

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 540 215**

51 Int. Cl.:

**H04R 25/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.04.2012 E 12712126 (7)**

97 Fecha y número de publicación de la concesión europea: **01.04.2015 EP 2702587**

54 Título: **Método de estimación de diferencia inter-canal y dispositivo de codificación de audio espacial**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**09.07.2015**

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 2 540 215 T3**

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

## DESCRIPCIÓN

Método de estimación de diferencia inter-canal y dispositivo de codificación de audio espacial

## 5 CAMPO DE LA INVENCION

La presente invención se refiere a un método para estimación de diferencia inter-canal (ICD) y un dispositivo de codificación de audio espacial o de codificación multicanal paramétrico, en particular, para la codificación de audio de multicanal paramétrico.

10

## ANTECEDENTES DE LA INVENCION

La codificación de audio multicanal paramétrica se describe en Faller, C., Baumgarte, F.: "Representación eficiente de audio espacial utilizando una parametrización perceptual", Proc. IEEE Workshop on Appl. of Sig. Proc. to Audio and Acoust., octubre 2001, páginas 199-202. Señales de audio de mezcla reductora pueden mezclarse de forma ascendente para sintetizar señales de audio multicanal, utilizando pistas espaciales para generar más canales de audio de salida que las señales de audio de mezcla reductora. En condiciones normales, las señales de audio de mezcla reductora se generan por superposición de una pluralidad de señales de canal de audio de una señal de audio multicanal, a modo de ejemplo, una señal de audio estéreo. Las señales de audio de mezcla reductora están codificadas en forma de onda y se introducen en un flujo de bits de audio junto con datos auxiliares en relación con las denominadas pistas espaciales. El decodificador utiliza los datos auxiliares para sintetizar las señales de audio multicanal sobre la base de los canales de audio codificados en forma de onda.

15

20

25

30

Existen varias pistas espaciales o parámetros que pueden utilizarse para sintetizar señales de audio multicanal. En primer lugar, la diferencia de nivel inter-canal (ILD) indica una diferencia entre los niveles de señales de audio en dos canales a compararse. En segundo lugar, la diferencia de tiempo inter-canal (ITD) indica la diferencia en el tiempo de llegada del sonido entre los oídos de un usuario que escucha. El valor de ITD es importante para la localización del sonido, puesto que proporciona una pista para identificar la dirección o ángulo de incidencia de la fuente de sonido en relación con los oídos de la persona que escucha. En tercer lugar, la diferencia de fase inter-canal (ICD) especifica la diferencia de fase relativa entre los dos canales que se van a comparar. Un valor de ICD de sub-banda puede utilizarse como una estimación del valor de ITD de sub-banda. Por último, la coherencia inter-canal (ICC) se define como la correlación cruzada inter-canal normalizada después de una alineación de fase en conformidad con los valores de ITD o ICD. El valor de ICC puede utilizarse para estimar la anchura de una fuente de sonido.

35

40

Los valores de ILD, ITD, ICD e ICC son parámetros importantes para la codificación/decodificación multicanal espacial, en particular para señales de audio estéreo y especialmente las señales de audio binaurales. La diferencia ITD puede cubrir, a modo de ejemplo, la gama de retardos audibles entre -1.5 ms a 1.5 ms. La diferencia ICD puede cubrir la gama completa de diferencia de fase entre  $-\pi$  y  $\pi$ . ICC puede cubrir la gama de correlación y puede especificarse en un valor porcentual entre 0 y 1 u otros factores de correlación entre -1 y +1. En sistemas de codificación estéreo paramétricos actuales, ILD, ITD, ICD e ICC se suelen estimar en el dominio de la frecuencia. Para cada sub-banda, ILD, ITD, ICD e ICC se calculan, cuantifican, se incluyen en la sección de parámetros de un flujo de bits de audio y son objeto de transmisión.

45

50

Debido a las limitaciones en tasas de bits para sistemas de codificación de audio paramétricos, no existen, a veces, suficientes bits en la sección de parámetros del flujo de bits de audio para transmitir la totalidad de los valores de los parámetros de codificación espaciales. A modo de ejemplo, el documento US 2006/0153408 A1 da a conocer un codificador de audio en donde se generan códigos de pistas combinados para una pluralidad de canales de audio a incluirse como información secundaria en un flujo de bits de audio de mezcla reductora. El documento US 8,054,981 B2 da a conocer un método para la codificación de audio espacial utilizando una regla de cuantización asociada con la relación de niveles de una medida de energía de un canal de audio y la medida de energía de una pluralidad de canales de audio. El documento US 2011/0046964 da a conocer un método para determinar las pistas espaciales sobre la base de un valor medio de las pistas espaciales para diferentes sub-bandas.

55

## SUMARIO DE LA INVENCION

Una idea inventiva de la presente invención es calcular los valores de diferencia inter-canal, ICD, para cada sub-banda de frecuencias o contenedor de frecuencias entre cada par de una pluralidad de señales de canal de audio y para calcular un valor medio ponderado sobre la base de los valores de ICD. Dependiendo del sistema de ponderación, las sub-bandas de frecuencia perceptualmente importantes o los contenedores de frecuencias se tienen en cuenta con una más alta prioridad que las menos importantes.

60

En una forma de realización preferida, la importancia perceptual o de la energía se tiene en cuenta con esta técnica, de modo que el sonido ambiental o el sonido difuso no afectarán a la estimación de ICD. Esto es especialmente conveniente para representar, de forma significativa, la imagen espacial de sonidos que tienen una componente directa fuerte tal como datos de audio de voz.

65

Además, el método propuesto reduce el número de parámetros de codificación espacial a incluirse en un flujo de bits de audio, con lo que se reduce la complejidad de la estimación y la tasa binaria de transmisión.

5 En consecuencia, un primer aspecto de la presente invención se refiere a un método para la estimación de diferencias inter-canal, ICD, comprendiendo dicho método la aplicación de una transformación desde un dominio temporal a un dominio frecuencial para una pluralidad de señales de canal de audio, calculando una pluralidad de valores de ICD para las diferencias ICD entre al menos una de la pluralidad de señales de canal de audio y una señal de canal de audio de referencia a través de una gama de frecuencias predeterminada, siendo cada valor de ICD calculado sobre una parte de la gama de frecuencias predeterminada, calculando, para cada uno de la pluralidad de valores de ICD, un valor de ICD ponderado multiplicando cada uno de la pluralidad de valores de ICD con un factor de ponderación dependiente de la frecuencia correspondiente y calculando un valor de la gama ICD para la gama de frecuencias predeterminada añadiendo la pluralidad de valores de ICD ponderados.

