

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 555 136**

51 Int. Cl.:

H04S 3/00 (2006.01)

G10L 19/00 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **17.02.2012** **E 12707055 (5)**

97 Fecha y número de publicación de la concesión europea: **23.09.2015** **EP 2702776**

54 Título: **Codificador paramétrico para codificar una señal de audio multicanal**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
29.12.2015

73 Titular/es:

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

72 Inventor/es:

LANG, YUE;
VIRETTE, DAVID y
XU, JIANFENG

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 555 136 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

Codificador paramétrico para codificar una señal de audio multicanal

5 **CAMPO DE LA INVENCION**

La presente invención se refiere a la codificación de audio.

10 **ANTECEDENTES DE LA INVENCION**

10 La codificación de audio paramétrica en estéreo o multicanal según se describe, p.e., en C. Faller y F. Baumgarte, "Representación eficiente de la señal de audio espacial utilizando una parametrización perceptual", en Proc. IEEE Workshop on Appl. of Sig. Proc. para audio y acústica, octubre 2001, páginas 199-202, utiliza pistas espaciales para sintetizar señales de audio multicanal a partir de señales de audio mezcladas reducidas - normalmente mono o
15 estéreo - teniendo las señales de audio multicanal más canales que las señales de audio mezcladas reducidas. En condiciones normales, las señales de audio mezcladas reducidas resultan de una superposición de una pluralidad de señales de canales de audio de una señal de audio multicanal, p.e., de una señal de audio estéreo. Estos canales están codificados en forma de onda e información secundaria, esto es, las pistas espaciales, referidas a las relaciones de señales de canales originales se añaden como parámetros de codificación a los canales de audio
20 codificados. El decodificador utiliza esta información secundaria para regenerar el número original de canales de audio sobre la base de los canales codificados en forma de onda decodificados.

25 Un codificador estéreo paramétrico básico puede utilizar las diferencias de niveles intercanales (ILD) como una pista necesaria para generar la señal de estéreo a partir de la señal de audio mezclada reducida monoaural. Codificadores más sofisticados pueden utilizar también la coherencia intercanal (ICC), que puede representar un grado de similitud entre las señales de canales de audio, esto es, canales de audio. Además, cuando se codifican señales estéreo binaurales, p.e., para señales de audio en 3D o la prestación circundante basada en auriculares, también una diferencia de fase intercanal (IPD) puede desempeñar una función importante para reproducir
30 diferencias de fase/retardo entre los canales.

30 La síntesis de pistas de ICC puede ser de importancia para la mayoría de los contenidos de audio y musicales para regenerar el entorno ambiental, reverberación de estéreo, anchura de fuente y otras percepciones relacionadas con la impresión especial según se describe en el documento de J. Blauert, Audiencia espacial: La psicofísica de la localización del sonido humano, The MIT Press, Cambridge, Massachusetts, Estados Unidos, 1997.
35

35 La síntesis de coherencia puede realizarse utilizando los denominados de-correladores en el dominio de la frecuencia según se describe en el documento de E. Schuijers, W. Oomen, B. den Brinker y J. Breebaart, "Avances en la codificación paramétrica para audio de alta calidad", en Preprint 114th Conv. Aud. Eng. Soc., marzo 2003. Sin embargo, los métodos de síntesis conocidos para la estimación de las pistas espaciales y la sintetización de señales de audio multicanal pueden sufrir una complejidad incrementada, véase documento EP 1565036. Además, el uso de parámetros de ICC, p.e., además de otros parámetros, tales como diferencias de nivel entre canales (ICLDs) y las diferencias de fase entre canales (ICPDs) pueden aumentar una sobrecarga de tasa binaria.
40

45 **SUMARIO DE LA INVENCION**

45 Es el objetivo de la invención dar a conocer un concepto para la estimación de parámetros de codificación que representen relaciones intercanal entre canales de una señal de audio multicanal para una codificación eficiente de las señales de audio.

50 Este objetivo se consigue mediante las características establecidas en las reivindicaciones independientes. Otras formas de realización son evidentes a partir de las reivindicaciones subordinadas, la descripción y las figuras.

Con el fin de poder describir la invención en detalle, se utilizarán los términos, abreviaturas y notaciones siguientes:

55 **BCC:** Codificación de pistas binaurales, codificación de señales estéreo o de multicanal utilizando una mezcla reducida y pistas binaurales (o parámetros espaciales) para describir las relaciones entre canales.

Pistas

60 **binaurales:** Pistas intercanales entre las señales de entrada en el oído izquierdo y en el oído derecho (véase también ITD, ILD, e IC).

CLD: Diferencia de nivel de canal, lo mismo que ICLD.

FFT: Realización rápida de la DFT, indicada como Transformada de Fourier Rápida.

65 **STFT:** Transformada de Fourier de corta duración.

- HRTF: Función de transferencia relacionada con auriculares, transducción de modelización de sonido desde una fuente a las entradas del oído izquierdo y del oído derecho en un campo libre.
- 5 IC: Coherencia interaural, esto es, grado de similitud entre las señales de entrada desde el oído izquierdo y el oído derecho. Este término se refiere también, a veces, como IAC o correlación cruzada interaural (IACC).
- 10 ICC: Coherencia intercanal, correlación intercanal.
- ICPD: Diferencia de fase intercanal. Diferencia de fase media entre un par de señales.
- ICLD: Diferencia de nivel de intercanal.
- 15 ICTD: Diferencia temporal intercanal.
- ILD: Diferencia de nivel interaural, esto es, diferencia de nivel entre las señales de entrada del oído izquierdo y del oído derecho. Se refiere también a veces como diferencia de intensidad interaural (IID)
- 20 IPD: Diferencia de fase interaural, esto es, diferencia de fase entre las señales de entrada del oído izquierdo y del oído derecho.
- ITD: Diferencia temporal interaural, esto es, diferencia temporal entre las señales de entrada del oído izquierdo y del oído derecho. Se refiere también, a veces, como un retardo interaural.
- 25 Mezcla: Dado un número de señales origen (p.e., instrumentos registrados por separado, registro multipista), el proceso de generar señales estereero o de audio multicanal previstas para la reproducción de audio espacial que se indica por mezcla.
- 30 Audio espacial: Señales de audio que, cuando se reproducen a través de un sistema de reproducción adecuado, recuperan operativamente una imagen espacial del auditorio.
- Pistas
- 35 espaciales: Pistas importantes para la percepción espacial. Este término se utiliza para pistas entre pares de canales de una señal de audio multicanal o estéreo (véase también ICTD, ICLD e ICC), también indicadas como parámetros espaciales o pistas binaurales.
- 40 En conformidad con el primer aspecto de la idea inventiva, la invención se refiere a un codificador de audio paramétrico para generar un parámetro de codificación para una señal de canal de audio para 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, siendo el parámetro de codificación un parámetro de coherencia intercanal, comprendiendo el codificador de audio paramétrico un generador de parámetros, estando configurado el generador de parámetros para
- 45 - determinar para la señal de canal de audio de la pluralidad de señales de canales de audio un primer conjunto de parámetros de codificación a partir de los valores de señales de canales de audio de los valores de la señal de audio de referencia y de señal de canal de audio 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, en donde el primer conjunto de parámetros de codificación son parámetros de diferencia de fase intercanal o parámetros de diferencia de fase intercanal de sub-banda,
- 50 - para determinar para la señal de canal de audio una primera media de parámetros de codificación sobre la base del primer conjunto de parámetros de codificación de la señal de canal de audio, refiriéndose la primera media de parámetros de codificación a una trama actual de la señal de canal de audio, en donde el generador de parámetros está configurado para determinar la primera media de parámetros de codificación de la señal de canal de audio como una media del primer conjunto de parámetros de codificación de la señal de canal de audio a través de las bandas de frecuencia o de las sub-bandas de frecuencia,
- 55 - para determinar para la señal de canal de audio, una segunda media de parámetros de codificación sobre la base de la primera media de parámetros de codificación de la señal de canal de audio y al menos una otra primera media de parámetros de codificación de la señal de canal de audio, refiriéndose a la al menos una otra primera de parámetros de codificación a una trama anterior de la señal de canal de audio, y
- 60 - para determinar para la señal de canal de audio, una segunda media de parámetros de codificación sobre la base de la primera media de parámetros de codificación de la señal de canal de audio y al menos una otra primera media de parámetros de codificación de la señal de canal de audio, refiriéndose a la al menos una otra primera de parámetros de codificación a una trama anterior de la señal de canal de audio, y
- 65 - para determinar el parámetro de codificación sobre la base de la primera media de parámetros de codificación de la señal de canal de audio y la segunda media de parámetros de codificación de la señal de canal de audio;

en donde el generador de parámetros está configurado, además,

- 5 - para determinar un valor absoluto de una diferencia entre la segunda media de parámetros de codificación y la primera media de parámetros de codificación, y
- para determinar el parámetro de codificación como una función del valor absoluto determinado.

10 Utilizando las tramas actuales y anteriores de la señal de canal de audio se puede realizar eficientemente la operación de promedio a largo plazo.

15 La señal de audio de referencia puede ser una de las señales de canales de audio de la señal de audio multicanal. En particular, la señal de audio de referencia puede ser una señal de canal de audio del oído izquierdo o del oído derecho de una señal estéreo que constituye una forma de realización de la señal multicanal de dos canales. Sin embargo, la señal de audio de referencia puede ser cualquier señal que constituya una referencia para determinar los parámetros de codificación. Dicha señal de referencia puede estar constituida por una señal de audio mezclada reducida monoaural después de la mezcla reducida de los canales de la señal de audio multicanal o uno de entre el canal de una señal de audio mezclada reducida después de la mezcla reducida de los canales de la señal de audio multicanal.

20 Dichos parámetros como diferencia de fase intercanal o diferencia de fase intercanal de sub-banda representan un grado de similitud entre las señales de audio y por ello, se pueden utilizar por el codificador para reducir la información a transmitirse y de este modo, reducir la complejidad del cálculo.

25 Mediante esa promediación del primer conjunto de parámetros de codificación de la señal de canal de audio a través de las bandas de frecuencias o de las sub-bandas de frecuencia, el codificador de audio paramétrico proporciona una media de corta duración de la señal de audio en donde se consideran todas las componentes de la frecuencia.

30 Mediante esa diferencia entre la segunda media de parámetros de codificación y la primera media de parámetros de codificación, el codificador de audio paramétrico proporciona una medida para la diferencia entre la media a largo plazo y la media de corta duración y por lo tanto, es capaz de predecir el comportamiento de la voz o de la música.

35 Cuando el parámetro de codificación se proporciona como una función del valor absoluto determinado, existe una relación entre el parámetro de codificación y el valor absoluto determinado, que se puede utilizar para calcular eficientemente el parámetro de codificación. La complejidad de cálculo se reduce de este modo.

40 El codificador de audio paramétrico puede tener una baja complejidad puesto que no requiere una coherencia o cálculo de correlación. Incluso proporciona una estimación exacta de la relación entre los canales de audio cuando el valor de ICC se cuantifica con un cuantificador aproximado que requiere solamente unas pocas etapas. En particular, para las señales musicales, pero también para las señales de voz, la utilización del parámetro de codificación para la codificación de las señales de audio es importante porque la música de salida suena más natural con la anchura de escena acústica correcta y no “en seco”. Para un sistema de codificación de audio estéreo paramétrico de baja tasa binaria, el presupuesto de bits está limitado y solamente se transmite una ICC de banda completa, siendo el parámetro de codificación capaz de representar la correlación global entre los canales.

45 En una primera posible forma de puesta en práctica del codificador de audio paramétrico en conformidad con el primer aspecto de la idea inventiva, el generador de parámetros está configurado para determinar las diferencias de fase de los valores de señales de canales de audio siguientes para obtener el primer conjunto de parámetros de codificación.

50 Las diferencias de fase de las señales de canales de audio siguientes se requieren para reproducir las diferencias de fase y/o retardo entre los canales. Cuando se reproducen las diferencias de fase, la voz y la música suenan de forma más natural.

55 En una segunda posible forma de puesta en práctica del codificador de audio paramétrico en conformidad con el primer aspecto de la idea inventiva o en conformidad con la forma de puesta en práctica precedente del primer aspecto de la idea inventiva, la señal de canal de audio y la señal de audio de referencia son señales de dominio de la frecuencia y los valores de las señales de canales de audio y los valores de la señal de audio de referencia están asociados con la bandas de frecuencia o las sub-bandas de frecuencia.

60 La resolución de frecuencia utilizada está ampliamente motivada por la resolución de frecuencia del sistema del auditorio. La psicoacústica sugiere que la percepción espacial está muy probablemente basada en una representación de banda crítica de la señal de entrada acústica. Esta resolución de frecuencia se considera utilizando un banco de filtros invertible con sub-bandas con anchos de banda iguales o proporcionales al ancho de banda crítico del sistema del auditorio. De este modo, el codificador de audio paramétrico puede estar bien adaptado a la percepción humana.

