de bits acoplado al generador de parámetros 405 y al codificador de audio 409 para formar un flujo de bits 403 a partir del parámetro de codificación 415 y la señal codificada 413.

El codificador de audio paramétrico 400 pone en práctica un sistema de codificación de audio para señales de audio multicanal y estéreo, que solamente transmite un canal de audio único, p.ej., la representación de mezcla descendente del canal de audio de entrada junto con parámetros adicionales que describen "diferencias perceptualmente pertinentes" entre los canales de audio, x_1, x_2, \dots, x_M . El sistema de codificación está en conformidad con la codificación de pistas binaurales (BCC) porque las pistas binaurales desempeñan un papel importante a este respecto. Según se indica en la Figura, los canales de audio de entrada x_1, x_2, \dots, x_M son objeto de mezcla descendente a un canal de audio único 411, también con la indicación de señal suma. Canales de audio x_1, x_2, \dots, x_M , el parámetro de codificación 415, p.ej., una diferencia de tiempo entre canales (ICTD), una diferencia de niveles entre canales (ICLD), y/o una coherencia entre canales (ICC), se estima como una función de la frecuencia y del tiempo y se transmite como información complementaria al decodificador 500 descrito en la Figura 5.

El generador de parámetros 405 que pone en práctica BCC procesa la señal de audio multicanal 401 con una determinada resolución de tiempo y frecuencia. La resolución de frecuencia utilizada está ampliamente motivada por la resolución de frecuencia del sistema auditivo. La psicoacústica sugiere que la percepción espacial es muy probable que esté basada en una representación de banda crítica de la señal de entrada acústica. Esta resolución de frecuencia se considera utilizando un banco de filtros invertibles con sub-bandas con anchos de bandas iguales o proporcionales al ancho de banda crítico del sistema auditivo. Es importante que la señal suma transmitida 411 contenga todos los componentes de señales de la señal de audio multicanal 401. El objetivo es que se mantenga completamente cada componente de señal. La suma simple de los canales de entrada de audio x_1, x_2, \dots, x_M de la señal de audio multicanal 401 suele dar lugar a la amplificación o atenuación de componentes de señales. Dicho de otro modo, la potencia de los componentes de señal en la suma "simple" suele ser mayor o menor que la suma de la potencia de la componente de señal correspondiente de cada canal x_1, x_2, \dots, x_M . Por lo tanto, una técnica de mezcla descendente se utiliza aplicando el dispositivo de mezcla descendente 407 que ecualiza la señal suma 411 de modo que la potencia de los componentes de señal en la señal suma 411 sea aproximadamente la misma que la potencia correspondiente en todos los canales de audio de entrada x_1, x_2, \dots, x_M de la señal de audio multicanal 401. Los canales de audio de entrada x_1, x_2, \dots, x_M están descompuestos en varias sub-bandas. Una de dichas sub-banda se indica como $X_i[b]$ (obsérvese que para simplicidad notacional no se utiliza un índice de sub-banda). Un procesamiento similar se aplica de forma independiente a todas las sub-bandas, normalmente las señales de sub-bandas se muestrean en sentido descendente. Las señales de cada sub-banda de cada canal de entrada se suman y luego se multiplican con un factor de normalización de potencia.

Dada la señal suma 411, el generador de parámetros 405 extrae parámetros de codificación espacial 415 tales como ICTD, ICLD, y/o ICC que se aproximan a la pistas correspondientes de la señal de audio multicanal original 401.

Cuando se considera la respuesta de impulsos ambiente binaurales (BRIRs) de una sola fuente, existe una relación entre la anchura de la envolvente del oyente y del evento auditivo y el valor de IC estimado para las partes iniciales y finales de la respuesta de impulsos de salas binaurales. Sin embargo, la relación entre IC o ICC y estas propiedades para señales generales y no solamente los valores de BRIRs no es directa. Las señales de audio multicanal y estéreo contienen una mezcla compleja de señales origen concurrentemente activas superpuestas por componente de señal reflejadas resultantes del registro en espacios cerrados o añadidas por el técnico de registro para crear artificialmente una impresión espacial. Diferentes señales de fuentes acústicas y sus reflexiones ocupan diferentes zonas en el plazo de tiempo-frecuencia. Esta circunstancia se refleja por los valores de ICTD, ICLD, e ICC que varían como una función del tiempo y de la frecuencia. En este caso, la relación entre los valores instantáneos de ICTD, ICLD, e ICC y las direcciones de eventos auditivos y la impresión espacial no es obvia. La estrategia del generador de parámetros 405 es una extracción a ciegas de estas pistas, de modo que se aproximen a las pistas correspondientes de la señal de audio original.

En una forma de realización, el codificador de audio paramétrico 400 utiliza bancos de filtros con sub-bandas de anchos de banda iguales a dos veces el ancho de banda rectangular equivalente. Una escucha informal descubrió que la calidad de audio de BCC no se mejoró notablemente cuando se eligió una más alta resolución de frecuencia. Una resolución de frecuencia más baja es favorable puesto que da lugar a valores menores de ICTD, ICLD, e ICC que necesitan transmitirse al decodificador y de este modo, en una tasa binaria más baja. Con respecto a la resolución en el tiempo, ICTD, ICLD, e ICC se consideran a los temporales periódicos. En una forma de puesta en práctica, los valores de ICTD, ICLD y de ICC se consideran aproximadamente cada 4-16 milisegundos. Conviene señalar que a no ser que las pistas se consideren en intervalos a muy corto plazo, no se considera directamente el efecto de prioridad.

La diferencia perceptualmente pequeña, frecuentemente conseguida, entre la señal de referencia y la señal sintetizada implica que las pistas relacionadas con una amplia gama de atributos de imágenes espaciales auditivas se consideran implícitamente sintetizando los valores de ICTD, ICLD, e ICC a intervalos periódicos. La tasa binaria requerida para la transmisión de estas pistas especiales es solamente de unos pocos kb/s y de este modo, el codificador de audio paramétrico 400 es capaz de transmitir señales de audio multicanal y estéreo a tasas binarias próximas a lo que se requiere para un canal de audio único. Las Figuras 1a y 2 ilustran un método en el que se

estima el valor de ITD como el parámetro de codificación 415. Las Figuras 1b y 3 ilustran un método en el que se estima el valor de CLD como el parámetro de codificación 415.

5 El codificador de audio paramétrico 400 comprende el generador de señales de mezcla descendente 407 para la superposición de al menos dos de las señales de canal de audio de la señal de audio multicanal 401 para obtener la señal de mezcla descendente 411, el codificador de audio 409, en particular un codificador monoaural, para codificar la señal de mezcla descendente 411 para obtener la señal de audio codificada 413 y el dispositivo combinador 417 para combinar la señal de audio codificada 413 con un parámetro de codificación 415 correspondiente.

10 El codificador de audio paramétrico 400 genera el parámetro de codificación 415 para una señal de canal de audio de la pluralidad de señales de canal de audio indicadas como x_1, x_2, \dots, x_M de la señal de audio multicanal 401. Cada una de las señales de canal de audio x_1, x_2, \dots, x_M puede ser una señal digital que comprende valores de señal de canal de audio digitales indicados como $x_1[n], x_2[n], \dots, x_M[n]$.

15 Una señal de canal de audio, a modo de ejemplo, para la que el codificador de audio paramétrico 400 genera el parámetro de codificación 415 es la primera señal de canal de audio x_1 con valores de señales $x_1[n]$. El generador de parámetros 405 determina el parámetro de codificación ITD a partir de los valores de señal de canal de audio $x_1[n]$ de la primera señal de audio x_1 y a partir de los valores de la señal de audio de referencia $x_2[n]$ de una señal de audio de referencia x_2 .

20 Una señal de canal de audio que se utiliza como una señal de audio de referencia es la segunda señal de canal de audio x_2 , a modo de ejemplo. De forma similar, cualquier otra de entre las señales de canal de audio x_1, x_2, \dots, x_M puede servir como señal de audio de referencia. En conformidad con un primer aspecto de la idea inventiva, la señal de audio de referencia es otra señal de canal de audio de las señales de canal de audio que no es igual a la señal de canal de audio x_1 para la que se genera el parámetro de codificación 415.

25 En conformidad con un segundo aspecto de la idea inventiva, la señal de audio de referencia es una señal de audio de mezcla descendente derivada a partir de al menos dos señales de canal de audio de entre la pluralidad de señales de audio multicanal 401, p.ej., derivadas de la primera señal de canal de audio x_1 y la segunda señal de canal de audio x_2 . En una forma de puesta en práctica, la señal de audio de referencia es la señal de mezcla descendente 411, también denominada señal suma generada por el dispositivo de mezcla descendente 407. En una forma de realización, la señal de audio de referencia es la señal codificada 413 proporcionada por el codificador 409.

30 Una señal de audio de referencia, a modo de ejemplo, utilizada por el generador de parámetros 405 es la segunda señal de canal de audio x_2 con valores de señales $x_2[n]$.