15 Según una primera puesta en práctica del primer aspecto de la idea inventiva, las diferencias ICDs son diferencias de fase inter-canal, IPD, o diferencias de tiempo inter-canal, ITD. Estos parámetros de codificación espacial son particularmente ventajosos para la reproducción de datos de audio para el oído humano.

20 Según una segunda puesta en práctica del primer aspecto de la idea inventiva, la transformación desde un dominio temporal a un dominio frecuencial comprende uno de entre el grupo de Transformación de Fourier Rápida, FFT, banco de filtros de modulación cosenoidal, Transformación de Fourier Discreta, DFT y banco de filtros complejo.

25 Según una tercera puesta en práctica del primer aspecto de la idea inventiva, la gama de frecuencias predeterminada comprende uno de entre el grupo de una banda de frecuencia completa de la pluralidad de señales de canal de audio, un intervalo de frecuencia predeterminado dentro de la banda de frecuencias completa de la pluralidad de señales de canal de audio y una pluralidad de intervalos de frecuencia predeterminados dentro de la banda de frecuencias completa de la pluralidad de señales de canal de audio.

30 En conformidad con una primera puesta en práctica de la tercera puesta en práctica del primer aspecto de la idea inventiva, el intervalo de frecuencias predeterminado está ubicado entre 200 Hz y 600 Hz o entre 300 Hz y 1.5 kHz. Estas gamas de frecuencias corresponden con la sensibilidad dependiente de la frecuencia del oído humano, en donde los parámetros de ICD tienen mayor importancia.

35 En conformidad con una cuarta puesta en práctica del primer aspecto de la idea inventiva, la señal de canal de audio de referencia comprende una de las señales de canal de audio o una señal de audio de mezcla reductora derivada de al menos dos señales de canales de audio de la pluralidad de señales de canal de audio.

40 En conformidad con una quinta puesta en práctica del primer aspecto de la idea inventiva, el cálculo de la pluralidad de los valores de ICD comprende el cálculo de la pluralidad de valores de ICD sobre la base de sub-bandas de frecuencias.

45 En conformidad con una primera puesta en práctica de la quinta puesta en práctica del primer aspecto de la idea inventiva, los factores de ponderación dependientes de la frecuencia se determinan sobre la base de la energía de las sub-bandas de frecuencia normalizadas sobre la base de la energía global a través de la gama de frecuencias predeterminada.

50 En conformidad con una segunda puesta en práctica de la quinta puesta en práctica del primer aspecto de la idea inventiva, los factores de ponderación dependientes de la frecuencia se determinan sobre la base de una curva de enmascaramiento para la distribución de energía de las frecuencias de las señales de canal de audio normalizadas a través de la gama de frecuencias predeterminada.

55 En conformidad con una tercera puesta en práctica de la quinta puesta en práctica del primer aspecto de la idea inventiva, los factores de ponderación dependientes de la frecuencia se determinan sobre la base de los valores de entropía perceptual de las sub-bandas de las señales de canal de audio normalizadas a través de una gama de frecuencias predeterminada.

60 En conformidad con una sexta puesta en práctica del primer aspecto de la idea inventiva, los factores de ponderación dependientes de la frecuencia son atenuados entre al menos dos tramas consecutivas. Lo que antecede puede ser conveniente puesto que los valores de ICD estimados son relativamente estables entre tramas consecutivas debido a que la imagen en estéreo no suele cambiar mucho durante un periodo de tiempo corto.

65 En conformidad con un segundo aspecto de la idea inventiva, un dispositivo de codificación de audio espacial comprende un módulo de transformación configurado para aplicar una transformación desde un dominio temporal a un dominio frecuencial para una pluralidad de señales de canal de audio y un módulo de estimación de parámetros configurado para calcular una pluralidad de valores de ICD para las diferencias ICDs entre al menos una de la pluralidad de señales de canal de audio y una señal de canal de audio de referencia a través de una gama de frecuencias predeterminada, para calcular, para cada uno de la pluralidad de valores de ICD, un valor de ICD

ponderado multiplicando cada uno de la pluralidad de valores de ICD con un factor de ponderación dependiente de la frecuencia correspondiente y para calcular un valor de la gama de ICD para la gama de frecuencias predeterminada añadiendo la pluralidad de valores de ICD ponderados.

5 En conformidad con una primera puesta en práctica del segundo aspecto de la idea inventiva, el dispositivo de codificación de audio espacial comprende, además, un módulo de mezcla reductora configurado para generar una señal de canal de audio de mezcla reductora mezclando en sentido descendente la pluralidad de señales de canal de audio.

10 En conformidad con una segunda puesta en práctica del segundo aspecto de la idea inventiva, el dispositivo de codificación de audio espacial comprende, además, un módulo de codificación acoplado al módulo de mezcla reductora y configurado para generar un flujo de bits de audio codificado que comprende el flujo de bits de audio de mezcla descendente codificado.

15 En conformidad con una tercera puesta en práctica del segundo aspecto de la idea inventiva, el dispositivo de codificación de audio espacial comprende, además, un módulo de flujos de bits acoplado al módulo de estimación de parámetros y configurado para generar un flujo de bits de audio que comprende un flujo de bits de audio de mezcla descendente y datos auxiliares que comprenden valores de la gama de ICD para la pluralidad de señales de canal de audio.

20 En conformidad con una primera puesta en práctica de la tercera puesta en práctica del segundo aspecto de la idea inventiva, el módulo de difusión de flujo continuo está configurado, además, para establecer un indicador en el flujo de bits de audio, indicando dicho indicador la presencia de datos auxiliares que comprenden los valores de la gama de ICD en el flujo de bits de audio.

25 En conformidad con una cuarta puesta en práctica del segundo aspecto de la idea inventiva, el indicador se establece para el flujo de bits de audio completo o se incluye en los datos auxiliares comprendidos en el flujo de bits de audio.