5 En una tercera posible forma de puesta en práctica del codificador de audio paramétrico en conformidad con el primer aspecto de la idea inventiva o en conformidad con cualquiera de las formas de puesta en práctica precedentes del primer aspecto, el codificador de audio paramétrico comprende, además, un transformador para transformar una pluralidad de señales de canales de audio en el dominio temporal en el dominio de la frecuencia para obtener la pluralidad de señales de canales de audio.

10 La ecualización de la respuesta a impulsos de canales puede realizarse eficientemente en el dominio de la frecuencia puesto que la convolución en el dominio temporal es una multiplicación en el dominio de la frecuencia. De este modo, la realización de los cálculos del codificador de audio paramétrico en el dominio de la frecuencia puede dar lugar a una más alta eficiencia con respecto a la complejidad de cálculo o en una mayor exactitud.

15 En una cuarta posible forma de realización del codificador de audio paramétrico, en conformidad con el primer aspecto de la idea inventiva o en conformidad con cualquiera de las formas de realización precedentes del primer aspecto, el generador de parámetros está configurado para determinar el primer conjunto de parámetros de codificación para cada contenedor de frecuencia o para cada sub-banda de frecuencia de las señales de canales de audio.

20 El codificador de audio paramétrico puede limitar la determinación del primer conjunto de parámetros de codificación para bandas de frecuencia o sub-bandas de frecuencias que sean percibibles por el oído humano y de este modo, se reduce la complejidad.

25 En una quinta posible forma de realización del codificador de audio paramétrico en conformidad con el primer aspecto de la idea inventiva o en conformidad con cualquiera de las formas de realización precedentes del primer aspecto, el generador de parámetros está configurado para determinar la segunda media de parámetros de codificación de la señal de canal de audio como una media de una pluralidad de las primeras medias de parámetros de codificación a través de una pluralidad de tramas de la señal de canal de audio, en donde cada primera media de parámetros de codificación está asociada a una trama y de la señal de audio multicanal.

30 Mediante esa promediación, el codificador de audio paramétrico proporciona una media a largo plazo de la señal de audio en donde se consideran las propiedades características de la señal de voz o de la señal musical.

35 En una sexta posible forma de realización del codificador de audio paramétrico en conformidad con el primer aspecto de la idea inventiva o en conformidad con cualquiera de las formas de realización precedentes del primer aspecto, el generador de parámetros está configurado para determinar el parámetro de codificación a partir de una diferencia entre un primer valor paramétrico y el valor absoluto determinado multiplicado por un segundo valor paramétrico.

40 Cuando el parámetro de codificación se proporciona como una diferencia entre el primer valor paramétrico y el valor absoluto determinado, existe una relación entre el parámetro de codificación y el valor absoluto determinado, que puede utilizarse para calcular eficientemente el parámetro de codificación. De este modo, se reduce la complejidad del cálculo.

45 En una séptima posible forma de realización del codificador de audio paramétrico en conformidad con el primer aspecto de la idea inventiva o en conformidad con cualquiera de las formas de realización precedentes del primer aspecto, el generador de parámetros está configurado para establecer el primer valor paramétrico a uno y para establecer el segundo valor paramétrico a uno.

50 Mediante esa relación, el codificador de audio paramétrico es capaz de calcular eficientemente el parámetro de codificación. De este modo, se reduce la complejidad del cálculo.

55 En una octava forma de realización posible del codificador de audio paramétrico en conformidad con el primer aspecto de la idea inventiva o en conformidad con cualquiera de las formas de realización precedentes del primer aspecto, el codificador de audio paramétrico comprende, además, un generador de señales mezcladas reducidas para la superposición de al menos dos de las señales de canales de audio de la señal de audio multicanal para obtener una señal mezclada reducida, un codificador de audio, en particular, un codificador monoaural, para codificar la señal mezclada reducida para obtener una señal de audio codificada y un combinador para combinar la señal de audio codificada con un parámetro de codificación correspondiente.

60 La señal mezclada reducida y la señal de audio codificada pueden utilizarse como una señal de referencia para el generador de parámetros. Ambas señales incluyen la pluralidad de señales de canales de audio y de este modo, proporciona una mayor exactitud que una señal de canal única tomada como señal de referencia.

65 En una novena forma de realización posible del codificador de audio paramétrico en conformidad con el primer aspecto de la idea inventiva, la trama actual de la señal de canal de audio es contigua a la trama anterior o la señal de canal de audio.

5 Cuando ambas tramas son contiguas, se detectan los denominados picos de tensión en las señales de canales de audio en la media y pueden considerarse en el codificador de audio paramétrico. De este modo, la codificación es más precisa que una codificación en donde no pueden detectarse dichos picos de tensión.

10 En conformidad con un segundo aspecto de la idea inventiva, la invención se refiere a un codificador de audio paramétrico para generar 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, siendo el parámetro de codificación un parámetro de coherencia intercanal, comprendiendo el codificador de audio paramétrico un generador de parámetros, estando el generador de parámetros configurado

15 - para determinar para la señal de canal de audio de la pluralidad de señales de canales de audio un primer conjunto de parámetros de codificación a partir de los valores de señales de canales de audio de la señal de canal de audio y los 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 mezclada reducida derivada de al menos dos señales de canales de audio de entre la pluralidad de señales de audio multicanal, en donde el primer conjunto de parámetros de codificación son parámetros de diferencia de fase intercanal o parámetros de diferencia de fase entre canales de sub-banda,

20 - para determinar para la señal de canal de audio una primera media de parámetros de codificación basada en el primer conjunto de parámetros de codificación de la señal de canal de audio, refiriéndose la primera media de parámetros de codificación a una trama actual de la señal de canal de audio, en donde el generador de parámetros está configurado para determinar la primera media de parámetros de codificación de la señal de canal de audio como una media del primer conjunto de parámetros de codificación de la señal de canal de audio a través de las bandas de frecuencia o de las sub-bandas de frecuencia,

25 - para determinar para la señal de canal de audio una segunda media de parámetros de codificación basada en la primera media de parámetros de codificación de la señal de canal de audio y al menos una otra primera media de parámetros de codificación de la señal de canal de audio, refiriéndose la a menos otra primera media de parámetros de codificación a una trama anterior de la señal de canal de audio y

30 - para determinar el parámetro de codificación sobre la base de la primera media de parámetros de codificación de la señal de canal de audio y la segunda media de parámetros de codificación de la señal de canal de audio;

35 en donde el generador de parámetros está configurado además

40 - para determinar un valor absoluto de una diferencia entre la segunda media de parámetros de codificación y la primera media de parámetros de codificación, y

45 - para determinar el parámetro de codificación como una función del valor absoluto determinado.

Utilizando las tramas actuales y anteriores de la señal de canal de audio, se puede realizar eficientemente la operación de promediación a largo plazo.

50 La señal de audio de referencia puede ser una de las señales de canales de audio de la señal de audio multicanal. En particular, la señal de audio de referencia puede ser una señal de canal del oído izquierdo o del oído derecho de una señal estéreo que constituye una forma de realización de una señal multicanal de dos canales. Sin embargo, la señal de audio de referencia puede ser cualquier señal que constituya una referencia para determinar los parámetros de codificación. Dicha señal de referencia puede formarse mediante una señal de audio mezclada reducida después de la mezcla reducida de los canales de la señal de audio multicanal o una salida de un codificador monoaural.

55 Parámetros tales como la diferencia de fase intercanal o la diferencia de fase entre canales de sub-banda representan un grado de similitud entre las señales de audio y por ello, se pueden utilizar por el codificador para reducir la información a transmitirse y de este modo, reducir la complejidad del cálculo.

60 Mediante esa promediación del primer conjunto de parámetros de codificación de la señal de canal de audio a través de bandas de frecuencia o de sub-bandas de frecuencia, el codificador de audio paramétrico proporciona una media a corto plazo de la señal de audio en donde se consideran todas las componentes de la frecuencia.

65 Mediante esa diferencia entre la segunda media de parámetros de codificación y la primera media de parámetros de codificación, el identificador de audio paramétrico proporciona una medida de la diferencia entre la media a largo plazo y la media a corto plazo y por lo tanto, es capaz de predecir el comportamiento de la voz o de la música.

Cuando el parámetro de codificación se proporciona como una función del valor absoluto determinado, existe una

relación entre el parámetro de codificación y el valor absoluto determinado, que se puede utilizar para calcular eficientemente el parámetro de codificación. De este modo, se reduce la complejidad del cálculo.

5 El codificador de audio paramétrico puede tener una baja complejidad puesto que no requiere un cálculo de correlación o coherencia. Incluso proporciona una estimación exacta de la relación entre los canales de audio cuando se cuantifica el valor de ICC con un cuantificador aproximado que requiere solamente unas pocas etapas. En particular, para las señales musicales, pero también para las señales de voz, la utilización del parámetro de codificación para la codificación de las señales de audio es importante porque la música de salida suena más natural con la anchura de escena acústica correcta y no “en seco”. Para un sistema de codificación de audio estéreo
10 paramétrico de muy baja tasa binaria, el presupuesto de bits está limitado y solamente se transmite una ICC de banda completa, siendo el parámetro de codificación capaz de representar la correlación global entre los canales.

15 En una primera posible forma de realización del codificador de audio paramétrico en conformidad con el segundo aspecto de la idea inventiva, el generador de parámetros está configurado para determinar diferencias de fase de valores de señales de canales de audio siguientes con el fin de obtener el primer conjunto de parámetros de codificación. Las diferencias de fase de las señales de canales de audio siguientes se requieren para reproducir las diferencias de fase y/o retardo entre los canales. Cuando se reproducen las diferencias de fase, la voz y la música tienen un sonido más natural.

20 En una segunda posible forma de realización del codificador de audio paramétrico en conformidad con el segundo aspecto de la idea inventiva o en conformidad con la forma de realización precedente del segundo aspecto, la señal de canal de audio y la señal de audio de referencia son señales de dominio de la frecuencia y los valores de las señales de canales de audio y los valores de las señales de audio de referencia están asociados con las bandas de frecuencia o las sub-bandas de frecuencia.

25 La resolución de frecuencia utilizada está ampliamente motivada por la resolución de frecuencia del sistema del auditorio. La psicoacústica sugiere que la percepción espacial está muy probablemente basada en una representación de banda crítica de la señal de entrada acústica. Esta resolución de frecuencia se considera utilizando un banco de filtros invertible con sub-bandas y anchos de banda iguales o proporcionales al ancho de banda crítico del sistema del auditorio. De este modo, el codificador de audio paramétrico puede adaptarse
30 adecuadamente para la percepción humana.

35 En una tercera posible forma de realización del codificador de audio paramétrico en conformidad con el segundo aspecto de la idea inventiva o en conformidad con cualquiera de las formas de realización precedentes del segundo aspecto, el codificador de audio paramétrico comprende, además, un transformador para transformar una pluralidad de señales de canales de audio del dominio temporal en el dominio de la frecuencia con el fin de obtener la pluralidad de señales de canales de audio.

40 La ecualización de la respuesta de impulso de canal puede realizarse eficientemente en el dominio de la frecuencia puesto que la convolución en el dominio temporal es una multiplicación en el dominio de la frecuencia. De este modo, la realización de los cálculos del codificador de audio paramétrico en el dominio de la frecuencia puede dar lugar a una más alta eficiencia con respecto a la complejidad del cálculo o en una mayor exactitud.

45 En una cuarta posible forma de realización del codificador de audio paramétrico en conformidad con el segundo aspecto de la idea inventiva o en conformidad con cualquiera de las formas de realización precedentes del segundo aspecto, el generador de parámetros está configurado para determinar el primer conjunto de parámetros de codificación para cada contenedor de frecuencia o para cada sub-bandas de frecuencia de las señales de canales de audio.

50 El codificador de audio paramétrico puede limitar la detección del primer conjunto de parámetros de codificación a bandas de frecuencia o sub-bandas de frecuencia que sean percibibles por el oído humano y de este modo, se reduce la complejidad.

55 En una quinta posible forma de realización del codificador de audio paramétrico en conformidad con el segundo aspecto de la idea inventiva o en conformidad con cualquiera de las formas de realización precedentes del segundo aspecto, el generador de parámetros está configurado para determinar la segunda media de parámetros de codificación de la señal de canal de audio como una media de una pluralidad de primeras medias de parámetros de codificación durante una pluralidad de tramas de la señal de canal de audio, en donde cada primera media de parámetros de codificación está asociada con una trama de la señal de audio multicanal.

60 Mediante esa promediación, el codificador de audio paramétrico proporciona una media a largo plazo de la señal de audio en donde se consideran las propiedades características de la señal de voz o de la señal de música.