35 El generador de parámetros 405 determina una transformada de frecuencia de los valores de la señal de canal de audio $x_1[n]$ de la señal de canal de audio x_1 y una transformada de frecuencia de los valores de la señal de audio de referencia $x_2[n]$ de la señal de audio de referencia x_2 . La señal de audio de referencia es otra señal de canal de audio x_2 de la pluralidad de señales de canal de audio o una señal de audio de mezcla descendente derivada de al menos dos señales de canal de audio x_1, x_2 de entre la pluralidad de señales de canal de audio. El generador de parámetros 405 determina la diferencia entre canales para al menos cada sub-banda de frecuencias de un subconjunto de sub-bandas de frecuencia. Cada diferencia entre canales indica una diferencia de tiempo ITD[b] o diferencia de fase IPD[b] o una diferencia de nivel CLD[b] entre una parte de señal de banda limitada de la señal de canal de audio y una parte de señal de banda limitada de la señal de audio de referencia en la respectiva sub-banda de frecuencias a la que está asociada la diferencia entre canales

40 Una diferencia de fase entre canales (ICPD) es una diferencia de fase media entre un par de señales. Una diferencia de niveles entre canales (ICLD) es la misma que una diferencia de niveles inter-aural (ILD), esto es, una diferencia de niveles entre las señales de entrada a los oídos izquierdo y derecho, pero definida más en general entre cualquier par de señales, p.ej., un par de señales de altavoces, un par de señales de entradas de los oídos, etc. Una coherencia entre canales o una correlación entre canales es la misma que una coherencia inter-aural (IC), esto es, el grado de similitud entre las señales de entrada a los oídos izquierdo y derecho, pero definida más en general entre cualquier par de señales, p.ej., par de señales de altavoces, par de señales de entradas a los oídos. Una diferencia de tiempo entre canales (ICTD) es la misma que una diferencia de tiempo inter-aural (ITD), a veces también referida como un retardo inter-aural, esto es, una diferencia de tiempo entre las señales de entrada a los oídos izquierdo y derecho pero definida más en general entre cualquier par de señales, p.ej., par de señales de altavoces, par de señales de entrada a los oídos, etc. La diferencia de niveles entre canales de sub-banda, la diferencia de fase entre canales de sub-banda, las coherencias entre canales de sub-banda y las diferencias de intensidad entre canales de sub-banda están relacionadas con los parámetros anteriormente especificados con respecto al ancho de banda de sub-banda.

45 El generador de parámetros 405 está configurado para poner en práctica uno de los métodos descritos con respecto a las Figuras 1a, 1b, 2 y 3.

50 En una forma de realización, el generador de parámetros 405 comprende:

un primer dispositivo de determinación para determinar, para la señal de canal de audio (x_1), un conjunto de funciones ($c[b]$) a partir de los valores de la señal de canal de audio ($x_1[n]$) de la señal de canal de audio (x_1) y de los valores de la señal de audio de referencia ($x_2[n]$) de una señal de audio de referencia (x_2), en donde la señal de audio de referencia es otra señal de canal de audio (x_2) de entre la pluralidad de señales de canal de audio o una señal de audio de mezcla descendente derivada a partir de al menos dos señales de canal de audio (x_1 , x_2) de la pluralidad de señales de audio multicanal;

un segundo dispositivo de determinación para determinar un primer conjunto de parámetros de codificación (ITD[b], CLD[b]) sobre la base de un suavizado del conjunto de funciones ($c[b]$) con respecto a una secuencia de tramas (i) de la señal de audio multicanal, estando el suavizado basado en un primer coeficiente de suavizado (SMW_1);

un tercer dispositivo de determinación para determinar el segundo conjunto de parámetros de codificación (ITD_inst[b], CLD_inst[b]) sobre la base de un suavizado del conjunto de funciones ($c[b]$) con respecto a la secuencia de tramas (i) de la señal de audio multicanal, con el suavizado basado en un segundo coeficiente de suavizado (SMW_2); y

un dispositivo de determinación de parámetros de codificación para determinar el parámetro de codificación (ITD, CLD) sobre la base de un criterio de calidad con respecto al primer conjunto de parámetros de codificación (ITD[b], CLD[b]) y/o el segundo conjunto de parámetros de codificación (ITD_inst[b], CLD_inst[b]).

La Figura 5 ilustra un diagrama de bloques de un decodificador de audio paramétrico 500 en conformidad con una forma de realización. El decodificador de audio paramétrico 500 recibe un flujo de bits 503 transmitido a través de un canal de comunicaciones como señal de entrada y proporciona una señal de audio multicanal decodificada 501 como la señal de salida. El decodificador de audio paramétrico 500 comprende un decodificador de flujo de bits 517 acoplado al flujo de bits 503 para decodificar el flujo de bits 503 en un parámetro de codificación 515 y una señal codificada 513, un decodificador 509 acoplado al decodificador de flujo de bits 517 para generar una señal suma 511 a partir de la señal codificada 513, un denominado resolvidor de parámetros 505 acoplado al decodificador de flujo de bits 517 para la resolución de un parámetro 521 a partir del parámetro de codificación 515 y un sintetizador 505 acoplado al resolvidor de parámetros 505 y al decodificador 509 para sintetizar la señal de audio multicanal decodificada 501 a partir del parámetro 521 y la señal suma 511.

El decodificador de audio paramétrico 500 genera los canales de salida de su señal de audio multicanal 501 tal como ICTD, ICLD, y/o ICC entre los canales que se aproximan a los de la señal de audio multicanal original. El sistema descrito es capaz de representar señales de audio multicanal a una tasa binaria solamente ligeramente más alta que la que se requiere para representar una señal de audio monoaural. Esto es así, porque los valores estimados de ICTD, ICLD, e ICC entre un par de canales contienen aproximadamente dos órdenes de magnitud de menos información que una forma de onda de audio. No solamente la tasa binaria baja sino también el aspecto de compatibilidad con versiones anteriores es de interés. La señal suma transmitida corresponde a una mezcla descendente monoaural de la señal multicanal o estéreo.

La Figura 6 ilustra un diagrama de bloques de un codificador de audio estéreo paramétrico 601 y un decodificador 603 en conformidad con una forma de realización. El codificador de audio estéreo paramétrico 601 corresponde al codificador de audio paramétrico 400 según se describe con referencia a la Figura 4, pero la señal de audio multicanal 401 es una señal de audio estéreo con un canal de audio izquierdo 605 y derecho 607.

El codificador de audio estéreo paramétrico 601 recibe la señal de audio estéreo 605, 607 como una señal de entrada y proporciona un flujo de bits como la señal de salida 609. El codificador de audio estéreo paramétrico 601 comprende un generador de parámetros 611 acoplado a la señal de audio estéreo 605, 607 para generar parámetros espaciales 613, un generador de señales de mezcla descendente 615 acoplado a la señal de audio estéreo 605, 607 para generar una señal de mezcla descendente 617 o señal suma, un codificador monoaural 619 acoplado al generador de señales de mezcla descendente 615 para codificar la señal de mezcla descendente 617 para proporcionar una señal de audio codificada 621 y un dispositivo de combinación de flujo de bits 623 acoplado al generador de parámetros 611 y al codificador monoaural 619 para combinar el parámetro de codificación 613 y la señal de audio codificada 621 para un flujo de bits con el fin de proporcionar la señal de salida 609. En el generador de parámetros 611 los parámetros espaciales 613 son extraídos y cuantificados antes de multiplexarse en el flujo de bits.

El decodificador de audio paramétrico estéreo 603 recibe el flujo de bits, esto es, la señal de salida 609 del codificador de audio estéreo paramétrico 601 transmitida a través de un canal de comunicaciones, como una señal de entrada y proporciona una señal de audio estéreo decodificada con el canal izquierdo 625 y el canal derecho 627 como señal de salida. El decodificador de audio paramétrico estéreo 603 comprende un decodificador de flujo de bits 629 acoplado al flujo de bits 609 recibido para decodificar el flujo de bits 609 en parámetros de codificación 631 y una señal codificada 633, un decodificador monoaural 635 acoplado al decodificador de flujo de bits 629 para generar una señal suma 637 desde la señal codificada 633, un resolvidor de parámetros espaciales 639 acoplado al decodificador de flujo de bits 629 para la resolución de los parámetros espaciales 641 a partir de los parámetros de

codificación 631 y un sintetizador 643 acoplado al resolvidor de parámetros espaciales 639 y al decodificador monoaural 635 para sintetizar la señal de audio estéreo decodificada 625, 627 a partir de los parámetros espaciales 641 y la señal suma 637.

5 El procesamiento en el decodificador de audio paramétrico estéreo 603 es capaz de introducir retardos y modificar el nivel de las señales de audio de forma adaptativa en tiempo y frecuencia para generar los parámetros espaciales 631, p.ej., las diferencias de tiempo entre canales (ICTDs) y las diferencias de niveles entre canales (ICLDs). Además, el decodificador de audio paramétrico estéreo 603 realiza eficientemente el filtrado adaptativo del tiempo para la síntesis de la coherencia entre canales (ICC). En una forma de realización, el codificador estéreo paramétrico
10 utiliza un banco de filtros basado en la denominada Transformada de Fourier a corto plazo (STFT) para poner en práctica eficientemente, sistemas de codificación de pistas binaurales (BCC) con baja complejidad de cálculo. El procesamiento en el codificador de audio estéreo paramétrico 601 tiene baja complejidad de cálculo y bajo retardo, lo que hace a la codificación de audio estéreo paramétrica adecuada para una puesta en práctica asequible en microprocesadores o procesadores de señales digitales para aplicaciones en tiempo real.