30 Según un tercer aspecto de la idea inventiva de la presente invención, se proporciona un programa informático, comprendiendo dicho programa informático un código de programa para realizar el método según el primer aspecto de cualquiera de sus puestas en práctica cuando se ejecuta en un ordenador.

35 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 circuitos de hardware dentro de un circuito integrado específico de la aplicación (ASIC).

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

45 Formas de realización y puestas en práctica adicionales pueden entenderse fácilmente a partir de la descripción siguiente. En particular, cualesquiera características de las formas de realización, aspectos y puestas en práctica de la idea inventiva según se establecen a continuación pueden combinarse con cualesquiera otras características de las formas de realización, aspectos y puestas en práctica, a no ser que se indique concretamente de otro modo.

#### BREVE DESCRIPCIÓN DE LOS DIBUJOS

50 Los dibujos adjuntos están incluidos para proporcionar un mejor entendimiento de la idea inventiva. Ilustran formas de realización y pueden ayudar a explicar los principios de la invención en conjunción con la descripción. Otras formas de realización y muchas de las ventajas previstas, principios contemplados y funcionalidades se apreciarán mejor haciendo referencia a la descripción detallada dada a continuación. Los elementos de los dibujos no están necesariamente dibujados a escala en relación entre sí. En general, las referencias numéricas similares designan partes similares correspondientes.

55 La Figura 1 ilustra, de forma esquemática, un sistema de codificación de audio espacial.

La Figura 2 ilustra, de forma esquemática, un dispositivo de codificación de audio espacial.

60 La Figura 3 ilustra, de forma esquemática, un dispositivo de decodificación de audio espacial.

La Figura 4 ilustra, de forma esquemática una forma de realización de un método para la estimación de diferencias inter-canal.

65 La Figura 5 ilustra, de forma esquemática, una variante de una estructura de flujo de bits para un flujo de bits de audio.

## DESCRIPCIÓN DETALLADA DE LA INVENCÓN

En la descripción detallada siguiente, se hace referencia a los dibujos adjuntos y en donde a modo de instrucción, se ilustran formas de realización específicas. Debe ser evidente que otras formas de realización pueden utilizarse y se pueden efectuar cambios estructurales o lógicos sin desviarse por ello del alcance de protección de la presente invención. A no ser que se indique concretamente de otro modo, las funciones, los principios y los detalles de cada forma de realización pueden combinarse con otras formas de realización. En general, esta aplicación está prevista para cubrir cualesquiera adaptaciones o variaciones de las formas de realización específicas aquí descritas. Por consiguiente, la siguiente descripción detallada no ha de tomarse en un sentido de limitación y el alcance de la presente invención se define por las reivindicaciones adjuntas.

Las formas de realización pueden incluir métodos y procesos que pueden materializarse dentro de instrucciones legibles por máquina proporcionadas por un soporte legible por máquina, incluyendo dicho soporte legible por máquina, sin limitación, dispositivos, aparatos, mecanismos o sistemas que son capaces de memorizar información que puede ser accesible para una máquina tal como un ordenador, un dispositivo de cálculo, una unidad de procesamiento, un dispositivo de gestión de redes, un ordenador portátil, un microprocesador o dispositivos similares. El soporte legible por máquina puede incluir soportes volátiles o no volátiles así como señales propagadas de cualquier forma tales como señales eléctricas, señales digitales, señales lógicas, señales ópticas, señales acústicas, señales acusto-ópticas o similares, siendo dichos soportes capaces de transmitir información a una máquina.

A continuación, se hace referencia a método y etapas de métodos, que se ilustran, de forma esquemática y a modo de ejemplo, en diagramas de flujo y diagramas de bloques. Debe entenderse que los métodos descritos en conjunción con dichos dibujos ilustrativos pueden realizarse fácilmente mediante también por formas de realización de sistemas, aparatos y/o dispositivos. En particular, debe ser evidente que los sistemas, aparatos y/o dispositivos capaces de realizar los diagramas de bloques detallados y/o los diagramas de flujo no están necesariamente limitados a los sistemas, aparatos y/o dispositivos ilustrados y detallados a continuación, sino que pueden ser también diferentes sistemas, aparatos y/o dispositivos. Los términos “primero”, “segundo”, “tercero” se utilizan simplemente como etiquetas y no están previstos para imponer exigencias numéricas sobre sus objetos o para establecer una determinada clasificación de importancia de sus objetos.

La Figura 1 ilustra, de forma esquemática, un sistema de codificación de audio espacial 100. El sistema de codificación de audio espacial 100 comprende un dispositivo de codificación de audio espacial 10 y un dispositivo de decodificación de audio espacial 20. Una pluralidad de señales de canal de audio 10a, 10b de las cuales solamente dos son a modo de ejemplo ilustrados en la Figura 1, se consideran como entrada al dispositivo de codificación de audio espacial 10. El dispositivo de codificación de audio espacial 10 codifica y efectúa una mezcla reductora de las señales de canal de audio 10a, 10b y genera un flujo de bits de audio 1 que se transmite al dispositivo de decodificación de audio espacial 20. El dispositivo de decodificación de audio espacial 20 decodifica y mezcla de forma ascendente los datos de audio incluidos en el flujo de bits de audio 1 y genera una pluralidad de señales de canal de audio de salida 20a, 20b de las que solamente dos son a modo de ejemplo ilustradas en la Figura 1. El número de señales de canal de audio 10a, 10b y 20a, 20b, respectivamente, está en principio no limitado. A modo de ejemplo, el número de señales de canal de audio 10a, 10b y 20a, 20b pueden ser dos para señales de estéreo binaurales. A modo de ejemplo, las señales de estéreo binaurales pueden utilizarse para señales de audio en 3D o la presentación circundante sobre la base de auriculares, a modo de ejemplo, con filtrado de HRTF.

El sistema de codificación de audio espacial 100 puede aplicarse para la codificación de la extensión en estéreo de ITU-T G.722, G.722 Anexo B, G.711.1 y/o G.711.1 Anexo D. Además, el sistema de codificación de audio espacial 100 puede utilizarse para la codificación/decodificación de voz y de audio en aplicaciones móviles, tal como se definen en el codificador-decodificador (*codec*) de 3GPP EVS (Servicios de Voz Mejorada).