65 En una sexta posible forma de realización del codificador de audio paramétrico en conformidad con el segundo aspecto de la idea inventiva o en conformidad con cualquiera de las formas de realización precedentes del segundo aspecto, el generador de parámetros está configurado para determinar el parámetro de codificación a partir de una

diferencia entre un primer valor paramétrico y el valor absoluto determinado multiplicado por un segundo valor paramétrico.

5 Cuando el parámetro de codificación se proporciona como una diferencia entre el primer valor paramétrico y el valor absoluto determinado, existe una relación entre el parámetro de codificación y el valor absoluto determinado, que puede utilizarse para calcular eficientemente el parámetro de codificación. De este modo, se reduce la complejidad del cálculo.

10 En una séptima posible forma de realización del codificador de audio paramétrico en conformidad con el segundo aspecto de la idea inventiva o en conformidad con cualquiera de las formas de realización precedentes del segundo aspecto, el generador de parámetros está configurado para establecer el primer valor paramétrico a uno y para establecer el segundo valor paramétrico a uno.

15 Mediante esa relación, el codificador de audio paramétrico es capaz de calcular eficientemente el parámetro de codificación. De este modo, se reduce la complejidad del cálculo.

20 En una octava posible forma de realización del codificador de audio paramétrico en conformidad con el segundo aspecto de la idea inventiva o en conformidad con cualquiera de las formas de realización precedentes del segundo aspecto, el codificador de audio paramétrico comprende, además, un generador de señales mezcladas reducidas para la superposición de al menos dos de las señales de canales de audio de la señal de audio multicanal para obtener una señal mezclada reducida, un codificador de audio, en particular un codificador monoaural, para codificar la señal mezclada reducida con el fin de obtener una señal de audio codificada y un combinador para combinar la señal de audio codificada con un parámetro de codificación correspondiente.

25 La señal mezclada reducida y la señal de audio codificada pueden utilizarse como una señal de referencia para el generador de parámetros. Ambas señales incluyen la pluralidad de señales de canales de audio y de este modo, proporcionan una exactitud mayor que una señal de canal único tomada como una señal de referencia.

30 En una novena forma de realización del codificador de audio paramétrico en conformidad con el segundo aspecto de la idea inventiva, la trama actual de la señal de canal de audio es contigua a la trama anterior de la señal de canal de audio.

35 Cuando ambas tramas son contiguas, se detectan picos de tensión en las señales de canales de audio en la promediación y pueden considerarse en el codificador de audio paramétrico. De este modo, la codificación es más precisa que una codificación en donde no puedan detectarse picos de tensión.

40 En conformidad con un tercer aspecto de la idea inventiva, la invención se refiere a un método para generar 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, siendo el parámetro de codificación un parámetro de coherencia intercanal, comprendiendo el método:

45 - determinar para la señal de canal de audio de la pluralidad de señales de canales de audio un primer conjunto de parámetros de codificación a partir de los valores de señales de canales de audio de la señal de canal de audio y de los valores de señales 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, en donde el primer conjunto de parámetros de codificación son parámetros de diferencia de fase intercanal o parámetros de diferencia de fase intercanal de sub-banda,

50 - determinar para la señal de canal de audio una primera media de parámetros de codificación sobre la base del primer conjunto de parámetros de codificación de la señal de canal de audio, refiriéndose la primera media de parámetros de codificación a una trama actual de la señal de canal de audio, en donde la primera media de parámetros de codificación se determina como una media del primer conjunto de parámetros de codificación de la señal de canal de audio a través de bandas de frecuencia o de sub-bandas de frecuencia,

55 - determinar para la señal de canal de audio una segunda media de parámetros de codificación sobre la base de la primera media de parámetros de codificación de la señal de canal de audio y al menos otra primera media de parámetros de codificación de la señal de canal de audio, refiriéndose al menos a otra primera media de parámetros de codificación a una trama anterior de la señal de canal de audio, y

60 - determinar el parámetro de codificación sobre la base de la primera media de parámetros de codificación de la señal de canal de audio y la segunda media de parámetros de codificación de la señal de canal de audio;

65 en donde la determinación del parámetro de codificación sobre la base de la primera media de parámetros de codificación de la señal de canal de audio y la segunda media de parámetros de codificación de la señal de canal de audio comprende:

- determinar un valor absoluto de una diferencia entre la segunda media de parámetros de codificación y la primera media de parámetros de codificación, y
- determinar el parámetro de codificación como una función del valor absoluto determinado.

5

El método puede realizarse eficientemente en un procesador.

La señal de audio de referencia puede ser una de las señales de canales de audio de la señal de audio multicanal. En particular, la señal de audio de referencia puede ser una señal de canal de audio izquierdo o derecho de una señal estéreo que constituye una forma de realización de una señal multicanal de dos canales. Sin embargo, la señal de audio de radiofrecuencia puede ser cualquier señal que constituya una referencia para determinar los parámetros de codificación. Dicha señal de referencia puede estar constituida por una señal de audio mezclada reducida monoaural después de la mezcla reducida de los canales de la señal de audio multicanal o uno del canal de una señal de audio mezclada reducida después de la mezcla reducida de los canales de la señal de audio multicanal.

10

15

En conformidad con un cuarto aspecto de la idea inventiva, la invención se refiere a un método para generar 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, siendo el parámetro de codificación un parámetro de coherencia intercanal, comprendiendo el método:

20

- determinar para la señal de canal de audio de la pluralidad de señales de canales de audio un primer conjunto de parámetros de codificación a partir de los valores de señales de canales 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 mezclada reducida derivada de al menos dos señales de canales de audio de entre la pluralidad de señales de audio multicanal, en donde el primer conjunto de parámetros de codificación son parámetros de diferencia de fase intercanal o parámetros de diferencia de fase intercanal de sub-banda,

25

- determinar para la señal de canal de audio una primera media de parámetros de codificación sobre la base del primer conjunto de parámetros de codificación de la señal de canal de audio, refiriéndose la primera media de parámetros de codificación a una trama actual de la señal de canal de audio, en donde la primera media de parámetros de codificación se determina como media del primer conjunto de parámetros de codificación de la señal de canal de audio a través de bandas de frecuencia o de sub-bandas de frecuencia,

30

35

- determinar para la señal de canal de audio una segunda media de parámetros de codificación sobre la base de la primera media de parámetros de codificación de la señal de canal de audio y al menos una otra primera media de parámetro de codificación de la señal de canal de audio, refiriéndose la a menos una otra primera media de parámetros de codificación a una trama anterior de la señal de canal de audio y

40

- determinar el parámetro de codificación sobre la base de la primera media de parámetros de codificación de la señal de canal de audio y la segunda media de parámetros de codificación de la señal de canal de audio;

en donde la determinación del parámetro de codificación sobre la base de la primera media de parámetros de codificación de la señal de canal de audio y la segunda media de parámetros de codificación de la señal de canal de audio comprende:

45

- determinar un valor absoluto de una diferencia entre la segunda media de parámetros de codificación y la primera media de parámetros de codificación, y
- determinar el parámetro de codificación como una función del valor absoluto determinado.

50

El método puede realizarse de forma eficiente en un procesador.

La señal de audio de referencia puede ser una de entre las señales de canales de audio de la señal de audio multicanal. En particular, la señal de audio de referencia puede ser una señal de canal de audio izquierdo o derecho de una señal estéreo que constituye una forma de realización de una señal multicanal de dos canales. Sin embargo, la señal de audio de referencia puede ser cualquier señal que constituya una referencia para determinar los parámetros de codificación. Dicha señal de referencia puede estar constituida por una señal de audio mezclada reducida monoaural después de la mezcla reducida de los canales de la señal de audio multicanal o uno de los canales de una señal de audio mezclada reducida después de la mezcla reducida de los canales de la señal de audio multicanal.

55

60

En conformidad con un quinto aspecto de la idea inventiva, la invención se refiere a un programa informático que está configurado para poner en práctica el método en conformidad con uno del tercero y cuarto aspectos de la idea inventiva cuando se ejecuta en un ordenador.

65

El programa informático tiene una complejidad reducida y de este modo, se puede poner en práctica eficientemente en un terminal móvil en donde debe economizarse la vida de la batería. La duración de la vida útil de la batería se aumenta cuando el programa informático se ejecuta en un terminal móvil.

5 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 secundario o como circuito de hardware dentro de un circuito integrado específico de la aplicación (ASIC).

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

BREVE DESCRIPCIÓN DE LOS DIBUJOS

15 Formas de realización adicionales de la invención se describirán, a continuación, con respecto a las Figuras siguientes, en las que:

La Figura 1 ilustra un diagrama de bloques de un codificador de audio paramétrico en conformidad con una forma de realización;

20 La Figura 2 ilustra un diagrama de bloques de un decodificador de audio paramétrico en conformidad con una forma de realización;

La Figura 3 ilustra un diagrama de bloques de un codificador de audio estéreo paramétrico y un decodificador en conformidad con una forma de realización; y

25 La Figura 4 ilustra un diagrama esquemático de un método para generar un parámetro de codificación para una señal de canal de audio en conformidad con una forma de realización.

DESCRIPCIÓN DETALLADA DE LAS FORMAS DE REALIZACIÓN DE LA INVENCION

30 La Figura 1 ilustra un diagrama de bloques de un codificador de audio paramétrico 100 en conformidad con una forma de realización. El codificador de audio paramétrico 100 recibe una señal de audio multicanal 101 como señal de entrada y proporciona un flujo de bits como señal de salida 103. El identificador de audio paramétrico 100 comprende un generador de parámetros 105 acoplado a la señal de audio multicanal 101 para generar un parámetro de codificación 115, un generador de señales mezcladas reducidas 107 acoplado a la señal de audio multicanal 101 para generar una señal mezclada reducida 111 o una señal suma, un codificador de audio 109 acoplado a la generador de señales mezcladas reducidas 107 para codificar la señal mezclada reducida 111 para proporcionar una señal de audio codificada 113 y un combinador 117, p.e., un dispositivo formador de flujos de bits acoplado al generador de parámetros 105 y al codificador de audio 109 para formar un flujo de bits 103 a partir del parámetro de codificación 115 y la señal codificada 113.

45 El codificador de audio paramétrico 100 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.e., el canal de audio de mezcla reducida junto con parámetros adicionales que describen “diferencias perceptualmente pertinentes” entre los canales de audio $X_1[b]$, $X_2[b]$, ..., $X_M[b]$. El sistema de codificación está en conformidad con la codificación de pistas binaurales (BCC) porque las pistas binaurales desempeñan una función importante a este respecto. Según se indica en la Figura, la pluralidad M de canales de audio de entrada $X_1[b]$, $X_2[b]$, ..., $X_M[b]$ de la señal de audio multicanal 101 son objeto de mezcla reducida a un canal de audio único 111, también indicado como la señal suma. Para una señal de audio estéreo M igual a 2, habida cuenta de las “diferencias perceptualmente pertinentes” entre los canales de audio $X_1[b]$, $X_2[b]$, ..., $X_M[b]$, el parámetro de codificación 115, p.e., una diferencia de tiempo intercanal (ICTD), una diferencia de nivel intercanal (ICLD) y/o una coherencia intercanal (ICC) se estima como una función de la frecuencia y del tiempo y se transmite como información secundaria al decodificador 200 descrito en la Figura 2.

55 El generador de parámetros 105 que pone en práctica BCC procesa la señal de audio multicanal 101 con un determinado periodo de tiempo y resolución de frecuencia. La resolución de frecuencia utiliza es ampliamente motivada por la resolución de frecuencia del sistema del auditorio. La psicoacústica sugiere que la percepción espacial está muy probablemente basada en una representación de banda crítica de la señal de entrada acústica. Esta resolución de frecuencia se considera utilizando un banco de filtros invertible con sub-bandas con anchos de banda iguales o proporcionales al ancho de banda crítico del sistema del auditorio. Es importante que la suma transmitida 111 contenga todas las componentes de señales de la señal de audio multicanal 101. El objetivo es que cada componente de señal se mantenga completamente. La suma simple de los canales de entrada de audio $X_1[b]$, $X_2[b]$, ..., $X_M[b]$ de la señal de audio multicanal 101 suele dar lugar a la amplificación o atenuación de componentes de señales. Dicho de otro modo, la potencia de las componentes de señales en la suma “simple” suele ser mayor o menor que la suma de la potencia de las componentes de señal correspondientes de cada canal $X_1[b]$, $X_2[b]$, ..., $X_M[b]$. Por lo tanto, una técnica de mezcla reducida se utiliza aplicando el dispositivo de mezcla reducida 107 que equaliza la señal suma 111 de modo que la potencia de las componentes de señales en la señal suma 111 sea

aproximadamente la misma que la potencia correspondiente en todos los canales de audio de entrada $X_1[b]$, $X_2[b]$, ..., $X_M[b]$ de la señal de audio multicanal 101. Los canales de audio de entrada $X_1[b]$, $X_2[b]$, ..., $X_M[b]$ representan las señales de canales para la sub-banda b . El canal de audio de entrada del dominio de la frecuencia se indica por $X_1[k]$, $X_2[k]$, ..., $X_M[k]$, en donde k representa el índice de frecuencia (contenedor de frecuencia), una sub-banda b estando normalmente constituida por varias bandas de frecuencia k .