15 El generador de parámetros 611 ilustrado en la Figura 6 es funcionalmente el mismo que el generador de parámetros correspondiente 405 descrito con respecto a la Figura 4, con la excepción de que se ha añadido la cuantización y codificación de las pistas espaciales. La señal suma 617 se codifica con un codificador de audio monoaural 619 convencional. En una forma de realización, el codificador de audio estéreo paramétrico 601 utiliza una transferencia de tiempo-frecuencia basada en STFT para transformar la señal de canal de audio 605, 607 en el dominio de la frecuencia. La transformada STFT aplica una Transformada de Fourier Discreta (DFT) a partes en ventanas operativas de una señal de entrada $x(n)$. Una trama de señales de N muestras se multiplica con una ventana de longitud W antes de que se aplique una DFT de N puntos. Las ventanas adyacentes están en solapamiento y se desplazan en $W/2$ muestras. La ventana se elige de modo que las ventanas en solapamiento se
20 añadan a un valor constante de 1. Por lo tanto para la transformada inversa, no existe necesidad de creación de ventanas adicionales. Una transformada de DFT inversa simple de tamaño N con anticipación de tramas sucesivas de $W/2$ muestras se utiliza en el decodificador 603. Si el espectro no es modificado, se consigue una reconstrucción perfecta mediante solapamiento/adición.

30 Puesto que la resolución espectral uniforme de la transformada STFT no está bien adaptada para la percepción humana, la salida de coeficientes espectrales uniformemente espaciados de la transformada STFT se agrupan en B particiones no solapadas con anchos de banda mejor adaptados para la percepción. Una partición corresponde conceptualmente a una "sub-banda" en conformidad con la descripción con respecto a la Figura 4. En una forma de realización alternativa, el codificador de audio estéreo paramétrico 601 utiliza un banco de filtros no uniforme para transformar la señal de canal de audio 605, 607 en el dominio de la frecuencia.

35 En una forma de realización, el dispositivo de mezcla descendente 615 determina los coeficientes espectrales de una partición b o de una sub-banda b de la señal suma ecualizada $S_m(k)$ 617 mediante

$$S_m(k) = e_b(k) \sum_{c=1}^C X_{c,m}(k),$$

40 en donde $X_{c,m}(k)$ son los espectros de los canales de audio de entrada 605, 607 y $e_b(k)$ es un factor de ganancia calculado como

$$e_b(k) = \sqrt{\frac{\sum_{c=1}^C p_{\tilde{x}_{c,b}}(k)}{p_{\tilde{x}_b}(k)}},$$

45 con estimaciones de potencia de las particiones,

$$p_{\tilde{x}_{c,b}}(k) = \sum_{m=A_{b-1}}^{A_b-1} |X_{c,m}(k)|^2$$

$$p_{\tilde{x}_b}(k) = \sum_{m=A_{b-1}}^{A_b-1} \left| \sum_{c=1}^C X_{c,m}(k) \right|^2.$$

50 Para impedir que resulten artefactos operativos a partir de factores de amplia ganancia cuando la atenuación de la suma de las señales de sub-bandas es importante, los factores de ganancia $e_b(k)$ están limitados a 6 dB, esto es,

eb(k) < 2.

En una forma de realización del codificador de audio estéreo paramétrico 601 y del decodificador 603, el tipo de información de ITD (banda completa) se señala a los decodificadores distantes 603. En una forma de realización, la señalización del tipo se realiza mediante una señalización implícita por medio de datos auxiliares transportados en al menos un flujo de bits. En una forma de realización alternativa, la señalización se realiza mediante una señalización explícita por medio de un indicador del tipo del flujo de bits respectivo. En una forma de realización, es posible conmutar entre una primera opción de señalización que comprende una señalización implícita y una segunda opción de señalización que comprende una señalización explícita. En una forma de realización de la señalización implícita, un indicador señala una presencia de la información de canal secundario en datos auxiliares de al menos un flujo de bits compatible en sentido inverso. El decodificador de legado no comprueba si está presente un indicador, o no lo está, y solamente decodifica el flujo de bits compatible en sentido inverso. A modo de ejemplo, la señalización del flujo de bits de canal secundario puede incluirse en los datos auxiliares de un flujo de bits de AAC. Además, el flujo de bits secundario puede incluirse también en los datos auxiliares de un flujo de bits de AAC. En ese caso, un decodificador de AAC de legado decodifica solamente la parte compatible en sentido inverso del flujo de bits y rechaza los datos auxiliares. En una forma de realización del codificador de audio estéreo paramétrico 601 y del decodificador 603, se comprueba la presencia de dicho indicador y si el indicador está presente en el flujo de bits recibido, el decodificador 603 reconstruye la señal de audio multicanal sobre la base de la información de ITD de banda completa adicional.

En una forma de realización de la señalización explícita, un indicador de que el flujo de bits es un nuevo flujo de bits obtenido con un nuevo codificador no de legado se utiliza a este respecto. Un decodificador de legado no es capaz de decodificar el flujo de bits puesto que no conoce cómo interpretar este indicador. Sin embargo, el decodificador 603 en conformidad con una forma de realización tiene la capacidad para decodificar y para decidir decodificar la parte compatible en sentido inverso solamente o la señal de audio multicanal completa.

Una ventaja de dicha compatibilidad en sentido inverso puede deducirse como sigue. Un terminal móvil que comprende un decodificador 603 en conformidad con una forma de realización puede decidir decodificar la parte compatible en sentido inverso para economizar la vida útil de la batería de una batería integrada puesto que la carga de complejidad es más baja. Además, dependiendo del denominado sistema de renderizado, el decodificador 603 puede decidir qué parte decodificar del flujo de bits. A modo de ejemplo, para la operación de renderizado con un auricular, la parte compatible en sentido inverso de la señal recibida puede ser suficiente, mientras que la señal de audio multicanal se decodifica solamente cuando el terminal está conectado, a modo de ejemplo, a una estación de carga de acoplamiento con una denominada capacidad de renderizado multicanal.

En una forma de realización, el método según se describe con respecto a una de las Figuras 1a, 1b, 2 y 3 se aplica en un codificador de la extensión estéreo de ITU-T G.722, G.722 Anexo B, G.711.1 y/o G.711.1 Anexo D. Además, en una forma de realización, el método según se describe con respecto a una de las Figuras 1a, 1b, 2 y 3 se aplica para el codificador de audio y vocal para aplicación móvil según se define en el codificador-decodificador de 3GGP EVS (Servicios Vocales Mejorados).

En una forma de realización según se describe con referencia a una de las Figuras 1a, 1b, 2 y 3 se utiliza para análisis de escenas auditivas. En ese caso, una de las formas de realización de la estimación de ITD o de la estimación de CLD se utiliza sola o en combinación para evaluar la característica de la imagen espacial y para detectar la posición de la fuente de sonido en la escena operativa de audio.

La Figura 7 ilustra un diagrama esquemático de un algoritmo de selección de ITD en conformidad con una forma de realización.

En una primera etapa 701, el número Nb_{pos} de valores ITD positivos se comprueba con respecto al número Nb_{neg} de valores ITD negativos. Si Nb_{pos} es mayor que Nb_{neg} , se realiza la etapa 703; si Nb_{pos} no es mayor que Nb_{neg} , se realiza la etapa 705.

En la etapa 703, la desviación estándar ITD_{std_pos} de los valores ITDs positivos se comprueba contra la desviación estándar ITD_{std_neg} de ITDs negativos y el número Nb_{pos} de valores ITD positivos se comprueba con respecto al número Nb_{neg} de valores ITD negativos multiplicado por un primer factor A, p.ej., en conformidad con la expresión: $(ITD_{std_pos} < ITD_{std_neg}) \parallel (Nb_{pos} \geq A * Nb_{neg})$. Si $ITD_{std_pos} < ITD_{std_neg}$ o $Nb_{pos} > A * Nb_{neg}$, el valor de ITD se selecciona como el valor medio de los valores ITD positivos en la etapa 707. De no ser así, la relación entre valores ITD positivos y negativos se comprobará además en la etapa 709.

En la etapa 709, la desviación estándar ITD_{std_neg} de valores ITDs negativos se comprueba con respecto a la relación estándar ITD_{std_pos} de valores ITDs positivos multiplicado por un segundo factor B, p.ej., en conformidad con: $(ITD_{std_neg} < B * ITD_{std_pos})$. Si se verifica que $ITD_{std_neg} < B * ITD_{std_pos}$, el valor opuesto del valor medio de ITD negativo será seleccionado como el valor ITD de salida en la etapa 715. De no ser así, el valor ITD de la trama anterior (Pre_itd) se comprueba en la etapa 717.

En la etapa 717, el valor ITD de la trama anterior se comprueba para verificar si es mayor que cero, p.ej., en conformidad con "Pre_itd > 0". Si se verifica Pre_itd > 0, se selecciona el valor ITD de salida como el valor medio de ITD positivo en la etapa 723 y de no ser así, el valor ITD de salida es el valor opuesto del valor medio ITD negativo en la etapa 725.