La Figura 2 ilustra, de forma esquemática, el dispositivo de codificación de audio espacial 10 ilustrado en la Figura 1 con mayor detalle. El dispositivo de codificación de audio espacial 10 puede comprender un módulo de transformación 15, un módulo de estimación de parámetros 11 acoplado al módulo de transformación 15, un módulo de mezcla reductora 12 acoplado al módulo de transformación 15, un módulo de codificación 13 acoplado al módulo de mezcla reductora 12 y un módulo de flujo continuo 14 acoplado al módulo de codificación 13 y al módulo de estimación de parámetros 11.

El módulo de transformación 15 puede estar configurado para aplicar una transformación desde un dominio temporal a un dominio frecuencial para una pluralidad de señales de canal de audio 10a, 10b que son objeto de entrada al dispositivo de codificación espacial 10. El módulo de mezcla reductora 12 puede configurarse para recibir las señales de canal de audio transformadas 10a, 10b desde el módulo de transformación 15 y para generar al menos una señal de canal de audio de mezcla reductora efectuando una mezcla de la pluralidad de señales de canal de audio transformadas 10a, 10b. El número de señales de canal de audio objeto de mezcla reductora puede ser, a modo de ejemplo, menor que el número de señales de canal de audio transformadas 10a, 10b. A modo de ejemplo, el módulo de mezcla reductora 12 puede configurarse para generar solamente una señal de canal de audio de mezcla reductora. El módulo de codificación 13 puede configurarse para recibir las señales de canal de audio de

mezcla reductora y para generar un flujo de bits de audio codificado que comprende las señales de canal de audio de mezcla reductora codificadas.

El módulo de estimación de parámetros 11 puede configurarse para recibir la pluralidad de señales de canal de audio 10a, 10b como entrada y para calcular una pluralidad de valores de diferencia inter-canal, ICD, para las diferencias ICDs entre al menos una de la pluralidad de señales de canal de audio 10a, 10b y una señal de canal de audio de referencia a través de una gama de frecuencias predeterminada. La señal de canal de audio de referencia puede ser, a modo de ejemplo, una de la pluralidad de señales de canal de audio 10a, 10b. Como alternativa, puede ser posible utilizar una señal de audio de mezcla reductora derivada de al menos dos señales de canal de audio de la pluralidad de señales de canal de audio 10a, 10b. El módulo de estimación de parámetros 11 puede configurarse, además, para calcular, para cada uno de la pluralidad de valores de ICD, un valor de ICD ponderado multiplicando cada uno de la pluralidad de valores de ICD con un factor de ponderación dependiente de la frecuencia correspondiente y para calcular un valor de la gama ICD para la gama de frecuencias predeterminada añadiendo la pluralidad de valores de ICD ponderados.

El valor de la gama ICD puede ser luego objeto de entrada al módulo de flujo continuo 14 que puede configurarse para generar el flujo de bits de audio de salida 1 que comprende el flujo de bits de audio codificado procedimiento del módulo de codificación 13 y una sección de parámetros que comprende una representación cuantificada del valor de la gama de ICD. El módulo de flujo continuo 14 puede configurarse, además, para establecer un indicador del tipo de parámetro en la sección de parámetros del flujo de bits de audio 1 que indica el tipo de valor de gama ICD que está incluido en el flujo de bits de audio 1. Además, el módulo de flujo continuo 14 puede estar configurado, además, para establecer un indicador en el flujo de bits de audio 1, indicando dicho indicador la presencia del valor de la gama ICD en la sección de parámetros del flujo de bits de audio 1. Este indicador puede establecerse para el flujo de bits de audio completo 1 o incluirse en la sección de parámetros del flujo de bits de audio 1. De este modo, la señalización del valor de la gama de ICD, que está incluido en el flujo de bits de audio 1, puede señalizarse explícitamente, o de forma implícita, para el dispositivo de decodificación de audio espacial 20. Puede ser posible conmutar entre los sistemas de señalización explícito e implícito.

En el caso de señalización implícita, el indicador puede indicar la presencia de la información de canal secundario en los datos auxiliares en la sección de parámetros. Un dispositivo de decodificación de legado 20 no comprueba si dicho indicador está presente y por ello, solamente decodifica el flujo de bits de audio de mezcla reductora codificado. Por otro lado, un dispositivo de decodificación de actualización, esto es, no de legado 20 puede comprobar la presencia de dicho indicador en el flujo de bits de audio recibido 1 y reconstruir la señal de audio multicanal 20a, 20b sobre la base de los parámetros de codificación espaciales de banda completa adicionales, esto es, el valor de la gama de ICD incluido en la sección de parámetros del flujo de bits de audio 1.

Cuando se utiliza la señalización explícita, el flujo de bits de audio completo 1 puede indicarse como conteniendo un valor de la gama de ICD. De esa forma, un dispositivo de decodificación de legado 20 no es capaz de decodificar el flujo de bits y por ello, descarta el flujo de bits de audio 1. Por otro lado, un dispositivo de decodificación de actualización 20 puede decidir sobre si decodificar el flujo de bits de audio 1 como un conjunto total o solamente decodificar el flujo de bits de audio de mezcla reductora codificado 1 mientras se desprecia operativamente el valor de la gama de ICD. La ventaja de la señalización explícita puede considerarse por cuanto que, a modo de ejemplo, un nuevo terminal móvil puede decidir qué partes de un flujo de bits de audio decodificar con el fin de ahorrar energía y de este modo, prolongar la vida útil de la batería de una batería integrada. La decodificación de los parámetros de codificación espacial suele ser más compleja y requiere más energía. Además, dependiendo del sistema de presentación, el dispositivo de decodificación de actualización 20 puede decidir qué parte del flujo de bits de audio 1 debe decodificarse. A modo de ejemplo, para la presentación con auriculares puede ser suficiente decodificar solamente el flujo de bits de audio de mezcla reductora codificado, mientras que la señal de audio multicanal se decodifica solamente cuando el terminal móvil está conectado a una estación de acoplamiento con dicha capacidad de presentación multicanal.