Dada la señal suma 111, el generador de parámetros 105 sintetiza una señal de audio multicanal o estéreo 115 de modo que ICTD, ICLD y/o ICC se aproximen a las pistas correspondientes de la señal de audio multicanal original 101.

Cuando se consideran respuestas de impulsos espaciales binaurales (BRIRs) de un solo origen, existe una relación entre la anchura del evento operativo del auditorio y la envolvente de escuchas y el valor de IC estimado para las partes iniciales y finales de las respuesta BRIRs. Sin embargo, la relación entre IC (o ICC) y estas propiedades para señales generales (y no solamente las BRIRs) no es simple. Las señales de audio multicanal y de estéreo suelen contener una mezcla compleja de señales origen simultáneamente activas superpuestas por componentes de señales reflejadas que se derivan del registro en espacios cerrados o que se añaden por el técnico de registro para crear artificialmente una impresión espacial. Señales origen diferentes y sus reflexiones ocupan diferentes zonas en el plano de tiempo-frecuencia. Lo que antecede se refleja por 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 del auditorio y la impresión espacial no es obvia. La estrategia del generador de parámetros 105 es sintetizar, a ciegas, estas pistas de modo que se aproximan a las pistas correspondientes de la señal de audio original.

En una forma de realización, el codificador de audio paramétrico 100 utiliza bancos de filtros con sub-bandas de anchos de banda iguales a dos veces el ancho de banda rectangular equivalente. Una escucha informal reveló que la calidad de audio de BCC no mejoró notablemente cuando se eligió una más alta resolución de frecuencia. Una más baja resolución de frecuencia es favorable puesto que da lugar a menos valores 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 temporal, ICTD, ICLD e ICC se consideran a intervalos de tiempo periódicos. En una forma de realización ICTD, ICLD e ICC se consideran aproximadamente cada 4 – 16 ms. Conviene señalar que a no ser que las pistas se consideren a intervalos muy cortos, el efecto de prioridad no está directamente considerado.

La frecuentemente diferencia perceptualmente pequeña conseguida entre la señal de referencia y la señal sintetizada implica que las pistas relacionadas con una amplia gama de atributos de imágenes espaciales del auditorio se consideran implícitamente sintetizando ICTD, ICLD e ICC a intervalos de tiempo periódicos. La tasa binaria requerida para la transmisión de estas pistas espaciales es solamente de unos pocos kb/s y de este modo, el codificador de audio paramétrico 100 es capaz de transmitir señales de audio multicanal y estéreo a tasas binarias próximas a las que se requieren para un canal de audio único. La Figura 4 ilustra un método en el que se estima ICC como el parámetro de codificación 115.

El codificador de audio paramétrico 100 comprende un generador de señal mezclada reducida 107 para la superposición de al menos dos de entre las señales de canales de audio de la señal de audio multicanal 101 para obtener la señal mezclada reducida 111, el codificador de audio 109, en particular un codificador monoaural, para codificar la señal mezclada reducida 111 con el fin de obtener la señal de audio codificada 113 y el combinador 117 para combinar la señal de audio codificada 113 con un parámetro de codificación correspondiente 115.

El codificador de audio paramétrico 100 genera el parámetro de codificación 115 para una señal de canal de audio de entre la pluralidad de señales de canales de audio indicadas como $X_1[b]$, $X_2[b]$, ..., $X_M[b]$ de la señal de audio multicanal 101. Cada una de las señales de canales de audio $X_1[b]$, $X_2[b]$, ..., $X_M[b]$ puede ser una señal digital que comprende valores de señales de canales de audio digitales en el dominio de la frecuencia que se indica como $X_1[k]$, $X_2[k]$, ..., $X_M[k]$.

Una señal de canal de audio, a modo de ejemplo, para la que el codificador de audio paramétrico 100 genera el parámetro de codificación 115 es la primera señal de canal de audio $X_1[b]$ con valores de señales $X_1[k]$. El generador de parámetros 105 determina para la señal de canal de audio $X_1[b]$ un primer conjunto de parámetros de codificación indicado como IPD[b] a partir de los valores de señales de canales de audio $X_1[k]$ de la señal de canal de audio $X_1[b]$ y a partir de los valores de señales de audio de referencia de una señal de audio de referencia.

Una señal de canal de audio que se utiliza como una señal de audio de referencia es la segunda señal de canal de audio $X_2[b]$, a modo de ejemplo. De forma similar, cualquier otra de entre las señales de canales de audio $X_1[b]$, $X_2[b]$, ..., $X_M[b]$ puede servir como una 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 canales de audio que no es igual a la señal de canal de audio $X_1[b]$ para la que se genera el parámetro de codificación 115.

En conformidad con un segundo aspecto de la idea inventiva, la señal de audio de referencia es una señal de audio mezclada reducida derivada a partir de al menos dos señales de canales de audio de entre la pluralidad de señales

de audio multicanal 101, p.e., derivada de la primera señal de canal de audio $X_1[b]$ y de la segunda señal de canal de audio $X_2[b]$. En una forma de realización, la señal de audio de referencia es la señal mezclada reducida 111, también denominada señal suma generada por el dispositivo de mezcla reducida 107. En otra forma de realización, la señal de audio de referencia es la señal codificada 113 proporcionada por el codificador 109.

5 Una señal de audio de referencia, a modo de ejemplo, utilizada por el generador de parámetros 105 es la segunda señal de canal de audio $X_2[b]$ con valores de señales $X_2[k]$.

10 El generador de parámetros 105 determina para la señal de canal de audio $X_1[b]$ una primera media de parámetros de codificación, indicada como $IPD_{mean}[i]$ sobre la base del primer conjunto de parámetros de codificación $IPD[b]$ de la señal de canal de audio $X_1[b]$.

15 El generador de parámetros 105 determina para la señal de canal de audio $X_1[b]$ una segunda media de parámetros de codificación, indicada por $IPD_{mean_long_term}$, sobre la base de la primera media de parámetros de codificación $IPD_{mean}[i]$ de la señal de canal de audio $X_1[b]$ y al menos una otra primera media de parámetros de codificación indicada como $IPD_{mean}[i-1]$ de la señal de canal de audio $X_1[b]$. En una forma de realización, la primera media de parámetros de codificación $IPD_{mean}[i]$ se refiere a una trama actual i de la señal de canal de audio $X_1[b]$ y la otra primera media de parámetros de codificación $IPD_{mean}[i-1]$ se refiere a una trama anterior $i-1$ de la señal de canal de audio $X_1[b]$. En una forma de realización, la trama anterior $i-1$ de la señal de canal de audio $X_1[b]$ es la trama $i-1$ recibida antes de la trama actual i sin ninguna otra trama intermedia. En una forma de realización, la trama anterior $i-N$ de la señal de canal de audio $X_1[b]$ es una trama $i-N$ recibida antes de la trama actual i pero múltiples tramas llegaron en el periodo intermedio.

25 El generador de parámetros 105 determinar el parámetro de codificación 115, indicado como ICC, sobre la base de la primera media de parámetros de codificación $IPD_{mean}[i]$ de la señal de canal de audio $X_1[b]$ y sobre la base de la segunda media de parámetros de codificación $IPD_{mean_long_term}$ de la señal de canal de audio $X_1[b]$.

30 El primer conjunto de parámetros de codificación $IPD[b]$ son diferencias de fase entre canales, diferencias de nivel entre canales, coherencias entre canales, diferencias de intensidad entre canales, diferencias de niveles entre canales de sub-banda, diferencias de fase entre canales de sub-banda, coherencias entrecanales de sub-banda, diferencias de intensidad entre canales de sub-banda o una de sus combinaciones. Una diferencia de fase entre canales (ICPD) es una diferencia de fase media entre un par de señales. Una diferencia de nivel intercanal (ICLD) es la misma que una diferencia de nivel interaural (ILD), esto es, una diferencia de nivel entre señales de entrada del oído izquierdo y del oído derecho, pero definidas más en general entre cualquier par de señales, p.e., un par de señales de altavoces, un par de señales de entrada al oído, etc. Una coherencia intercanal o una correlación intercanal es la misma que una coherencia interaural (IC), esto es, el grado de similitud entre las señales de entrada el oído izquierdo y del oído derecho, pero definidas más en general entre cualquier par de señales, p.e., par de señales de altavoces, par de señales de entrada a los oídos, etc. Una diferencia temporal intercanal (ICTD) es la misma que una diferencia temporal interaural (ITD), a veces también referida como un retardo interaural, esto es, una diferencia temporal entre las señales de entrada de los oídos izquierdo y derecho pero definida más en general entre cualquier par de señales, p.e., pares de señales de altavoces, par de señales de entrada a los oídos, etc. Las diferencias de niveles entre canales de sub-banda, las diferencias de fase entre canales de sub-bandas, 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 101 determina las diferencias de fase de valores de señal de canal de audio $X_1[k]$ siguientes para obtener el primer conjunto de parámetros de codificación $IPD[b]$. En una forma de realización, la señal de canal de audio $X_1[b]$ y la señal de audio de referencia $X_2[b]$ son señales de dominio de la frecuencia y los valores de señales de canales de audio $X_1[k]$ y los valores de señales de audio de referencia $X_2[k]$ están asociados con bandas de frecuencia indicadas como $[k]$ o sub-bandas de frecuencia, indicadas como $[b]$. En una forma de realización, el codificador de audio paramétrico 100 comprende un transformador, p.e., un dispositivo de FFT para transformar una pluralidad de señales de canales de audio del dominio temporal $X_1[n]$, $X_2[n]$ en el dominio de la frecuencia con el fin de obtener la pluralidad de señales de canales de audio $X_1[b]$, $X_2[b]$. En una forma de realización, el generador de parámetros 101 determinar el primer conjunto de parámetros de codificación $IPD[b]$ para cada contenedor de frecuencia $[k]$ o para cada sub-banda de frecuencia $[b]$ de las señales de canales de audio $X_1[b]$, $X_2[b]$.

60 En una primera etapa, el generador de parámetros 105 aplica una transformada de tiempo-frecuencia en el canal de entrada del dominio temporal, p.e., el primer canal de entrada $x_1[n]$ y el canal de referencia del dominio temporal, p.e., el segundo canal de entrada $x_2[n]$. En el caso de reproducción estéreo, estos son los canales izquierdo y derecho. En una forma de realización preferida, la transformada de tiempo-frecuencia es una Transformada de Fourier Rápida (FFT). En una forma de realización alternativa, la transformada de frecuencia-tiempo es un banco de filtros de modulación cosenoidal o un banco de filtros completo.

65 En una segunda etapa, el generador de parámetros 105 calcula un espectro cruzado para cada contenedor de frecuencia $[b]$ de 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. El asterisco $*$ indica una conjugación completa. Para este caso, una sub-banda $[b]$ corresponde directamente a un contenedor de frecuencia $[k]$, mientras que el contenedor de frecuencia $[b]$ y $[k]$ representa exactamente el mismo contenedor de frecuencia.

como alternativa, el generador de parámetros 105 calcula el espectro cruzado por sub-banda $[b]$ 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 $[b]$ y $X_1[k]$ y $X_2[k]$ son los coeficientes de FFT de los dos canales. El asterisco $*$ indica una conjugación completa. k_b es la banda o sub-banda inicial b y k_{b+1} es la banda inicial de la sub-banda adyacente $b+1$. Por lo tanto, las bandas de frecuencia $[k]$ de FFT entre k_b y $k_{b+1}-1$ representan las sub-bandas $[b]$.

Para diferencias de fase entre canales (IPDs) se calculan por sub-banda sobre la base del espectro cruzado como:

$$IPD[b] = \angle c[b]$$

en donde la operación \angle es el operador de argumento para calcular el ángulo de $c[b]$.

En una forma de realización, el generador de parámetros 101 determina la primera media de parámetros de codificación $IPD_{mean}[i]$ de la señal de canal de audio $X_1[b]$ como una media del primer conjunto de parámetros de codificación $IPD[b]$ de la señal de canal de audio $X_1[b]$ por intermedio de los contenedores de frecuencia $[b]$ o sub-bandas de frecuencia $[b]$.

La IPD promediada (IPD_{mean}) a través de las bandas de frecuencia $[b]$ o las sub-bandas de frecuencia $[b]$ se calcula según se define en la ecuación siguiente:

$$IPD_{mean} = \frac{\sum_{k=1}^K IPD[k]}{K}$$

en donde K es el número de las bandas de frecuencia o de las sub-bandas de frecuencia que se tienen en cuenta para el cálculo de la media.