5 En la etapa 705, la desviación estándar ITD_{std_neg} de los valores ITDs negativos se comprueba con respecto a la desviación estándar ITD_{std_pos} de valores ITDs positivos y el número Nb_{neg} de valores ITDs negativos se comprueba con respecto al número Nb_{pos} de valores ITDs positivos multiplicado por un primer factor A, p.ej., en conformidad con: $(ITD_{std_neg} < ITD_{std_pos}) \parallel (Nb_{neg} \geq A * Nb_{pos})$. Si $ITD_{std_neg} < ITD_{std_pos}$ o $Nb_{neg} > A * Nb_{pos}$, se selecciona el valor ITD como el valor medio de ITD negativo en la etapa 711. De no ser así, la relación entre valores de ITD negativos y positivos se comprueba, además, en la etapa 713.

15 En la etapa 713, la desviación estándar ITD_{std_pos} de valores ITDs positivos se comprueba con respecto a la desviación estándar ITD_{std_neg} de valores ITDs negativos multiplicado por un segundo factor B, p.ej., en conformidad con: $(ITD_{std_pos} < B * ITD_{std_neg})$. Si $ITD_{std_pos} < B * ITD_{std_neg}$, el valor positivo del valor medio de ITD positivo se selecciona como el valor ITD de salida en la etapa 719. De no ser así, el valor de ITD desde la trama anterior (Pre_itd) se comprueba en la etapa 721.

20 En la etapa 721, el valor ITD de la trama anterior se comprueba para determinar si es mayor que cero, p.ej., en conformidad con "Pre_itd > 0". Si Pre_itd > 0, se selecciona el valor ITD de salida como el valor medio del ITD negativo en la etapa 727; de no ser así, el valor ITD de salida es el valor opuesto del valor medio de ITD positivo en la etapa 729.

25 La selección entre el ITD basado en la versión fuertemente suavizada del espectro cruzado (ITD_{mean}) y la versión de ITD débilmente suavizada del espectro cruzado (ITD_{mean_inst}) se obtiene por separado para el valor ITD positivo y el valor ITD negativo. Por último, la decisión sobre ITD se realiza según se describe en la Figura 7.

30 De lo que antecede resultará evidente para los expertos en esta técnica que se dan a conocer una diversidad de métodos, sistemas, programas informáticos en soporte de registro y dispositivos similares.

La presente invención soporta también un producto de programa informático que incluye un código ejecutable por ordenador o instrucciones ejecutables por ordenador que, cuando se ejecutan, hacen al menos que un ordenador ejecute las etapas de realización y cálculo aquí descritas.

35 La presente invención soporta también un sistema configurado para ejecutar las etapas de realización y cálculo aquí descritas.

40 Numerosas alternativas, modificaciones y variaciones serán evidentes para los expertos en esta técnica teniendo en cuenta las enseñanzas anteriores. Por supuesto, los expertos en esta técnica reconocen fácilmente que existen numerosas aplicables de la invención más allá de las aquí descritas. Aunque la presente invención ha sido descrita con referencia a una o más formas de realización particulares, los expertos en esta técnica reconocen que pueden efectuarse numerosos cambios a la misma sin desviarse por ello del alcance de la presente invención. Por lo tanto, ha de entenderse que dentro del alcance de las reivindicaciones adjuntas, la invención puede ponerse en práctica de cualquier otro modo distinto al modo aquí descrito.

45

50

REIVINDICACIONES

- 5 **1.** Un método (100) para determinar un parámetro de codificación (ITD) para una señal de canal de audio (x_1) de una pluralidad de señales de canal de audio (x_1, x_2) de una señal de audio multicanal, teniendo cada señal de canal de audio (x_1, x_2) valores de señal de canal de audio ($x_1[n], x_2[n]$), cuyo método comprende:
- la determinación (101) para la señal de canal de audio (x_1) de un conjunto de funciones ($c[b]$) a partir de los valores de la señal de canal de audio ($x_1[n]$) de la señal de canal de audio (x_1) y valores de señal de audio de referencia ($x_2[n]$) de una señal de audio de referencia (x_2), en donde la señal de audio de referencia es otra señal de canal de audio (x_2) de entre la pluralidad de señales de canal de audio o una señal de audio de mezcla descendente derivada de al menos dos señales de canal de audio (x_1, x_2) de la pluralidad de señales de audio multicanal;
- 10 la determinación (103) de un primer conjunto de parámetros de codificación (ITD[b]) sobre la base de un suavizado operativo del conjunto de funciones ($c[b]$) con respecto a una secuencia de tramas (i) de la señal de audio multicanal, estando la función de suavizado basada en un primer coeficiente de suavizado (SMW₁);
- 15 la determinación (105) de un segundo conjunto de parámetros de codificación (ITD_inst[b]) sobre la base de un suavizado del conjunto de funciones ($c[b]$) con respecto a la secuencia de tramas (i) de la señal de audio multicanal, estando el suavizado basado en un segundo coeficiente de suavizado (SMW₂); y
- 20 la determinación (107) del parámetro de codificación (ITD) sobre la base de un criterio de calidad con respecto al primer conjunto de parámetros de codificación (ITD[b]) y/o el segundo conjunto de parámetros de codificación (ITD_inst[b]).
- 25 **2.** El método según la reivindicación 1, en donde la determinación (101) del conjunto de funciones ($c[b]$) comprende:
- la determinación de una transformada de frecuencia ($X_1[k]$) de los valores señal de canal de audio ($x_1[n]$) de la señal de canal de audio (x_1);
- 30 la determinación (103) de una transformada de frecuencia ($X_2[k]$) de los valores de señal de audio de referencia ($x_2[n]$) de la señal de audio de referencia (x_2);
- 35 la determinación (105) del conjunto de funciones ($c[b]$) como un espectro cruzado o correlación cruzada para al menos cada sub-banda de frecuencias (b) de un subconjunto de sub-bandas de frecuencia, siendo cada función del conjunto de funciones ($c[b]$) calculada entre una parte de señal de banda limitada de la señal de canal de audio y una parte de señal de banda limitada de la señal de audio de referencia en la respectiva sub-banda de frecuencias (b), en donde la función del conjunto de funciones ($c[b]$) está asociada.
- 40 **3.** El método (100) según la reivindicación 2, en donde una sub-banda de frecuencias comprende una o una pluralidad de contenedores de frecuencias (k).
- 45 **4.** El método (100) según cualquiera de las reivindicaciones precedentes, en donde el primero (ITD[b]) y el segundo (ITD_inst[b]) conjuntos de parámetros de codificación comprenden diferencias entre canales (ICD[b]), en donde las diferencias entre canales (ICD[b]) comprenden diferencias de tiempo entre canales (ITD[b]) y/o diferencias de nivel entre canales (CLD[b]).
- 50 **5.** El método (100) según cualquiera de las reivindicaciones precedentes, en donde la determinación (107) del parámetro de codificación (ITD) sobre la base de un criterio de calidad comprende la determinación de un parámetro de estabilidad, siendo el parámetro de estabilidad utilizado por el criterio de calidad.
- 55 **6.** El método (100) según la reivindicación 5, en donde la determinación (107) del parámetro de codificación (ITD, CLD) comprende:
- la determinación de un parámetro de estabilidad del segundo conjunto de parámetros de codificación (ITD_inst[b], CLD_inst[b]) sobre la base de una comparación entre valores consecutivos del segundo conjunto de parámetros de codificación (ITD_inst[b], CLD_inst[b]) con respecto a la secuencia de tramas (i); y
- 60 la determinación del parámetro de codificación (ITD, CLD) dependiendo del parámetro de estabilidad.
- 7.** El método (100) según la reivindicación 5, en donde el parámetro de estabilidad está basado al menos en una desviación estándar (ITD_inst_{std}) del segundo conjunto de parámetros de codificación (ITD_inst[b], CLD_inst[b]).
- 65 **8.** El método (100) según la reivindicación 6 o la reivindicación 7, en donde el parámetro de estabilidad se determina durante una sola trama o durante múltiples tramas de la señal de audio multicanal.

9. El método (100) según la reivindicación 6, la reivindicación 7 o la reivindicación 8, en donde la determinación del parámetro de codificación (ITD, CLD) se determina sobre la base de un cruce de umbral del parámetro de estabilidad.

5 10. El método (100) según la reivindicación 9 que comprende, además:

la actualización del primer conjunto de parámetros de codificación (ITD[b], CLD[b]) con el segundo conjunto de parámetros de codificación (ITD_inst[b], CLD_inst[b]) si el parámetro de estabilidad cruza el umbral.

10 11. El método (100) según cualquiera de las reivindicaciones precedentes, en donde el suavizado del conjunto de funciones (c[b]) basado en un primero y segundo coeficiente de suavizado (SMW_1 , SMW_2) se calcula como una adición de un estado de memoria de las primera y segunda versiones suavizadas del conjunto de funciones multiplicado por un primer coeficiente basado en el primero y el segundo coeficiente de suavizado (SMW_1 , SMW_2) y el conjunto de funciones (c[b]) multiplicado por un segundo coeficiente basado en el primero y el segundo coeficiente de suavizado (SMW_1 , SMW_2).

15 12. El método (100) según la reivindicación 11, que comprende, además:

la actualización del estado de memoria de la primera versión suavizada del conjunto de funciones (c[b]) con el estado de memoria de la segunda versión suavizada del conjunto de funciones (c[b]) si el parámetro de estabilidad cruza el umbral.