La Figura 3 ilustra, de forma esquemática, el dispositivo de decodificación de audio espacial 20 ilustrado en la Figura 1 con mayor detalle. El dispositivo de decodificación de audio espacial 20 puede comprender un módulo de extracción de flujos de bits 26 un módulo de extensión de parámetros 21, un módulo de decodificación 22, un módulo de mezcla en sentido ascendente 24 y un módulo de transformación 25. El módulo de extracción de flujo de bits 26 puede configurarse para recibir un flujo de bits de audio 1 y separar la sección de parámetros y el flujo de bits de audio de mezcla reductora codificado encerrado en el flujo de bits de audio 1. El módulo de extracción de parámetros 21 puede configurarse para detectar un indicador del tipo de parámetro en la sección de parámetros de un flujo de bits de audio recibido 1 que indica que un valor de la gama ICD está incluido en el flujo de bits de audio 1. El módulo de extracción de parámetros 21 puede configurarse, además, para la lectura del valor de la gama ICD desde la sección de parámetros del flujo de bits de audio recibido 1.

El módulo de decodificación 22 puede configurarse para decodificar el flujo de bits de audio de mezcla reductora codificado y para introducir la señal de audio de mezcla reductora decodificada en el módulo de mezcla en sentido ascendente 24. El módulo de mezcla en sentido ascendente 24 puede acoplarse al módulo de extracción de parámetros 21 y configurarse para la mezcla en sentido ascendente de la señal de audio de mezcla reductora

5 decodificada para una pluralidad de señales de canal de audio utilizando el valor de la gama ICD objeto de lectura desde la sección de parámetros del flujo de bits de audio recibido 1 según se proporciona por el módulo de extracción de parámetros 21. Por último, el módulo de transformación 25 puede acoplarse al módulo de mezcla en sentido ascendente 24 y configurarse para transformar la pluralidad de señales de canal de audio desde un dominio

10 La Figura 4 ilustra, de forma esquemática, una forma de realización de un método 30 para la codificación espacial paramétrica. El método 30 comprende, en una primera etapa, la realización de una transformación de tiempo-frecuencia en canales de entrada, a modo de ejemplo, los canales de entrada 10a, 10b. En caso de una señal estéreo, se realiza una primera transformación en la etapa 30a y una segunda transformación se realiza en la etapa 30b. La transformación puede, en cada caso, realizarse utilizando la transformación de Fourier rápida (FFT). Como alternativa, puede realizarse la transformación de Fourier a corto plazo (STFT), el filtrado de modulación cosenoidal con un banco de filtros con modulación cosenoidal o un filtrado complejo con un banco de filtros complejos.

15 En una segunda etapa 31, puede calcularse un espectro cruzado  $c[b]$  por sub-banda  $b$  como

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

20 en donde  $X_1[k]$  y  $X_2[k]$  son los coeficientes de FFT de los dos canales 1 y 2, a modo de ejemplo, el canal izquierdo y el canal derecho en caso de estéreo. El asterisco "\*" indica la conjugación compleja.  $k_b$  indica el bit de inicio de la sub-banda  $b$  y  $k_{b+1}$  indica el bit de inicio de la sub-banda próxima  $b+1$ . Por consiguiente, los contenedores de frecuencias  $[k]$  de FFT desde  $k_b$  a  $k_{b+1}$  representan la sub-banda  $b$ .

25 Como alternativa, el espectro cruzado puede calcularse para cada contenedor de frecuencia  $k$  de FFT. En este caso, la sub-banda  $b$  corresponde directamente a un contenedor de frecuencias  $[k]$ .

En una tercera etapa 32, pueden calcularse las diferencias inter-canal por sub-banda  $b$  sobre la base del espectro cruzado. A modo de ejemplo, en caso de la diferencia de fase interaural, IPD, dicho cálculo puede realizarse como

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

30 en donde el valor de IPD por sub-banda  $b$  es el ángulo del espectro cruzado  $c[b]$  de la respectiva sub-banda  $b$ . Las etapas 31 y 32 aseguran que una pluralidad de valores de ICD, en particular valores de IPD, para los ICDs/IPDs entre al menos una de la pluralidad de señales de canal de audio y una señal de canal de audio de referencia, a través de una gama de frecuencias predeterminada, son objeto de cálculo. Además, cada valor de ICD se calcula sobre una parte de la gama de frecuencias predeterminada, que es una sub-banda de frecuencias  $b$  o al menos un contenedor de frecuencias único.

35 El sistema de cálculo según se detalla con respecto a las etapas 31 y 32 corresponde al método según se conoce a partir de los documentos de Breebart, J., van de Par, S. Kohlrausch, A., Schuijers, E.: "Codificación paramétrica de señal de audio estéreo", EURASIP Journal on Applied Signal Processing, 2005, nº 9, páginas 1305-1322.

40 El valor de IPD representa una diferencia de fase para una señal limitada en banda. Si el ancho de banda está suficientemente limitado, esta diferencia de fase puede considerarse como un retardo fraccionario entre las señales de entrada. Para cada sub-banda de frecuencia  $b$ , las diferencias IPD y las diferencias de tiempo inter-canal, ITD, representan la misma información. No obstante, para el banco completo, el valor de IPD difiere del valor de ITD: el valor IPD de banda completa es la diferencia de fase constante entre dos canales 1 y 2, mientras que el valor ITD de banda completa es la diferencia de tiempo constante entre dos canales.

45 Con el fin de calcular el valor de IPD de banda completa sobre la base del valor de IPD de sub-bandas, podría ser posible calcular el valor medio a través de los valores de IPD de sub-bandas para obtener el valor de IPD de banda completa, es decir, el valor de gama IPD a través de la gama de frecuencias completa de las señales de canal de audio. Sin embargo, este método de estimación puede llevar a una estimación incorrecta de un valor de gama IPD representativo, puesto que las sub-bandas de frecuencia tienen diferente importancia perceptual.

50 Para el cálculo de un valor de gama ICD, una gama de frecuencias predeterminada puede definirse. A modo de ejemplo, la gama de frecuencias predeterminada puede ser la banda de frecuencias completa de la pluralidad de señales de canal de audio. Como alternativa, uno o más intervalos de frecuencia predeterminados dentro de la banda de frecuencias completa de la pluralidad de señales de canal de audio pueden seleccionarse, cuyos intervalos de frecuencias predeterminados pueden ser coherentes o estar separados. La gama de frecuencias predeterminada puede incluir, a modo de ejemplo, la banda de frecuencias entre 200 Hz y 600 Hz o, como alternativa, entre 300 Hz y 1.5 kHz.