En una forma de realización, el generador de parámetros 101 determina la segunda media de parámetros de codificación $IPD_{mean_long_term}$ de la señal de canal de audio $X_1[b]$ como una media de una pluralidad de primera media de parámetros de codificación $IPD_{mean}[i]$ a través de una pluralidad de tramas de la señal de canal de audio $X_1[b]$, en donde cada primera media de parámetros de codificación $IPD_{mean}[i]$ está asociada a una trama $[i]$ de la señal de audio multicanal.

Sobre la base de la media IPD_{mean} anteriormente calculada, el generador de parámetros 105 calcula una media a largo plazo de IPD. La media $IPD_{mean_long_term}$ se calcula como la media a través de las N últimas tramas (a modo de ejemplo N puede establecerse en 10).

$$IPD_{mcan_long_term} = \frac{\sum_{i=1}^N IPD_{mean}[i]}{N}$$

En una forma de realización, el generador de parámetros 101 determina un valor absoluto IPD_{dist} de una diferencia entre la segunda media de parámetros de codificación $IPD_{mean_long_term}$ y la primera media de parámetros de codificación $IPD_{mean}[i]$.

Con el fin de evaluar la estabilidad del parámetro de IPD, la distancia entre IPD_{mean} e $IPD_{mean_long_term}$ (IPD_{dist}) se calcula a este respecto, lo que indica la evolución de IPD durante las N últimas tramas. En una forma de realización preferida, la distancia entre los valores de IPD local y a largo plazo se calcula como el valor absoluto de la diferencia entre la media local y a largo plazo.

$$IPD_{dist} = \text{abs}(IPD_{mean} - IPD_{mean_long_term})$$

5 Puede deducirse que si el parámetro de IPD_{mean} es estable a través de las tramas anteriores, la distancia IPD_{dist} se hace próxima a 0. La distancia es entonces igual a cero cuando la diferencia de fase es estable en el transcurso del tiempo. Esta distancia proporciona una buena estimación de la similitud de los canales.

10 En una forma de realización, el generador de parámetros 101 determina el parámetro de codificación ICC como una función del valor absoluto determinado IPD_{dist} . En una forma de realización, el generador de parámetros 101 determina el parámetro de codificación ICC a partir de una diferencia entre un primer valor paramétrico d y el valor absoluto determinado IPD_{dist} multiplicado por un segundo valor paramétrico e . En una forma de realización, el generador de parámetros 101 establece el primer valor paramétrico d a uno y establece el segundo valor paramétrico e a uno.

15 La coherencia o parámetro ICC se calcula como $ICC = 1 - IPD_{dist}$, puesto que ICC e IPD_{dist} tienen una relación inversa indirecta. El valor de ICC es próximo a 1 cuando los canales son similares e IPD_{dist} se hace igual a cero en ese caso.

20 Como alternativa, la ecuación para definir la relación entre ICC e IPD_{dist} se define como $ICC = d - e \cdot IPD_{dist}$ con los valores de d y e eligiéndose para representar mejor la relación inversa entre los dos parámetros. En otra forma de realización, la relación entre ICC e IPD_{dist} se obtiene a través de una base de datos amplia y se generaliza luego como $ICC = f(IPD_{dist})$.

25 Durante el segmento correlacionado de la señal de audio (a modo de ejemplo, para la señal de voz), el valor de IPD_{dist} es pequeño y durante las partes difusas de la entrada de audio (a modo de ejemplo, para señal de música), este parámetro de IPD_{dist} se hace mucho mayor y tendrá un valor próximo a 1 si los canales de entrada no están en correlación. De este modo, ICC e IPD_{dist} tienen una relación inversa indirecta.

30 La Figura 2 ilustra un diagrama de bloques de un decodificador de audio paramétrico 200 en conformidad con una forma de realización. El decodificador de audio paramétrico 200 recibe un flujo de bit 203 transmitido a través de un canal de comunicaciones como señal de entrada y proporciona una señal de audio multicanal decodificada 201 como señal de salida. El decodificador de audio paramétrico 200 comprende un decodificador de flujos de bit 217 acoplado al flujo de bit 203 para decodificar el flujo de bit 203 en un parámetro de codificación 215 y una señal codificada 213, un decodificador 209 acoplado al decodificador de flujo de bit 217 para generar una señal suma 211 a partir de la señal codificada 213, un decodificador paramétrico 205 acoplado al decodificador de flujo de bit 217 para decodificar un parámetro 221 a partir del parámetro de codificación 215 y un sintetizador 205 acoplado al decodificador paramétrico 205 y el decodificador 209 para sintetizar la señal de audio multicanal decodificada 201 a partir del parámetro 221 y de la señal suma 211.

40 El decodificador de audio paramétrico 200 genera los canales de salida de su señal de audio multicanal 201 tal como ICTD, ICLD y/o ICC entre los canales en proximidad 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 de ICTD, ICLD e ICC estimados entre un par de canales contienen aproximadamente dos órdenes de magnitud de menos información que una forma de onda de audio. No solamente la baja tasa binaria sino también el aspecto de la compatibilidad retroactiva es de interés. La señal suma transmitida corresponde a una mezcla reducida monoaural de la señal multicanal o estéreo.

50 La Figura 3 ilustra un diagrama de bloques de un codificador de audio estéreo paramétrico 301 y un decodificador 303 en conformidad con una forma de realización. El decodificador de audio estereero paramétrico 301 corresponde al decodificador de audio paramétrico 100 según se describe con respecto a la Figura 1 pero la señal de audio multicanal 101 es una señal de audio estéreo con canales de audio izquierdo 305 y derecho 307.

55 El codificador de audio estéreo paramétrico 301 recibe la señal de audio estéreo 305, 307 que comprende una señal de audio de canal izquierdo 305 y una señal de audio de canal de audio 307, como señal de entrada y proporciona un flujo de bit como señal de salida 309. El codificador de audio estéreo paramétrico 301 comprende un generador de parámetros 311 acoplado a la señal de audio estéreo 305, 307 para generar parámetros espaciales 313, un generador de señales mezclada reducida 315 acoplado a la señal de audio estéreo 305, 307 para generar una señal mezclada reducida 317 o señal suma, un codificador monoaural 319 acoplado al generador de señales mezcladas reducidas 315 para codificar la señal mezclada reducida 317 para proporcionar una señal de audio codificada 321 y un combinador de flujo de bits 323 acoplado al generador de parámetros 311 y al codificador monoaural 319 para combinar el parámetro de codificación 313 y la señal de audio codificada 321 para un flujo de bits con el fin de proporcionar la señal de salida 309. En el generador de parámetros 311, los parámetros espaciales 313 se extraen y cuantifican antes de que sean objeto de multiplexación en el flujo de bits.

65 El decodificador de audio estéreo paramétrico 303 recibe el flujo de bits, esto es, la señal de salida 309 del codificador de audio estéreo paramétrico 301 transmitida a través de un canal de comunicaciones, como una señal

de entrada y proporciona una señal de audio estéreo decodificada con canal izquierdo 325 y canal derecho 327 como señal de salida. El decodificador de audio estéreo paramétrico 303 comprende un decodificador de flujos de bits 329 acoplado al flujo de bit recibido 309 para decodificar el flujo de bits 309 en parámetros de codificación 331 y una señal codificada 333, un decodificador monoaural 335 acoplado al decodificador de flujo de bits 329 para generar una señal suma 337 a partir de la señal codificada 333, un decodificador paramétrico espacial 339 acoplado al decodificador de flujos de bit 329 para decodificar parámetros espaciales 341 a partir de los parámetros de codificación 331 y un sintetizador 343 acoplado al decodificador paramétrico espacial o sistema de resolución 339 y el decodificador monoaural 335 para sintetizar la señal de audio estéreo decodificada 325, 327 a partir de los parámetros espaciales 341 y de la señal suma 337.

El procesamiento en el codificador de audio estéreo paramétrico 301 es capaz de extraer retardos y calcular el nivel de las señales de audio de forma adaptativa en el tiempo y frecuencia para generar los parámetros espaciales 313, p.e., diferencias temporales entre canales (ICTDs) y diferencias de niveles entre canales (ICLDs). Además, el codificador de audio estéreo paramétrico 301 realiza un filtrado adaptativo temporal eficientemente para la síntesis de coherencia entre canales (ICC). En una forma de realización, el codificador estéreo paramétrico utiliza una Transformada de Fourier de Tiempo Corto (STFT) que sirve como base para un banco de filtros para la puesta en práctica eficiente de los sistemas de codificación de pistas binaurales (BCC) con baja complejidad de cálculo. El procesamiento en el codificador de audio estéreo paramétrico 301 tiene baja complejidad de cálculo y bajo retardo. Haciendo 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.

El generador de parámetros 311 ilustrado en la Figura 3 es funcionalmente el mismo que el generador de parámetros correspondiente 105 descrito con respecto a la Figura 1, con la excepción de que la cuantización y la codificación de las pistas espaciales se ha añadido para fines ilustrativos. La señal suma 317 está codificada con un codificador de audio monoaural convencional 319. En una forma de realización, el codificador de audio estéreo paramétrico 301 utiliza una transformación de tiempo-frecuencia basada en la transformada STFT para transformar la señal de canal de audio estéreo 305, 307 en el dominio de la frecuencia. La STFT aplica una Transformada de Fourier discreta (DFT) para las partes de ventanas temporales de una señal de entrada $x(n)$. Una trama de señal de N muestras se multiplica con una ventana de longitud W antes de que se aplique una DFT de N puntos. Las ventanas adyacentes son objeto de solapamiento y se desplazan en W/2 muestras. La ventana se elige de modo que las ventanas en solapamiento se añadan hasta un valor constante de 1. Por lo tanto, para la transformación inversa, no existe necesidad de establecer ventanas adicionales. Una DFT inversa sencilla de tamaño N con avance temporal de tramas sucesivas de W/2 muestras se utiliza en el decodificador 303. Si no se modifica el espectro, la reconstrucción perfecta se consigue mediante solapamiento/adición.

Puesto que la resolución espectral uniforme de la STFT no está bien adaptada a la percepción humana, los coeficientes espectrales uniformemente espaciados, objeto de salida de la STFT, se agrupan en B particiones no solapantes 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 1. En una forma de realización alternativa, el codificador de audio estéreo paramétrico 301 utiliza un banco de filtro no uniforme para transformar la señal de canal de audio estéreo 305, 307 en el dominio de la frecuencia.

En una forma de realización, el mezclador-reductor 315 determina los coeficientes espectrales de una partición b o de una sub-banda b de la señal suma ecualizada $S_m(k)$ 317 mediante

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

en donde $X_{c,m}(k)$ son los espectros de los canales de audio de entrada 305, 307 y $e_b(k)$ es un factor de ganancia calculado como

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

con estimaciones de la potencia de partición,

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

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

Para impedir la presencia de elementos de ayuda resultantes de factores de ganancia de gran magnitud cuando la atenuación de la suma de las señales de sub-banda es importante, los factores de ganancia $e_b(k)$ pueden limitarse a 6 dB, esto es, $e_b(k) \leq 2$.

En una forma de realización, el generador de parámetros 311 aplica una transformación de tiempo-frecuencia, p.e., la STFT según se describió anteriormente o una FFT sobre los canales de entrada, esto es, en el canal izquierdo 305 y en el canal derecho 307. En una forma de realización, la transformación de tiempo-frecuencia es una Transformada de Fourier Rápida (FFT). En una forma de realización alternativa, la transformación de tiempo-frecuencia se realiza en una banco de filtros de modulación cosenoidal o en un banco de filtros complejo.

El generador de parámetros 311 calcula un espectro cruzado para cada contenedor de frecuencia [b] de la FFT o de la STFT como

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

Para este caso, una sub-banda [b] corresponde directamente a un solo contenedor de frecuencia [k], representando un contenedor de frecuencia [b] y [k] exactamente el mismo contenedor de frecuencia.

Como alternativa, el generador de parámetros 311 calcula el espectro cruzado para la sub-banda [k] 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 banda b o la sub-banda k. $X_1[k]$ y $X_2[k]$ son los coeficientes de FFT del canal izquierdo 305 y del canal derecho 307. El operador de asterisco * indica una conjugación compleja. k_b es la banda inicial de la sub-banda k y k_{b+1} es la banda inicial de la sub-banda b+1 adyacente. En consecuencia, las bandas de frecuencia [k] de una FFT o STFT entre k_b y $k_{b+1}-1$ representan las sub-bandas [b].

Las diferencias de fase entrecanales (IPDs) se calculan por sub-banda sobre la base del espectro cruzado como:

$$IPD[b] = \angle c[b]$$

en donde la operación \angle es el operador de argumento para el cálculo del ángulo de $c[b]$.