20 13. El método (100) según cualquiera de las reivindicaciones precedentes, en donde el primer coeficiente de suavizado (SMW_1) es más alto que el segundo coeficiente de suavizado (SMW_2).

25 14. Un codificador de audio multicanal (400) para determinar un parámetro de codificación (ITD) para una señal de canal de audio (x_1) de una pluralidad de señales de canal de audio (x_1 , x_2) de una señal de audio multicanal, teniendo cada señal de canal de audio (x_1 , x_2) valores de señal de canal de audio ($x_1[n]$, $x_2[n]$), comprendiendo el codificador de audio multicanal:

30 un primer dispositivo de determinación para determinar, para la señal de canal de audio (x_1), un conjunto de funciones (c[b]) a partir de los valores de señal de canal de audio ($x_1[n]$) de la señal de canal de audio (x_1) y de los valores de señal de audio de referencia ($x_2[n]$) de una señal de audio de referencia (x_2), en donde la señal de audio de referencia es otra señal de canal de audio (x_2) de la pluralidad de señales de canal de audio o una señal de audio de mezcla descendente derivada de al menos dos señales de canal de audio (x_1 , x_2) de la pluralidad de señal de audio multicanales;

35 un segundo dispositivo de determinación para determinar un primer conjunto de parámetros de codificación (ITD[b]) sobre la base de una función de suavizado del conjunto de funciones (c[b]) con respecto a una secuencia de tramas (i) de la señal de audio multicanal, estando el suavizado basado en un primer coeficiente de suavizado (SMW_1);

40 un tercer dispositivo de determinación para determinar un segundo conjunto de parámetros de codificación (ITD_inst[b]) sobre la base de un suavizado del conjunto de funciones (c[b]) con respecto a la secuencia de tramas (i) de la señal de audio multicanal, estando el suavizado basado en un segundo coeficiente de suavizado (SMW_2); y

45 un dispositivo de determinación de parámetro de codificación para determinar el parámetro de codificación (ITD) sobre la base de un criterio de calidad con respecto al primer conjunto de parámetros de codificación (ITD[b]) y/o el segundo conjunto de parámetros de codificación (ITD_inst[b]).

50 15. Un programa informático que tiene un código de programa para realizar el método (100) según una de las reivindicaciones 1 a 13 cuando se ejecuta en un ordenador.