55 En una tercera etapa 33 y una cuarta etapa 34, en paralelo con las primera y segunda etapas 31 y 32, la energía

$E[b]$  de cada parte de la gama de frecuencias predeterminada, esto es, cada sub-banda de frecuencia  $b$  o contenedor de frecuencia  $b$  se calcula por

$$E[b] = X_1[b]^2 + X_2[b]^2$$

o como alternativa

5

$$E[b] = \sum_{k=k_b}^{k_{b+1}-1} (X_1[k]^2 + X_2[k]^2),$$

y posteriormente se normaliza sobre la energía  $E_G$  de la gama de frecuencias predeterminada, a modo de ejemplo, la banda completa:

10

$$E_G = \sum_{b=M_{\min}}^{M_{\max}} E[b],$$

en donde  $M_{\min}$  y  $M_{\max}$  son el índice de la sub-banda de frecuencia más baja y más alta o el contenedor de frecuencias dentro de la gama de frecuencias predeterminada, respectivamente.

15

En la etapa 35, para cada uno de la pluralidad de valores de ICD, a modo de ejemplo, los valores  $IPD[b]$ , un valor de ICD ponderado, a modo de ejemplo, un valor de IPD ponderado  $IPD_w[b]$ , se calcula multiplicando cada uno de la pluralidad de valores de ICD con un factor de ponderación independiente de la frecuencia correspondiente  $E_w[b]$ :

$$IPD_w[b] = IPD[b] \cdot E_w[b]$$

20 El factor de ponderación dependiente de la frecuencia puede ser, a modo de ejemplo, un valor de energía ponderado asociado  $E_w[b]$  según se calcula por

$$E_w[b] = E[b] / E_G.$$

Puede ser posible atenuar los factores de ponderación  $E_w[b]$  a través de tramas consecutivas, esto es, teniendo en cuenta una fracción de los factores de ponderación  $E_w[b]$  de tramas anteriores de la pluralidad de señales de canal de audio cuando se calculan los factores de ponderación actuales  $E_w[b]$ .

25

Por último, en la etapa 36, un valor de gama ICD, a modo de ejemplo, un valor IPD de banda completa  $IPD_F$ , puede calcularse para la gama de frecuencias predeterminada añadiendo la pluralidad de valores ICD ponderados:

30

$$IPD_F = \sum_{b=M_{\min}}^{M_{\max}} IPD_w[b]$$

Como alternativa, los factores de ponderación  $E_w[b]$  pueden derivarse a partir de una curva de enmascaramiento para la distribución de energía de las frecuencias de las señales de canal de audio normalizadas a través de la gama de frecuencias predeterminada. Dicha curva de enmascaramiento puede calcularse, a modo de ejemplo, como es conocido a partir del documento de Bosi, M. Goldberg, R.: "Introducción a la codificación de audio digital y sus normas", Kluwer Academic Publishers 2003. También es posible determinar los factores de ponderación dependientes de la frecuencia sobre la base de los valores de entropía perceptual de las sub-bandas  $b$  de las señales de canal de audio normalizadas a través de una gama de frecuencias predeterminada. En ese caso, la versión normalizada de la curva de enmascaramiento o la entropía perceptual puede utilizarse como función de ponderación.

35

40

El método según se ilustra en la Figura 4 puede aplicarse también a la codificación de audio paramétrica multicanal. Un espectro cruzado puede calcularse por sub-banda  $b$  y por cada canal  $j$  como:

45

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

en donde  $X_j[k]$  es el coeficiente de FFT del canal  $j$  y  $X_{ref}[k]$  es el coeficiente FFT de un canal de referencia. El canal de referencia puede ser un canal seleccionado de entre la pluralidad de canales  $j$ . Como alternativa, el canal de referencia puede ser el espectro de una señal de mezcla reductora única, que es la media sobre todos los canales  $j$ . En el caso anterior, se genera  $M-1$  pistas espaciales, mientras que en el último caso, se generan  $M$  pistas espaciales, como  $M$  siendo el número de canales  $j$ . El asterisco "\*" indica la conjugación compleja,  $k_b$  indica el contenedor de inicio de la sub-banda  $b$  y  $k_{b+1}$  indica el contenedor de inicio de la sub-banda próxima  $b+1$ . En consecuencia, los contenedores de frecuencias  $[k]$  de FFT desde  $k_b$  a  $k_{b+1}$  representan la sub-banda  $b$ .

50

Como alternativa, puede calcularse el espectro cruzado para cada contenedor de frecuencias  $k$  de FFT. En este caso, la sub-banda  $b$  corresponde directamente a un contenedor de frecuencias  $[k]$ .

5 Las diferencias inter-canal del canal  $j$  pueden calcularse por sub-banda  $b$  sobre la base del espectro cruzado. A modo de ejemplo, en caso de la diferencia de fase, IPD, dicho cálculo puede realizarse como

$$IPD_j[b] = \angle c_j[b],$$

en donde el valor  $IPD_j$  por sub-banda  $b$  y canal  $j$  es el ángulo del espectro cruzado  $q[b]$  de la respectiva sub-banda  $b$  y canal  $j$ .

10 El valor de  $E_j[b]$  de energía por canal  $j$  de cada parte de la gama de frecuencias predeterminada, esto es, cada sub-banda de frecuencia  $b$  o contenedor de frecuencia  $b$  se calcula mediante

$$E_j[b] = 2 \cdot X_j[b] \cdot X_{ref}[b]$$

o como alternativa

15 
$$E[b] = \sum_{k=k_b}^{k_{b+1}-1} (X_j[k]^2 + X_{ref}[k]^2),$$

y posteriormente, se normaliza sobre el valor  $E_{G_j}$  de energía de la gama de frecuencias predeterminada, a modo de ejemplo, la banda completa:

20 
$$E_{G_j} = \sum_{b=M_{\min}}^{M_{\max}} E_j[b],$$

en donde  $M_{\min}$  y  $M_{\max}$  son el índice de la sub-banda de frecuencia más baja y más alta o el contenedor dentro de la gama de frecuencias predeterminada, respectivamente.