A continuación, el generador de parámetros 311 calcula el valor promediado de IPD (IPD_{mean}) a través de las bandas de frecuencia o de las sub-bandas de frecuencia según se define en la ecuación siguiente.

$$IPD_{mean} = \frac{\sum_{k=1}^K IPD[k]}{K}$$

en donde K es el número de las bandas de frecuencia o de las sub-bandas de frecuencia que se tienen en cuenta para el cálculo de la media.

A continuación, sobre la base del valor de IPD_{mean} anteriormente calculado, el generador de parámetros 311 calcula una media a largo plazo de IPD. El $IPD_{mean_long_term}$ se calcula como la media a través de las N últimas tramas, en una forma de realización, siendo N establecido a un valor de 10.

$$\text{IPD}_{\text{mean_long_term}} = \frac{\sum_{i=1}^N \text{IPD}_{\text{mean}} [i]}{N}$$

5 Con el fin de evaluar la estabilidad del parámetro de IPD, el generador de parámetros 311 calcula la distancia IPD_{dist} entre IPD_{mean} y $\text{IPD}_{\text{mean_long_term}}$ que ilustra la evolución del IPD durante las N últimas tramas. En una forma de realización, la distancia entre el IPD local y a largo plazo se calcula como el valor absoluto de la diferencia entre la media local y la media a largo plazo:

$$\text{IPD}_{\text{dist}} = \text{abs}(\text{IPD}_{\text{mean}} - \text{IPD}_{\text{mean_long_term}})$$

10 Puede deducirse que si el parámetro de IPD_{mean} es estable a través de las tramas anteriores, el valor de la distancia IPD_{dist} se hace próximo a 0. La distancia es entonces igual a cero cuando la diferencia de fase es estable en el transcurso del tiempo. Esta distancia proporciona una buena estimación de la similitud de los canales.

15 En una forma de realización, el generador de parámetros 311 calcula la coherencia o parámetro ICC como $\text{ICC} = 1 - \text{IPD}_{\text{dist}}$ puesto que ICC e IPD_{dist} tienen una relación inversa indirecta. El valor de ICC está próximo a 1 cuando los canales son similares y el valor de IPD_{dist} se hace igual a 0 en ese caso.

20 Como alternativa, el generador de parámetros 311 utiliza la relación entre ICC e IPD_{dist} definida como $\text{ICC} = d - e \cdot \text{IPD}_{\text{dist}}$ con d y e siendo parámetros elegidos para representar mejor la relación inversa entre los dos parámetros ICC e IPD_{dist} . En una forma de realización alternativa, el generador de parámetros 311 obtiene la relación entre ICC e IPD_{dist} a través de una amplia base de datos que se generaliza como $\text{ICC} = f(\text{IPD}_{\text{dist}})$.

25 Durante un segmento en correlación de una señal de audio, a modo de ejemplo, para la señal de voz, el valor de IPD_{dist} es pequeño y durante partes difusas de la salida de audio, a modo de ejemplo, para señal de música, este parámetro de IPD_{dist} se hace mucho mayor y será próximo a 1 si los canales de entrada no están en correlación. De este modo, ICC e IPD_{dist} tienen una relación inversa indirecta.

30 El generador de parámetros 311 utiliza IPD_{dist} para una estimación aproximada del valor de ICC. El espectro cruzado requiere una más baja complejidad que el cálculo de la correlación. Además, en caso de cálculo del IPD en el codificador de audio espacial paramétrico, este espectro cruzado está ya calculado y se deduce, en consecuencia, la complejidad total.

35 La Figura 4 ilustra un diagrama esquemático de un método 400 para generar un parámetro de codificación en conformidad con una forma de realización. El método 400 para generar el parámetro de codificación ICC para una señal de canal de audio $x_1[n]$ de entre una pluralidad de señales de canales de audio $x_1[n]$, $x_2[n]$ de una señal de audio multicanal. Cada señal de canal de audio $x_1[n]$, $x_2[n]$ tiene valores de señales de canales de audio. La Figura 4 ilustra el caso de estéreo en donde la pluralidad de señales de canales de audio comprende un canal de audio izquierdo $x_1[n]$ y un canal de audio derecho $x_2[n]$. El método 400 comprende:

40 aplicar una transformación de FFT 401 a la señal de canal de audio izquierdo $x_1[n]$ y aplicar una transformación de FFT 403 a la señal de canal de audio derecho $x_2[n]$ para obtener señales de canales de audio en el dominio de la frecuencia $X_1[b]$ y $X_2[b]$, en donde $X_1[b]$ es la señal de canal de audio izquierdo y $X_2[b]$ es la señal de canal de audio derecho con respecto al contenedor de frecuencia [b] en el dominio de la frecuencia. como alternativa, se aplica una transformación de banco de filtros a la señal de canal de audio izquierdo $x_1[n]$ y a la señal de canal de audio derecho $x_2[n]$ para obtener señales de canales de audio $X_1[b]$, $X_2[b]$ en sub-bandas de la frecuencia, en donde [b] indica la sub-banda de frecuencia;

50 determinar 405 una correlación cruzada $c[b]$ de cada contenedor de frecuencia [b] de la señal de canal de audio izquierdo $X_1[b]$ y la señal de canal de audio derecho $X_2[b]$ o de forma alternativa, determinar 405 una correlación cruzada $c[b]$ de cada sub-banda de frecuencia [b] de la señal de canal de audio izquierdo $X_1[b]$ y de la señal de canal de audio derecho $X_2[b]$;

55 determinar 407 para la señal de canal de audio $X_1[b]$ de la pluralidad de señales de canales de audio un primer conjunto de parámetros de codificación $\text{IPD}[b]$ de los valores de señales de canales de audio de la señal de canal de audio $X_1[b]$ y los valores de señales de audio de referencia de una señal de audio de referencia $X_2[b]$, en donde la señal de audio de referencia es otra señal de canal de audio $X_2[b]$ de la pluralidad de señales de canales de audio o una señal de audio mezclada reducida derivada de al menos dos señales de canales de audio de la pluralidad de señales de audio multicanal. La Figura 4 ilustra el caso de estéreo, en donde la operación de determinación 407 determina para la señal de canal de audio izquierdo $X_1[b]$ el primer conjunto de parámetros de codificación $\text{IPD}[b]$ y en donde la señal de audio de referencia es la señal de canal de audio derecho $X_2[b]$;

60 determinar 409 para la señal de canal de audio $X_1[b]$ una primera media de parámetros de codificación $\text{IPD}_{\text{mean}}[i]$

sobre la base del primer conjunto de parámetros de codificación IPD[b] de la señal de canal de audio $X_1[b]$;

determinar 411 para la señal de canal de audio $X_1[b]$ una segunda media de parámetros de codificación $IPD_{mean_long_term}$ sobre la base de la primera media de parámetros de codificación $IPD_{mean}[i]$ de la señal de canal de audio $X_1[b]$ y al menos una otra media de parámetros de codificación $IPD_{mean}[i-1]$ de la señal de canal de audio $X_1[b]$. La otra primera media de parámetros de codificación $IPD_{mean}[i-1]$ se calcula a partir de las N-1 tramas anteriores de la señal de canal de audio $X_1[b]$; y

determinar 413 o calcular el parámetro de codificación ICC sobre la base de la primera media de parámetros de codificación $IPD_{mean}[i]$ de la señal de canal de audio $X_1[b]$ y la segunda media de parámetros de codificación $IPD_{mean_long_term}$ de la señal de canal de audio $X_1[b]$.

En una forma de realización, el primer conjunto de parámetros de codificación IPD[b] de la señal de canal de audio $X_1[b]$ está ya disponible y el método 400 se inicia con las etapas 409, 411 y 413 según se describió con anterioridad.

Aunque no se ilustra en la Figura 4, el método 400 es aplicable al caso general de señales de audio multicanal, siendo entonces la señal de referencia otra señal de canal de audio o una señal de audio mezclada reducida según se describió anteriormente con respecto a la Figura 1.

En una forma de realización, el método 400 se procesa como sigue:

en una primera etapa 401, 403, se aplica una transformación de tiempo-frecuencia en los canales de entrada (canal izquierdo y canal derecho en caso de estéreo). En una forma de realización preferida, la transformación de tiempo-frecuencia se realiza con una Transformada de Fourier rápida (FFT). En una forma de realización alternativa, la transformación de tiempo-frecuencia puede realizarse con un banco de filtros de modulación cosenoidal o un banco de filtros complejo.

En una segunda etapa 405, un espectro cruzado para cada contenedor de frecuencia de la FFT se calcula mediante

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

en donde una sub-banda [b] corresponde directamente a un solo contenedor de frecuencia [k], con el contenedor de frecuencia [b] y [k] representando exactamente el mismo contenedor de frecuencia.

Como alternativa, el espectro cruzado puede calcularse 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 banda b o 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 caso de estéreo). El asterisco * indica una conjugación completa. k_b es la banda inicial de sub-banda b y k_{b+1} es la banda inicial de la sub-banda adyacente b+1. Por lo tanto, las bandas de frecuencia [k] de la FFT entre k_b y $k_{b+1}-1$ representan las sub-bandas [b].

En una tercera etapa 407, las diferencias de fase entre canales (IPDs) se calculan por sub-banda, sobre la base del espectro cruzado como

$$IPD[b] = \angle c[b]$$

en donde la operación \angle es el operador de argumento para calcular el ángulo de $c[b]$.

En una cuarta etapa 409, el IPD promediado (IPD_{mean}) a través de las bandas de frecuencia (o sub-bandas de frecuencia) se calcula también según se define en la ecuación siguiente:

$$IPD_{mean} = \frac{\sum_{k=1}^K IPD[k]}{K}$$

en donde K es el número de las bandas de frecuencia o de las sub-bandas de frecuencia que se tienen en cuenta para el cálculo de la media.

En una quinta etapa 411, basada en el valor de IPD_{mean} anteriormente calculado se determina una media a largo plazo de IPD. El $IPD_{mean_long_term}$ se calcula como la media a través de las N últimas tramas (a modo de ejemplo, N puede establecerse a 10).

$$IPD_{mean_long_term} = \frac{\sum_{i=1}^N IPD_{mean} [i]}{N}$$

Con el fin de evaluar la estabilidad del parámetro de IPD, la distancia entre IPD_{mean} e $IPD_{mean_long_term}$ (IPD_{dist}) es objeto de cálculo, lo que muestra la evolución de IPD durante las N últimas tramas. En una forma de realización preferida, la distancia entre el IPD local y a largo plazo se calcula como el valor absoluto de la diferencia entre la media local y la media a largo plazo:

$$IPD_{dist} = \text{abs}(IPD_{mean} - IPD_{mean_long_term})$$

Puede deducirse que si el parámetro de IPD_{mean} es estable a través de las tramas anteriores, la distancia IPD_{dist} se hace próxima a 0. La distancia es entonces igual a cero cuando la diferencia de fase es estable en el transcurso del tiempo. Esta distancia proporciona una buena estimación de la similitud de los canales.

En una sexta etapa 413, el parámetro de ICC o la coherencia se calcula por $ICC = 1 - IPD_{dist}$, puesto que ICC e IPD_{dist} tienen una relación inversa indirecta. El valor de ICC es próximo a 1 cuando los canales son similares e IPD_{dist} se hace igual a cero en ese caso.

En una forma de realización alternativa de la sexta etapa 413, la ecuación para definir la relación entre ICC e IPD_{dist} se define como $ICC = d - e \cdot IPD_{dist}$ con los parámetros d y e eligiéndose para representar mejor la relación inversa entre los dos parámetros ICC e IPD_{dist} . En otra forma de realización de la sexta etapa 413, la relación entre ICC e IPD_{dist} se obtiene mediante una base de datos amplia y puede generarse entonces como $ICC = f(IPD_{dist})$.

Durante un segmento en correlación de una señal de audio (a modo de ejemplo, para señal de voz), el valor de IPD_{dist} es pequeño y durante partes difusas de la entrada de audio (a modo de ejemplo, para señal de música), este parámetro de IPD_{dist} se hace mucho mayor y su valor será próximo a 1 si los canales de entrada no están en correlación. De este modo, ICC e IPD_{dist} tienen una relación inversa indirecta.

De lo que antecede un experto en esta técnica deducirá que se dan a conocer una diversidad de métodos, sistemas, programas informáticos en soportes de registro y 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 que al menos un ordenador ejecute las etapas de realización y cálculo aquí descritas.

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

Numerosas alternativas, modificaciones y variantes 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 aplicaciones 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 sin desviarse por ello del alcance de protección de la presente invención. Por lo tanto, ha de entenderse que dentro del alcance de las reivindicaciones adjuntas y de sus equivalentes, puede realizarse la invención de forma distinta a la aquí concretamente descrita.