55

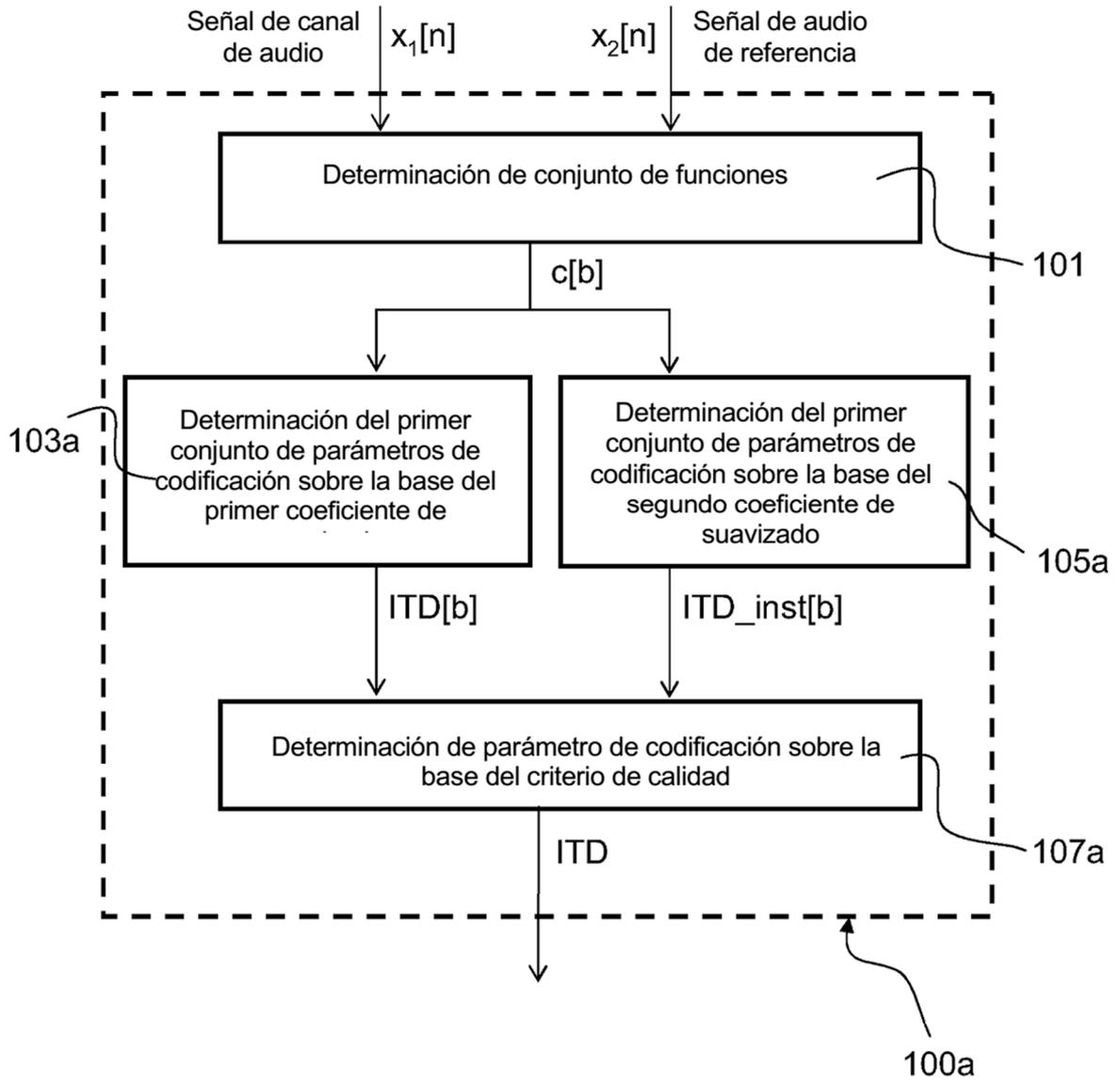


Fig. 1a

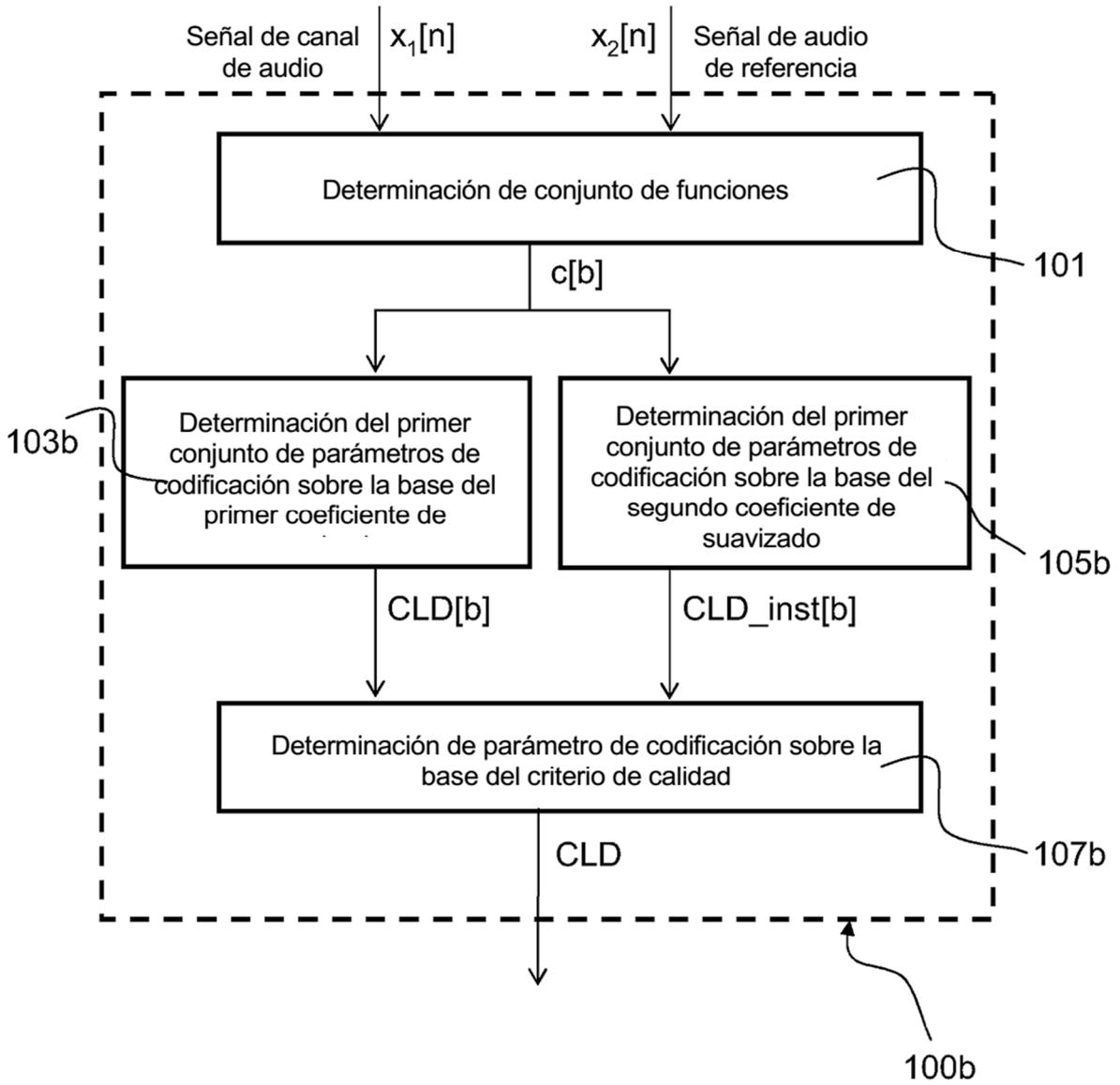


Fig. 1b

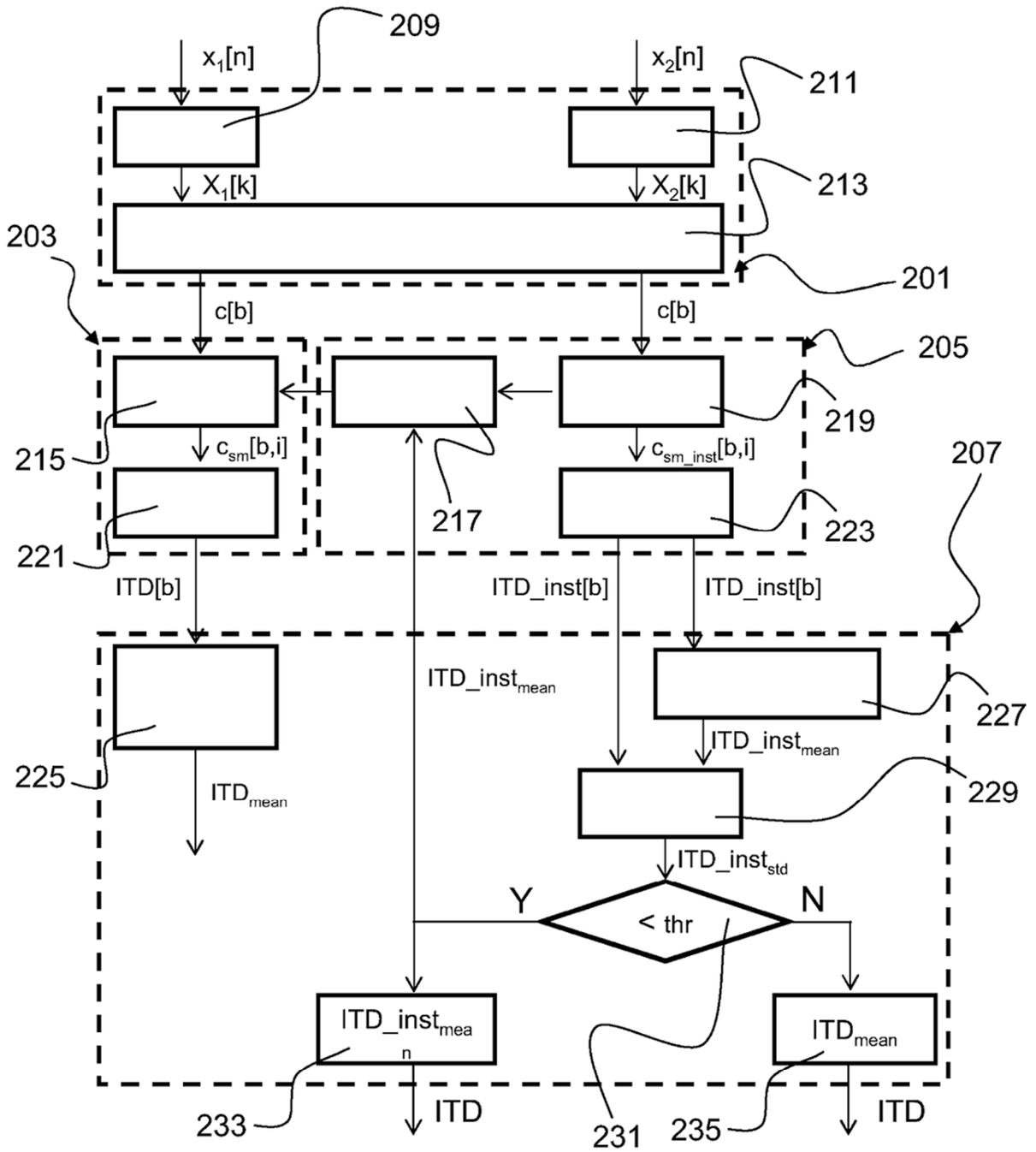


Fig. 2