25 Para cada uno de la pluralidad de valores ICD, a modo de ejemplo, los valores  $IPD_j[b]$ , un valor ICD ponderado, a modo de ejemplo, un valor IPD ponderado  $IPD_{w_j}[b]$ , se calcula multiplicando cada uno de la pluralidad de valores ICD con un factor de ponderación dependiente de la frecuencia correspondiente  $E_{w_j}[b]$ :

30 
$$IPD_{w_j}[b] = IPD_j[b] \cdot E_{w_j}[b]$$

El factor de ponderación dependiente de la frecuencia puede ser, a modo de ejemplo, un valor de energía ponderado asociado  $E_{w_j}[b]$  según se calcula por

35 
$$E_{w_j}[b] = E_j[b] / E_{G_j}.$$

Puede ser posible atenuar el valor de los factores de ponderación  $E_{w_j}[b]$  a través de tramas consecutivas, esto es, teniendo en cuenta una fracción de los factores de ponderación  $E_{w_j}[b]$  de tramas anteriores de la pluralidad de señales de canal de audio cuando se calculan los factores de ponderación actuales  $E_{w_j}[b]$ .

40 Por último, un valor de gama ICD, a modo de ejemplo, un valor IPD de banda completa  $IPD_{F_j}$  puede calcularse para la gama de frecuencias predeterminada añadiendo la pluralidad de valores ICD ponderados:

$$IPD_{F_j} = \sum_{b=M_{\min}}^{M_{\max}} IPD_{w_j}[b].$$

45 La Figura 5 ilustra, de forma esquemática, una estructura de flujo de bits de un flujo de bits de audio, a modo de ejemplo, el flujo de bits de audio 1 detallado en las Figuras 1 a 3. En la Figura 5, el flujo de bits de audio 1 puede incluir una sección de flujo de bits de audio de mezcla reductora codificada 1a y una sección de parámetros 1b. La sección de flujo de bits de audio de mezcla reductora codificada 1a y la sección de parámetros 1b pueden alternarse y su longitud combinada puede ser indicativa de la tasa binaria global del flujo de bits de audio 1. La sección de flujo de bits de audio de mezcla reductora codificada 1a puede incluir los datos de audio reales que se van a decodificar.

50 La sección de parámetros 1b puede comprender una o más representaciones cuantificadas de parámetros de codificación espacial tales como el valor de la gama ICD. El flujo de bits de audio 1 puede incluir, a modo de ejemplo, un bit de indicador de señalización 2 utilizado para la señalización explícita si el flujo de bits de audio 1 incluye datos auxiliares en la sección de parámetros 1b o no los incluye. Además, la sección de parámetros 1b puede incluir un bit indicador de señalización 3b utilizado para la señalización implícita si el flujo de bits de audio 1 incluye datos auxiliares en la sección de parámetros 1b o no lo incluye.

55

## REIVINDICACIONES

1. Un método (30) para la estimación de diferencias inter-canal, ICD, que comprende:
- 5 aplicar (30a, 30b) una transformación desde un dominio temporal a un dominio frecuencial para una pluralidad de señales de canal de audio;
- calcular (31, 32) una pluralidad de valores de ICD para las diferencias ICD entre al menos una de la pluralidad de señales de canal de audio y una señal de canal de audio de referencia en una gama de frecuencias predeterminada, calculándose cada valor de ICD en una parte de la gama de frecuencias predeterminada;
- 10 calcular (35), para cada uno de la pluralidad de valores de ICD, un valor de ICD ponderado multiplicando cada uno de la pluralidad de valores de ICD con factor de ponderación que depende de la frecuencia correspondiente; y
- 15 calcular (36) un valor de gama de ICD para la gama de frecuencias predeterminada añadiendo la pluralidad de valores de ICD ponderados.
2. El método (30) según la reivindicación 1, en donde los valores de ICD son diferencias de fase inter-canal, IPD, o diferencias de tiempo inter-canal, ITD.
- 20 3. El método (30) según una de las reivindicaciones 1 o 2, en donde la transformación desde un dominio temporal a un dominio frecuencial comprende una de entre el grupo de una Transformación de Fourier Rápida, FFT, una batería de filtros de modulación cosenoidal, una Transformación de Fourier Discreta, DFT y una batería de filtros complejos.
- 25 4. El método (30) según una de las reivindicaciones 1 a 3, en donde la gama de frecuencias predeterminada comprende una gama del grupo de una banda de frecuencias completa de la pluralidad de señales de canal de audio, un intervalo de frecuencias predeterminado dentro de la banda de frecuencias completa de la pluralidad de señales de canal de audio y una pluralidad de intervalos de frecuencias predeterminados dentro de la banda de frecuencias completa de la pluralidad de señales de canal de audio .
- 30 5. El método (30) según la reivindicación 4, en donde el intervalo de frecuencias predeterminado se sitúa entre 200 Hz y 600 Hz o entre 300 Hz y 1.5 kHz.
- 35 6. El método (30) según una de las reivindicaciones 1 a 5, en donde la señal de canal de audio de referencia comprende una de las señales de canal de audio o una señal de audio que haya sufrido una mezcla reductora derivada de al menos dos señales de canal de audio de la pluralidad de señales de canal de audio.
- 40 7. El método (30) según una de las reivindicaciones 1 a 6, en donde calcular la pluralidad de valores de ICD comprende calcular la pluralidad de valores de ICD sobre la base de sub-bandas de frecuencias.
8. El método (30) según la reivindicación 7, en donde los factores de ponderación dependientes de la frecuencia se determinan sobre la base de la energía de las sub-bandas de frecuencia normalizadas sobre la base de la energía global en la gama de frecuencias predeterminada.
- 45 9. El método (30) según la reivindicación 7, en donde los factores de ponderación dependientes de la frecuencia se determinan sobre la base de una curva de enmascaramiento para la distribución de la energía de las frecuencias de las señales de canal de audio normalizadas a través de la gama de frecuencias predeterminada.
- 50 10. El método (30) según la reivindicación 7, en donde los factores de ponderación dependientes de la frecuencia se determinan sobre la base de valores de entropía perceptual de las sub-bandas de las señales de canal de audio normalizadas en la gama de frecuencias predeterminada.
- 55 11. El método (30) según una de las reivindicaciones 1 a 10, en donde los factores de ponderación dependientes de la frecuencia son atenuados entre al menos dos tramas consecutivas.
12. Un dispositivo de codificación de audio espacial (10), que comprende:
- 60 un módulo de transformación (15) configurado para aplicar una transformación desde un dominio temporal a un dominio frecuencial para una pluralidad de señales de canal de audio (10a; 10b); y
- un módulo de estimación de parámetros (11) configurado para calcular una pluralidad de valores de ICD para las diferencias ICDs entre al menos una de la pluralidad de señales de canal de audio (10a; 10b) y una señal de canal de audio de referencia a través de una gama de frecuencias predeterminada, para calcular, para cada uno de la pluralidad de valores de ICD, un valor de ICD ponderado multiplicando cada uno de la pluralidad de valores de ICD con un factor de ponderación dependiente de la frecuencia correspondiente y para calcular un valor de la gama ICD
- 65

para la gama de frecuencias predeterminada añadiendo la pluralidad de valores de ICD ponderados.