Una forma de realización correspondiente de la presente invención puede aplicarse en el codificador de la extensión estéreo de las normas ITU-T G.722, G.722 Anexo B, G.711.1 y/o G.711.1 Anexo D. Además, el método descrito puede aplicarse también para un codificador de voz y de audio para aplicación móvil según se define en 3GPP EVS (servicio de voz mejorado) códec.

REIVINDICACIONES

1. Un codificador de audio paramétrico (100) para generar un parámetro de codificación (ICC) para una señal de canal de audio ($X_1[b]$) de una pluralidad de señales de canales de audio ($X_1[b]$, $X_2[b]$) de una señal de audio multicanal, teniendo cada señal de canal de audio ($X_1[b]$, $X_2[b]$) valores de señal de canal de audio ($X_1[k]$, $X_2[k]$), siendo el parámetro de codificación un parámetro de coherencia intercanales (ICC), comprendiendo el codificador de audio paramétrico (100) un generador de parámetros (105), estando el generador de parámetros (105) para
- para determinar para la señal de canal de audio ($X_1[b]$) de la pluralidad de señales de canales de audio un primer conjunto de parámetros de codificación ($IPD[b]$) a partir de los valores de señal de canal de audio ($X_1[k]$) de la señal de canal de audio ($X_1[b]$) y valores de señal de audio de referencia ($X_2[k]$) de una señal de audio de referencia ($X_2[b]$), en donde la señal de audio de referencia es otra señal de canal de audio ($X_2[b]$) de la pluralidad de señales de canales de audio o una señal de audio mezclada reducida derivada de al menos dos señales de canales de audio de la pluralidad de señales de audio multicanales, en donde el primer conjunto de parámetros de codificación ($IPD[b]$) son parámetros de diferencia de fase intercanales o parámetros de diferencia de fase intercanales de sub-banda,
 - para determinar para la señal de canal de audio ($X_1[b]$) un primer valor medio de parámetros de codificación ($IPD_{mean}[i]$) sobre la base del primer conjunto de parámetros de codificación ($IPD[b]$) del primer canal de señal de audio ($X_1[b]$), refiriéndose el primer valor medio de parámetros de codificación a una trama corriente de la señal de canal de audio, en donde el generador de parámetros (105) está configurado para determinar el primer valor medio de parámetros de codificación ($IPD_{mean}[i]$) de la señal de canal de audio ($X_1[b]$) como una media del primer conjunto de parámetros de codificación ($IPD[b]$) de la señal de canal de audio ($X_1[b]$) a través de bandas de frecuencias $[k]$ o de sub-bandas de frecuencia $[b]$,
 - para determinar para la señal de canal de audio ($X_1[b]$) una segunda media de parámetros de codificación ($IPD_{mean_long_term}$) en función de la primera media de parámetros de codificación ($IPD_{mean}[i]$) de la señal de canal de audio ($X_1[b]$) y al menos una otra primera media de parámetros de codificación ($IPD_{mean}[i-1]$) de la señal de canal de audio ($X_1[b]$), refiriéndose a la al menos una otra primera media de parámetros de codificación a una trama anterior de la señal de canal de audio, y
 - para determinar el parámetro de codificación (ICC) sobre la base de la primera media de parámetros de codificación ($IPD_{mean}[i]$) de la señal de canal de audio ($X_1[b]$) y la segunda media de parámetros de codificación ($IPD_{mean_long_term}$) de la señal de canal de audio ($X_1[b]$);
- en donde el generador de parámetros (105) está configurado, además
- para determinar un valor absoluto (IPD_{dist}) de una diferencia entre la segunda media de parámetros de codificación ($IPD_{mean_long_term}$) y la primera media de parámetros de codificación ($IPD_{mean}[i]$), y
 - para determinar el parámetro de codificación (ICC) como una función de valor absoluto determinado (IPD_{dist}).
2. El codificador de audio paramétrico (100) según la reivindicación 1, en donde el generador de parámetros (105) está configurado para determinar diferentes de fase de valores de señal de canal de audio siguientes ($X_1[k]$) con el fin de obtener el primer conjunto de parámetros de codificación ($IPD[b]$).
3. El codificador de audio paramétrico (100) según una de las reivindicaciones precedentes, en donde la señal de canal de audio ($X_1[b]$) y la señal de audio de referencia ($X_2[b]$) son señales del dominio de frecuencia y en donde los valores de la señal de canal de audio ($X_1[k]$) y los valores de la señal de audio de referencia ($X_2[k]$) están asociados con las bandas de frecuencia (k) o las sub-bandas de frecuencia (b).
4. El codificador de audio paramétrico (100) según una de las reivindicaciones precedentes, que comprende, además, un transformador (FFT) para transformar una pluralidad de señales de canales de audio en el dominio temporal ($x_1[n]$, $x_2[n]$) en el dominio de la frecuencia para obtener la pluralidad de señales de canales de audio ($X_1[b]$, $X_2[b]$).
5. El codificador de audio paramétrico (100) según una de las reivindicaciones precedentes, en donde el generador de parámetros (105) está configurado para determinar el primer conjunto de parámetros de codificación ($IPD[b]$) para cada contenedor de frecuencia ($[k]$) o para cada sub-banda de frecuencia ($[b]$) de las señales de canales de audio ($X_1[b]$, $X_2[b]$).
6. El codificador de audio paramétrico (100) según una de las reivindicaciones precedentes, en donde el generador de parámetros (105) está configurado para determinar la segunda media de parámetros de codificación ($IPD_{mean_long_term}$) de la señal de canal de audio ($X_1[b]$) como una media de una pluralidad de primeras medias de parámetros de codificación ($IPD_{mean}[i]$) en una pluralidad de tramas de la señal de canal de audio ($X_1[b]$), en donde cada primera media de parámetros de codificación ($IPD_{mean}[i]$) está asociada a una primera trama (i) de la señal de

audio multicanal.

7. El codificador paramétrico (100) según una de las reivindicaciones precedentes, en donde el generador de parámetros (105) está configurado para determinar el parámetro de codificación (ICC) a partir de una diferencia entre un primer valor paramétrico (d) y el valor absoluto determinado (IPD_{dist}) multiplicado por un segundo valor paramétrico (e).

8. El codificador de audio de paramétrico (100) según la reivindicación 7, en donde el generador de parámetros (105) está configurado para establecer el primer valor paramétrico (d) a uno y para establecer el segundo valor paramétrico (e) a uno.

9. El codificador de audio paramétrico (100) según una de las reivindicaciones precedentes, que comprende, además, un generador de señales mezcladas reducidas para superponer al menos dos de las señales de canales de audio de la señal de audio multicanal con el fin de obtener una señal mezclada reducida, un codificador de audio, en particular, un codificador mono, para codificar la señal mezclada reducida para obtener una señal de audio codificada y un combinador para combinar la señal de audio codificada con un parámetro de codificación correspondiente.

10. Un método (400) para generar un parámetro de codificación (ICC) para una señal de canal de audio ($X_1[b]$) de una pluralidad de señales de canales de audio ($X_1[b]$, $X_2[b]$) de una señal de audio multicanal, teniendo cada señal de canal de audio ($X_1[b]$, $X_2[b]$) valores de señales de canales de audio ($X_1[k]$, $X_2[k]$), siendo el parámetro de codificación un parámetro de coherencia intercanales (ICC), comprendiendo el método (400):

- la determinación (407) para la señal de canal de audio ($X_1[b]$) de la pluralidad de señales de canales de audio un primer conjunto de parámetros de codificación ($IPD[b]$) a partir de los valores de señales de canales de audio ($X_1[k]$) de la señal de canal de audio ($X_1[b]$) y los valores de la señal de audio de referencia ($X_2[k]$) de una señal de audio de referencia ($X_2[b]$), en donde la señal de audio de referencia es otra señal de canal de audio ($X_2[b]$) de la pluralidad de señales de canales de audio o una señal de audio mezclada reducida derivada de al menos dos señales de canales de audio de la pluralidad de señales de audio multicanales, en donde el primer conjunto de parámetros de codificación ($IPD[b]$) son parámetros de diferencia de fase intercanal o parámetros de diferencia de fase intercanal de sub-banda,

- la determinación (409) para la señal de canal de audio ($X_1[b]$) de una primera media de parámetros de codificación ($IPD_{mean[i]}$) sobre la base del primer conjunto de parámetros de codificación ($IPD[b]$) de la señal de canal de audio ($X_1[b]$), refiriéndose la primera media de parámetros de codificación a una trama corriente de la señal de canal de audio, en donde la primera media de parámetros de codificación ($IPD_{mean[i]}$) se determina como una media del primer conjunto de parámetros de codificación ($IPD[b]$) de la señal de canal de audio ($X_1[b]$) a través de bandas de frecuencia [k] o sub-bandas de frecuencia [b],

- la determinación (411) para la señal de canal de audio ($X_1[b]$) de una segunda media de parámetros de codificación ($IPD_{mean_long_term}$) sobre la base de la primera media de parámetros de codificación ($IPD_{mean[i]}$) de la señal de canal de audio ($X_1[b]$) y al menos una otra primera media de parámetros de codificación ($IPD_{mean[i-1]}$) de la señal de canal de audio ($X_1[b]$), refiriéndose a por lo menos otra primera media de parámetros de codificación a una trama anterior de la señal de canal de audio, y

- la determinación (413) del parámetro de codificación (ICC) sobre la base de la primera media de parámetros de codificación ($IPD_{mean[i]}$) de la señal de canal de audio ($X_1[b]$) y la segunda media de parámetros de codificación ($IPD_{mean_long_term}$) de la señal de canal de audio ($X_1[b]$);

en donde la determinación (413) del parámetro de codificación (ICC) sobre la base de la primera media de parámetros de codificación ($IPD_{mean[i]}$) de la señal de canal de audio ($X_1[b]$) y la segunda media de parámetros de codificación ($IPD_{mean_long_term}$) de la señal de canal de audio comprende:

- la determinación de un valor absoluto (IPD_{dist}) de una diferencia entre la segunda media de parámetros de codificación ($IPD_{mean_long_term}$) y la primera media de parámetros de codificación ($IPD_{mean[i]}$), y

- la determinación del parámetro de codificación (ICC) como una función del valor absoluto determinado (IPD_{dist}).

11. Un programa informático que está configurado para poner en práctica el método según la reivindicación 10 cuando se ejecuta en un ordenador.