200

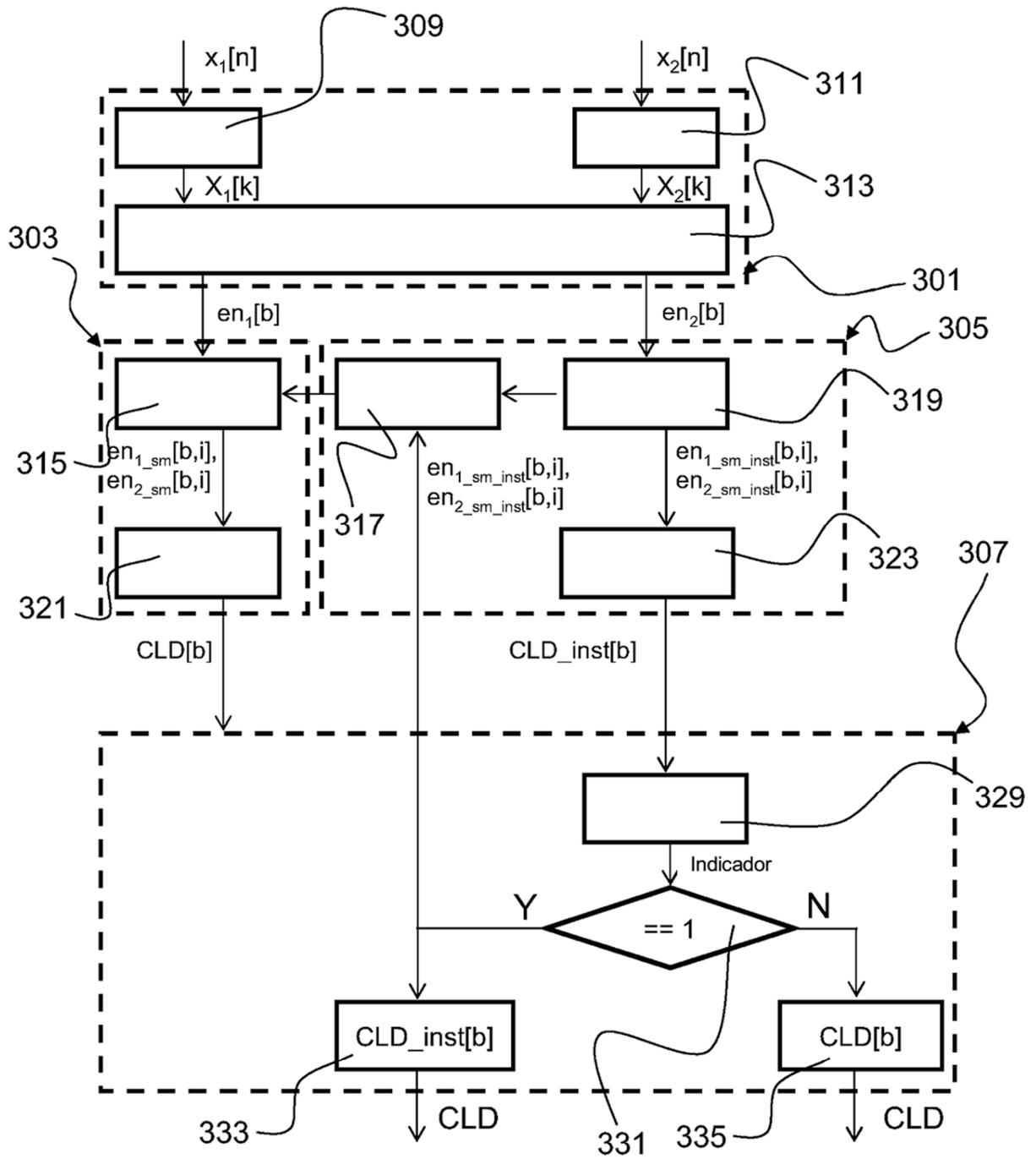


Fig. 3

300

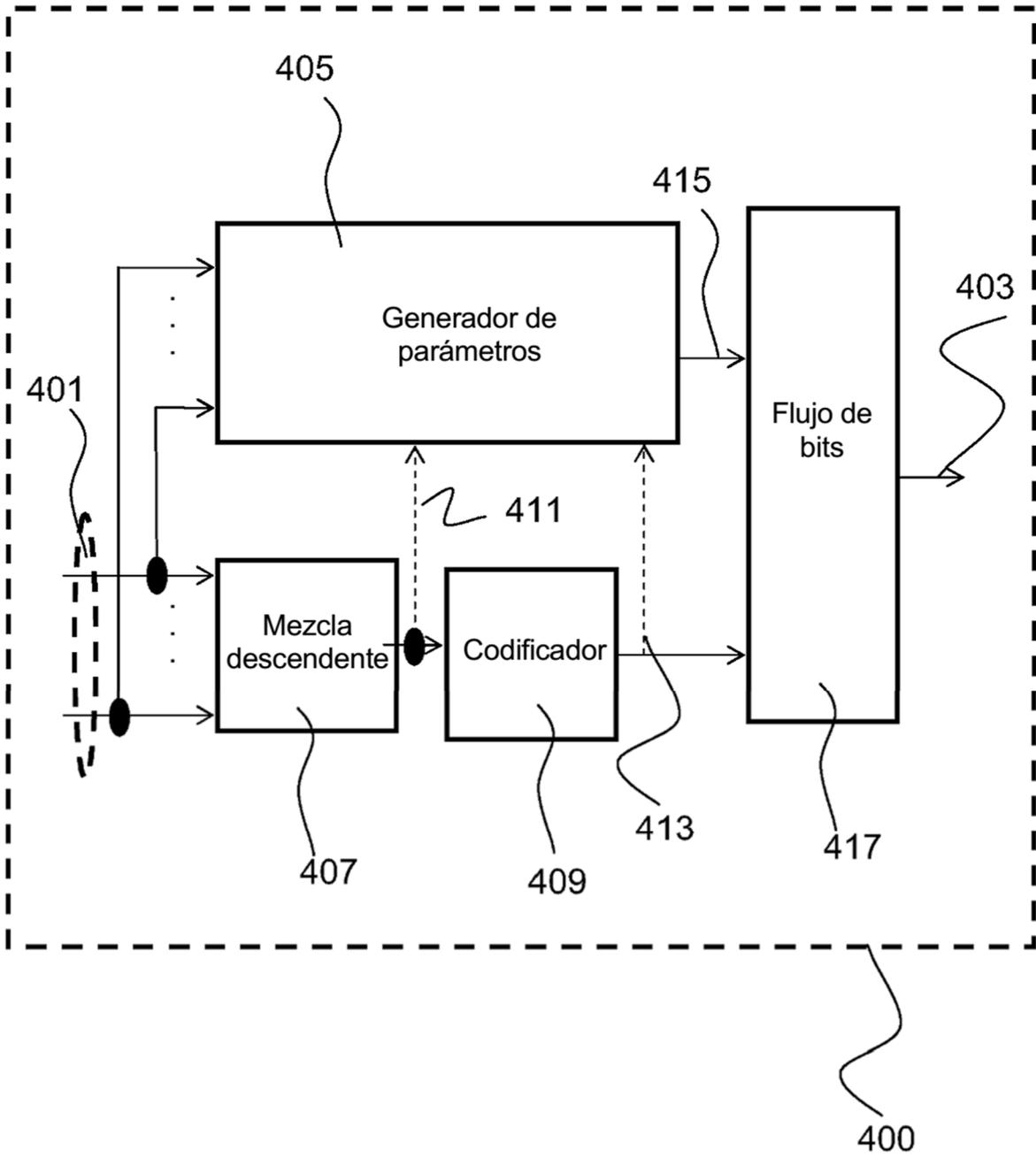


Fig. 4

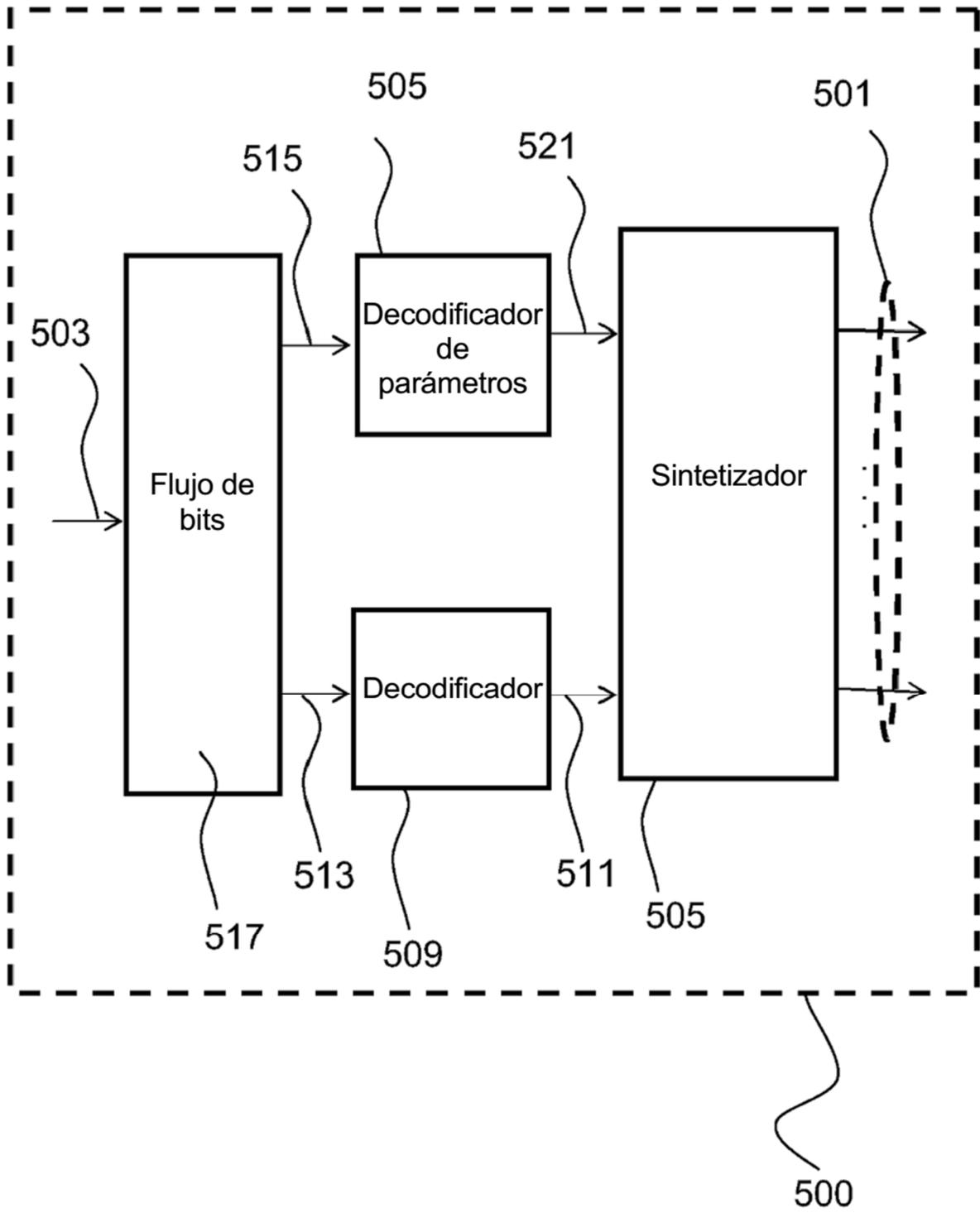