**13.** El dispositivo de codificación de audio espacial (10) según la reivindicación 12, que comprende, además:

5 un módulo de mezcla reductora (12) configurado para generar una señal de canal de audio reductora mediante una mezcla reductora de la pluralidad de señales de datos del canal de audio (10a; 10b).

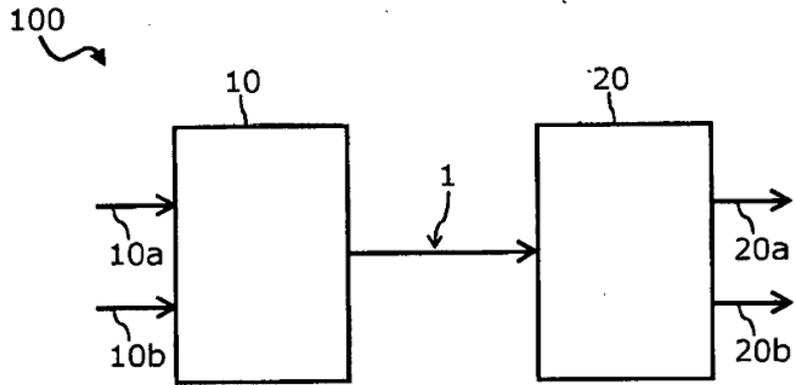
**14.** El dispositivo de codificación de audio espacial (10) según la reivindicación 13 que comprende, además:

10 un módulo de codificación (13) acoplado al módulo de mezcla reductora (12) y configurado para generar un flujo de bits de audio codificado que comprende el flujo de bits de audio de mezcla reductora codificado.

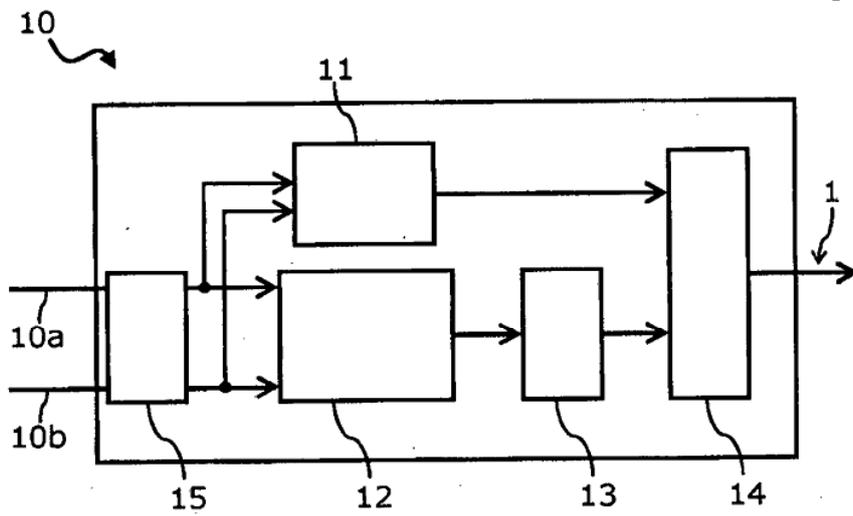
**15.** El dispositivo de codificación de audio espacial (10) según una de las reivindicaciones 12 a 14 que comprende, además:

15 un módulo de difusión en flujo continuo (14) acoplado al módulo de estimación de parámetros (11) y configurado para generar un flujo de bits de audio (1) que comprende un flujo continuo de bits de audio de mezcla reductora y datos auxiliares que comprenden los valores de la gama de ICD para la pluralidad de señales de canal de audio (10a; 10b).

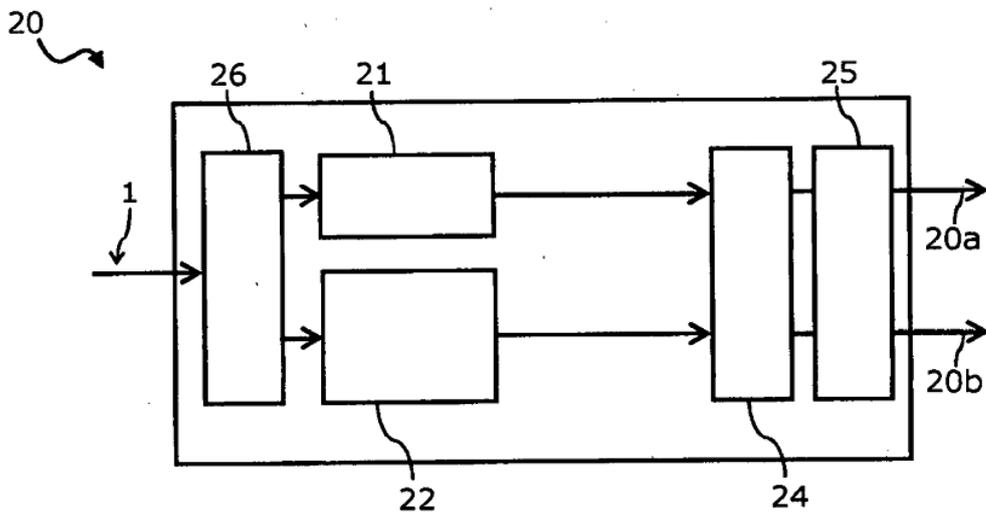
20 **16.** Un programa informático para ejecutar el método según una de las reivindicaciones 1 a 11.



**Fig. 1**



**Fig. 2**



**Fig. 3**