65

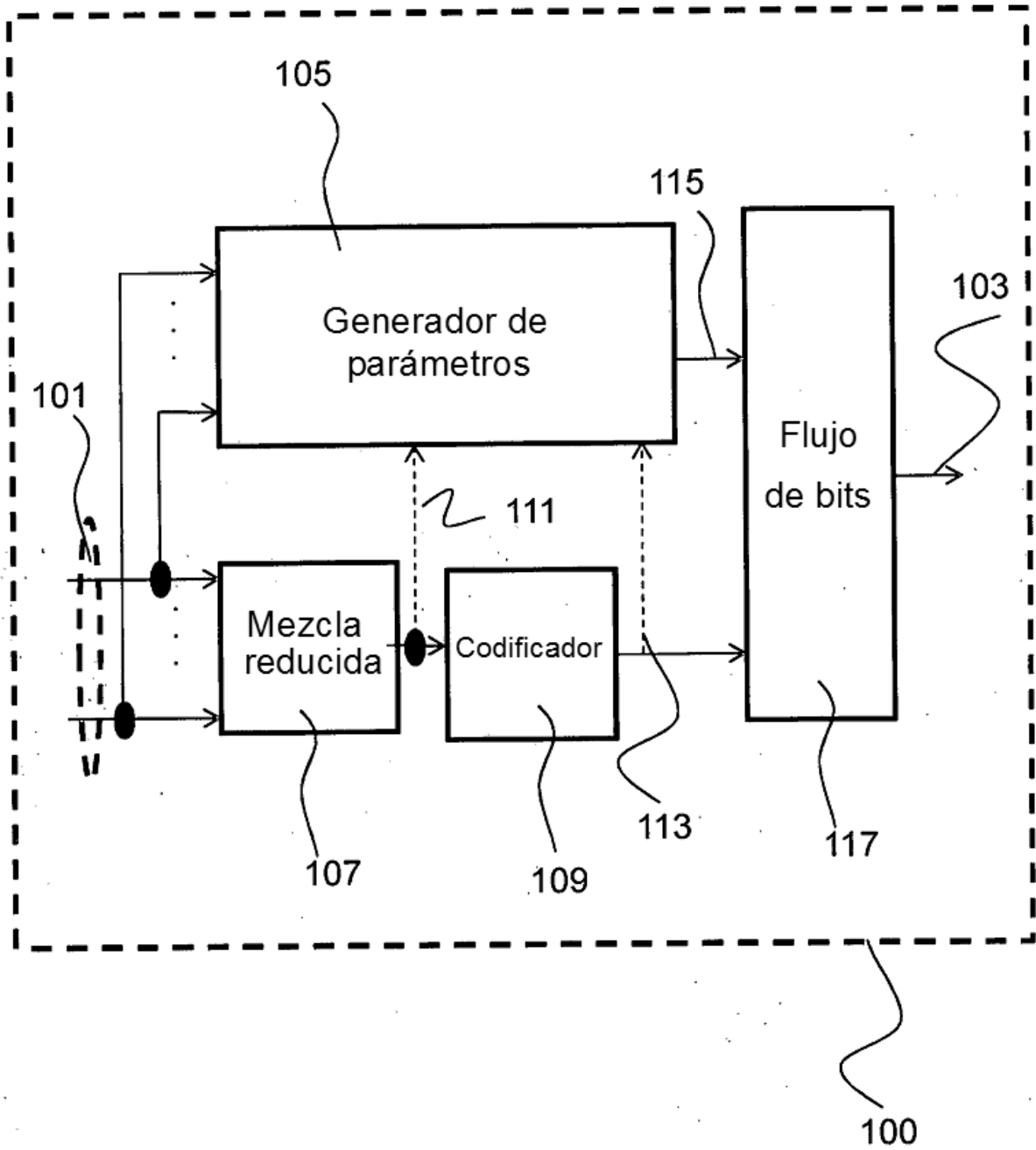


Fig. 1

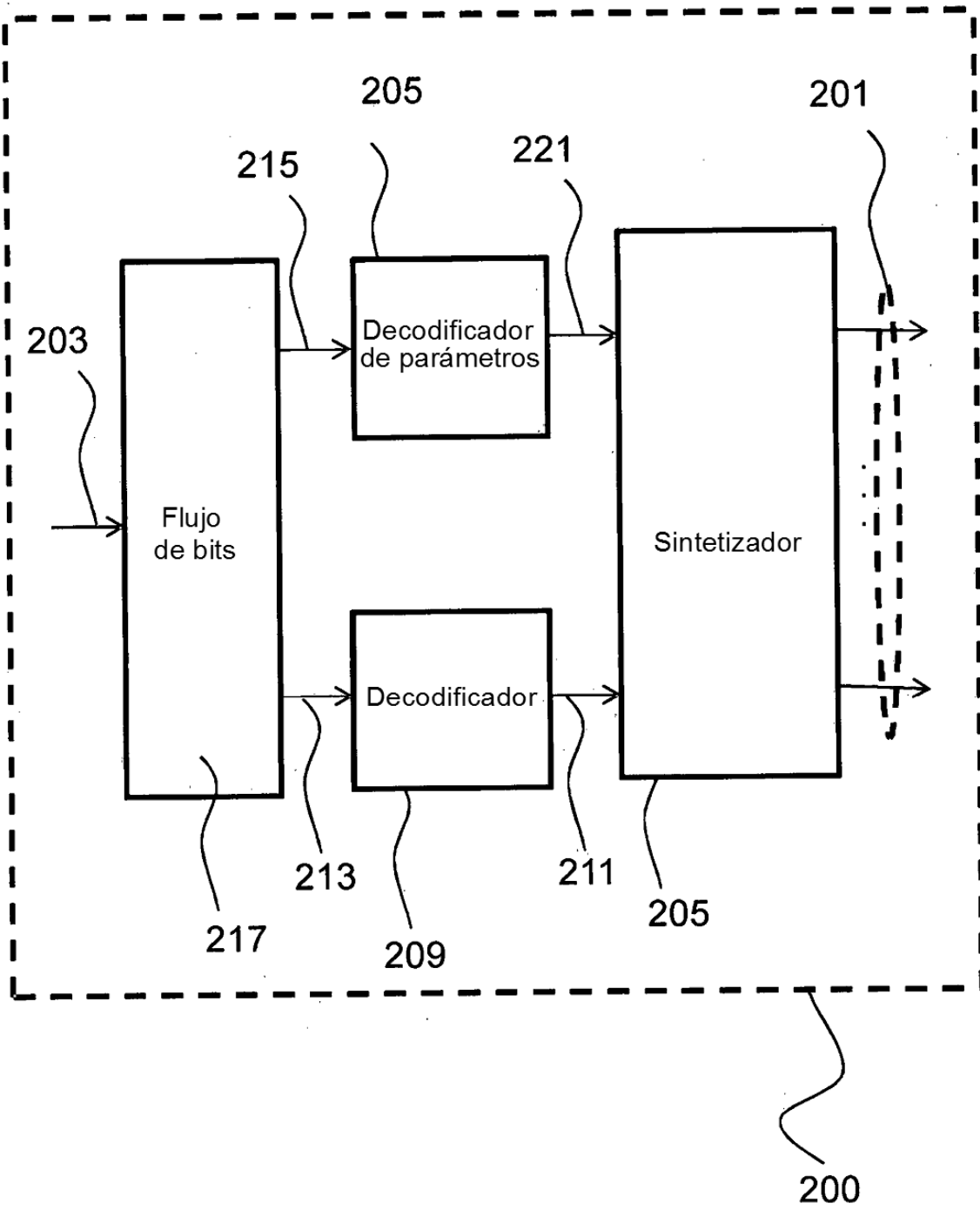


Fig. 2

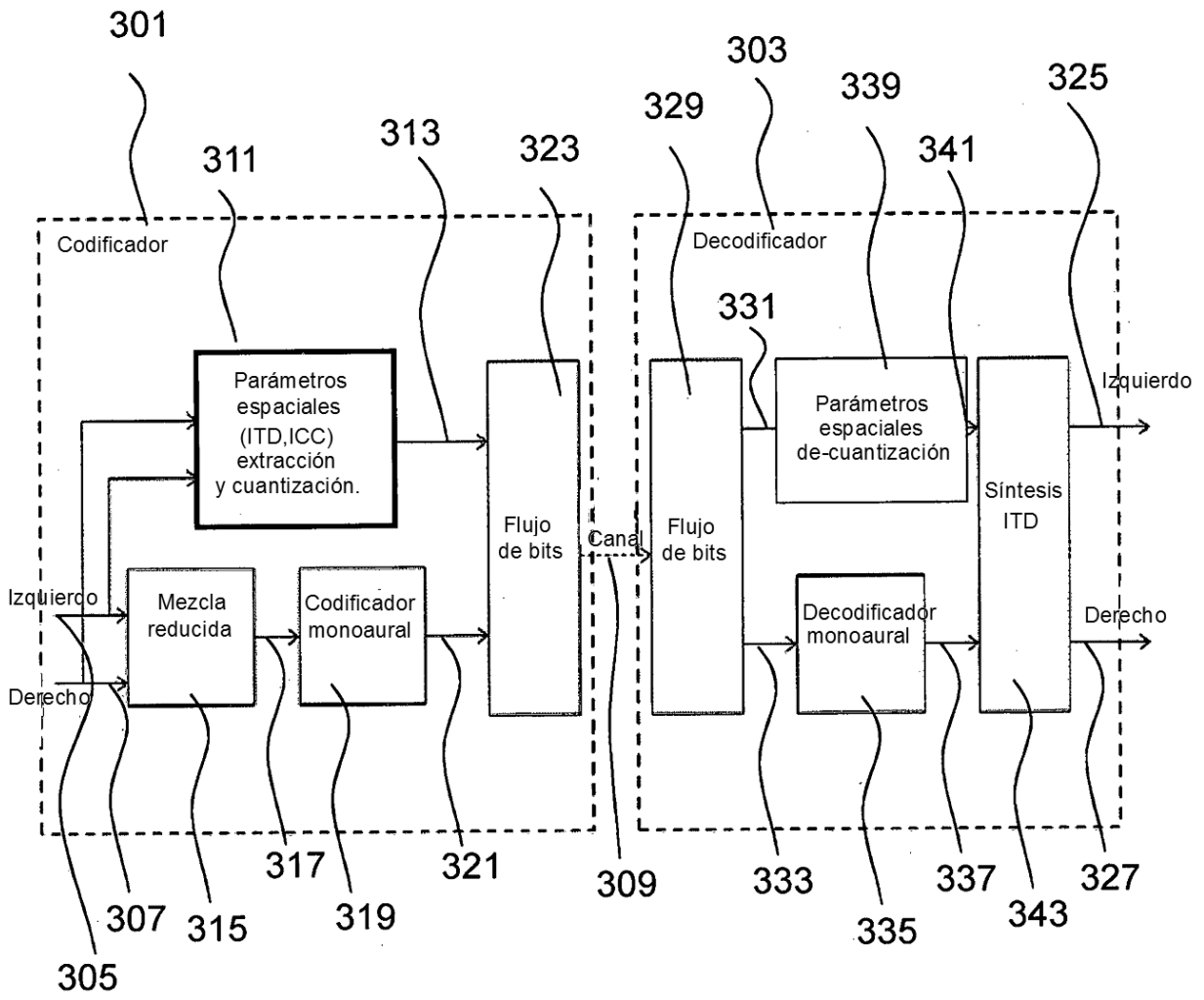


Fig. 3

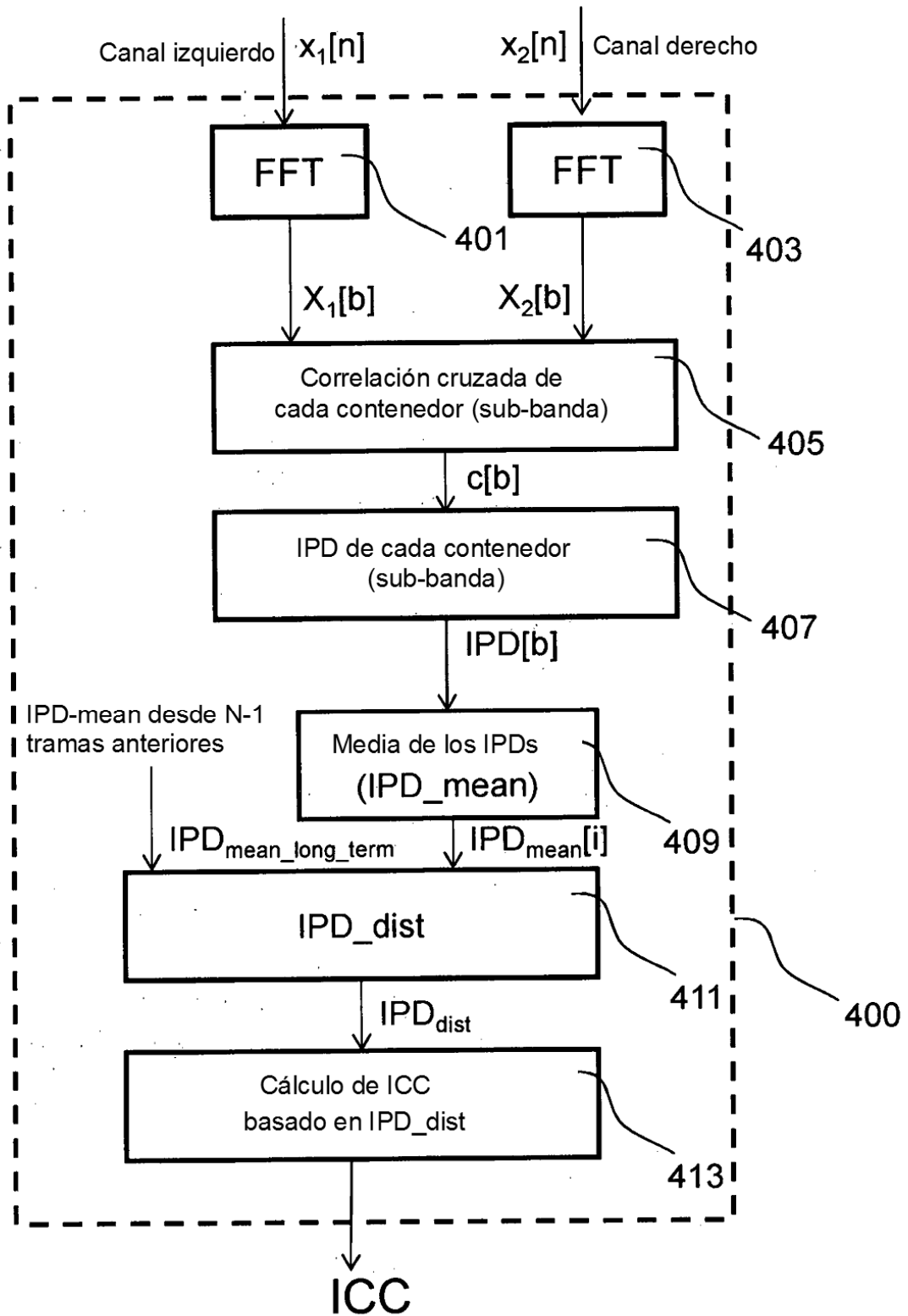


Fig. 4