Fig. 5

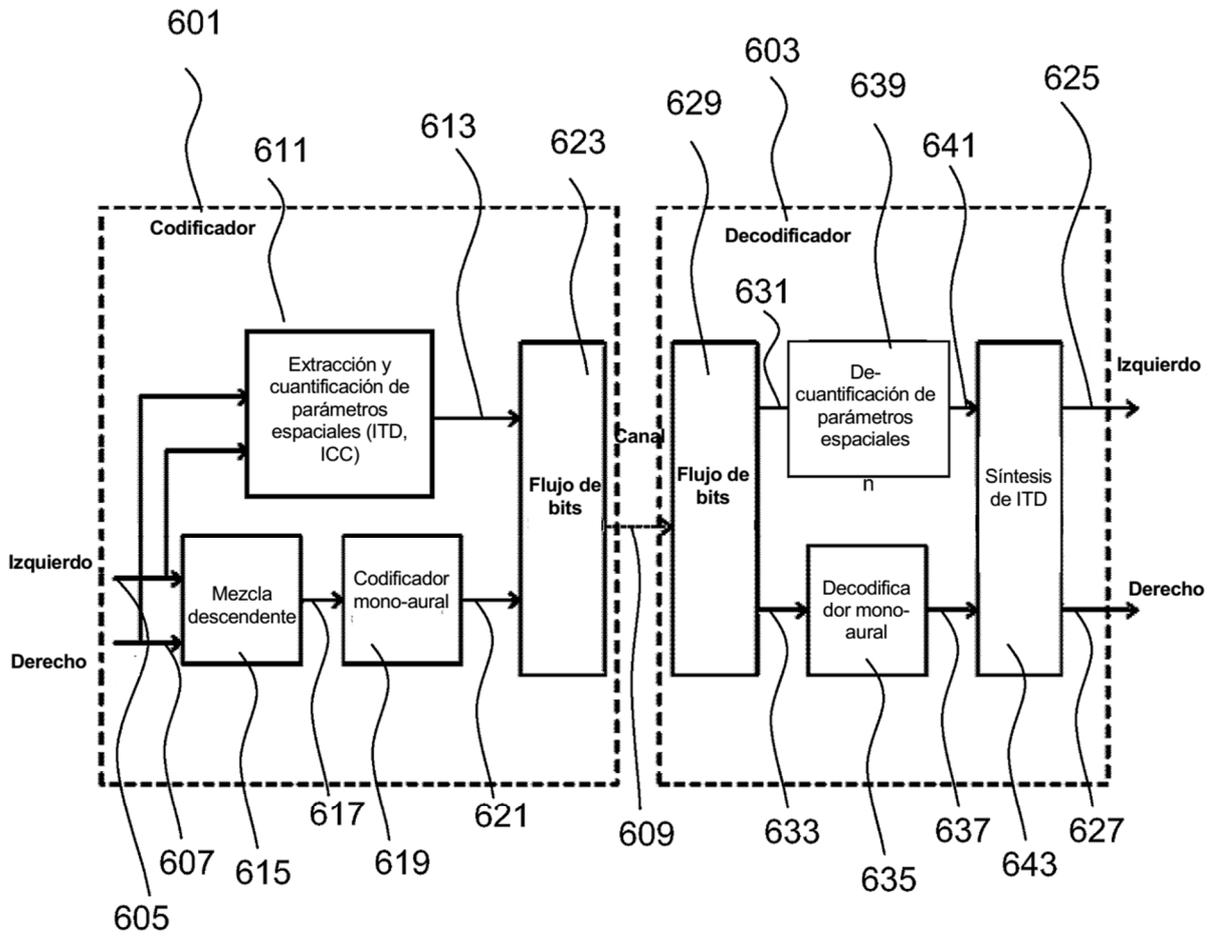


Fig. 6

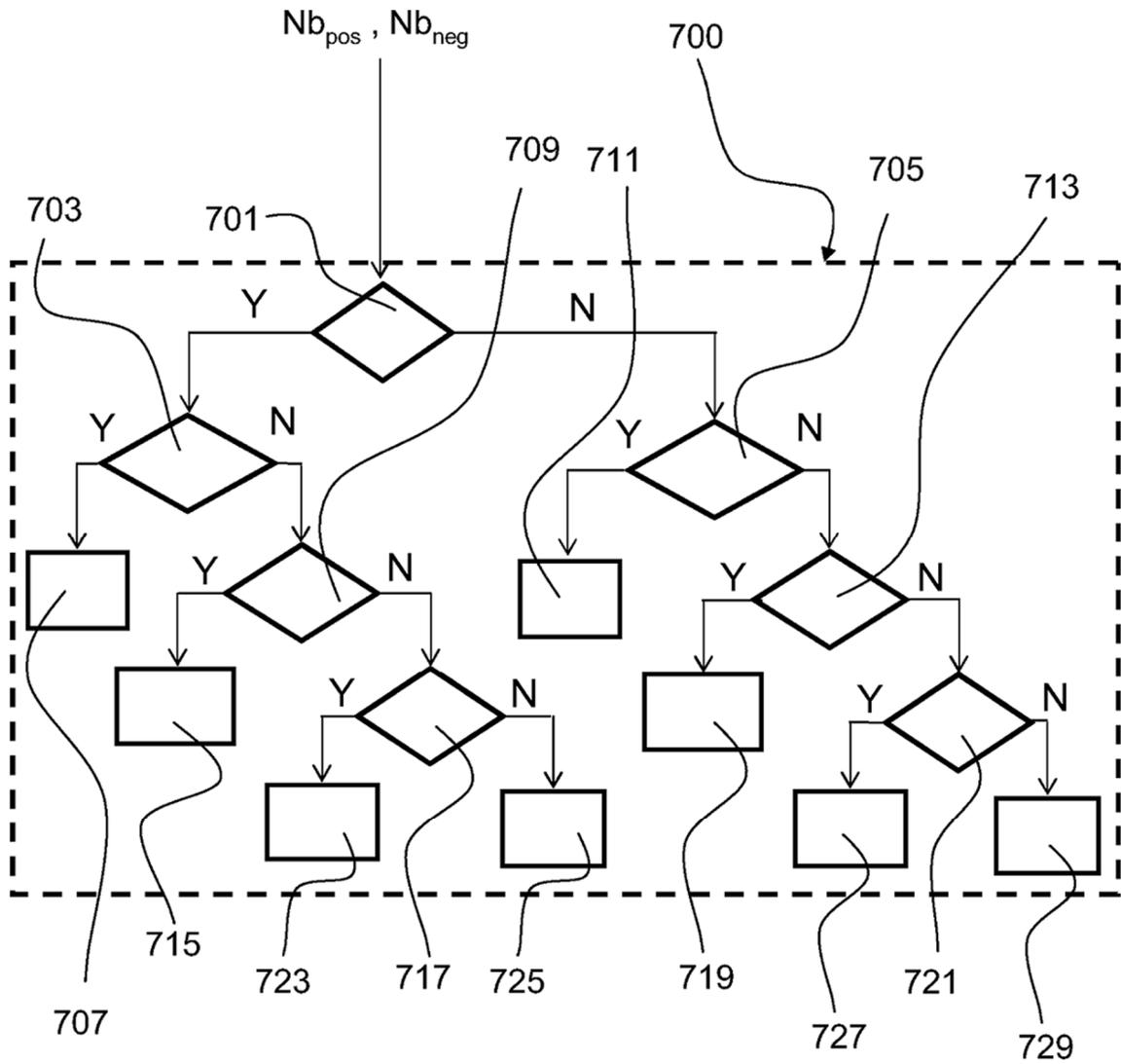


Fig. 7

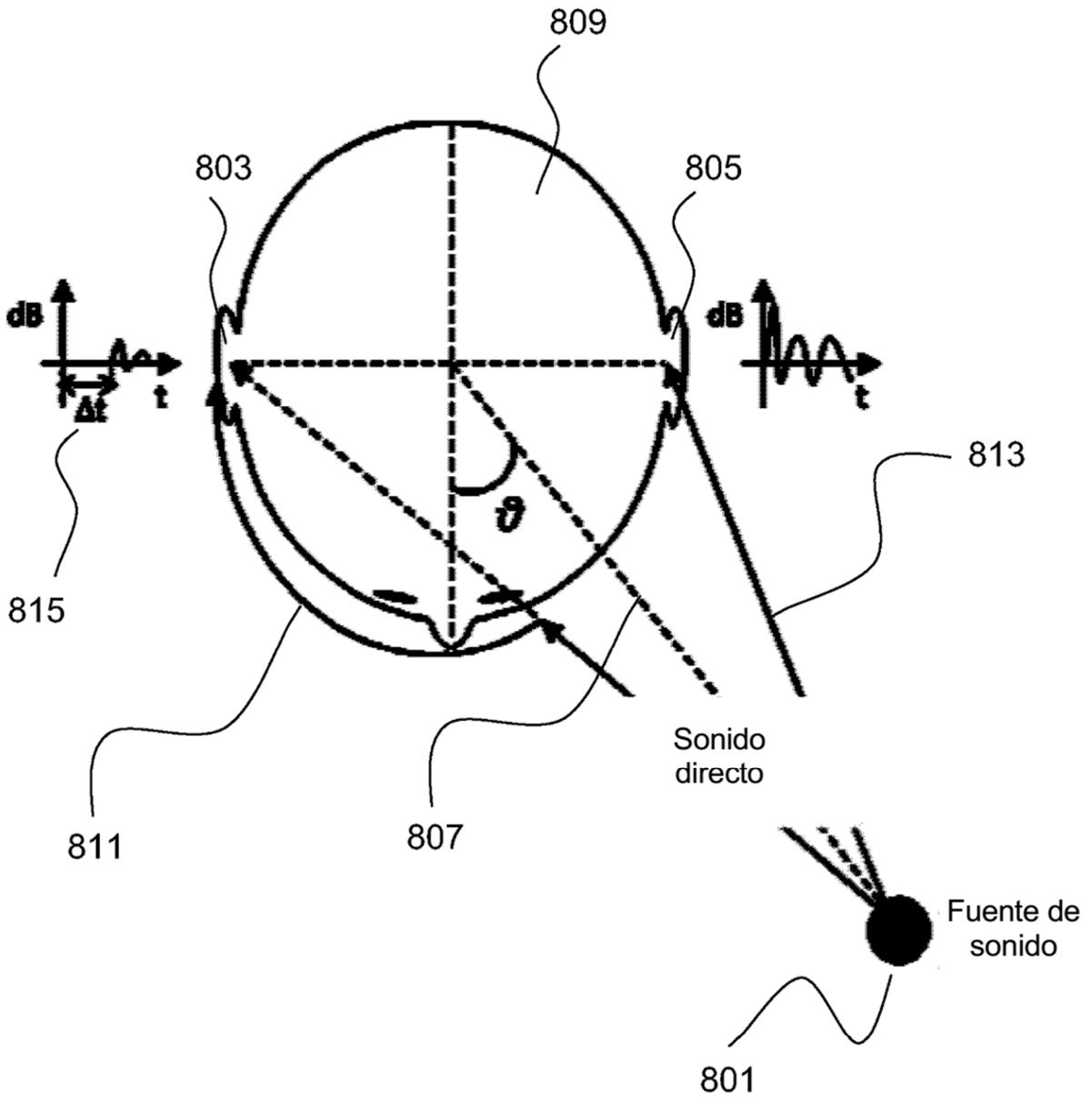


Fig. 8