

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 773 795**

51 Int. Cl.:

**G10L 19/008** (2013.01)

**G10L 19/028** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.02.2017 PCT/EP2017/053272**

87 Fecha y número de publicación internacional: **24.08.2017 WO17140666**

96 Fecha de presentación y número de la solicitud europea: **14.02.2017 E 17704485 (6)**

97 Fecha y número de publicación de la concesión europea: **25.12.2019 EP 3417452**

54 Título: **Aparato y método para llenado estéreo en codificación multicanal**

30 Prioridad:

**17.02.2016 EP 16156209**

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

**14.07.2020**

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR  
FÖRDERUNG DER ANGEWANDTEN  
FORSCHUNG E.V. (100.0%)  
Hansastraße 27c  
80686 München, DE**

72 Inventor/es:

**DICK, SASCHA;  
HELMRICH, CHRISTIAN;  
RETTTELBACH, NIKOLAUS;  
SCHUH, FLORIAN;  
FUEG, RICHARD y  
NAGEL, FREDERIK**

74 Agente/Representante:

**ARIZTI ACHA, Monica**

ES 2 773 795 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

Aparato y método para llenado estéreo en codificación multicanal

5 La presente invención se refiere a codificación de señales de audio y, en particular, a un aparato y método para llenado estéreo en la codificación multicanal.

La codificación de audio es el dominio de compresión que trata de explotar la redundancia e irrelevancia en señales de audio.

10 En MPEG USAC (véase, por ejemplo, [3]), la codificación estereofónica conjunta de dos canales se lleva a cabo usando predicción compleja, MPS 2–1–2 o estéreo unificado con señales residuales limitadas de banda o de banda completa. MPEG envolvente (véase, por ejemplo, [4]) combina jerárquicamente cajas One–To–Two (OTT) y Two–To–Three (TTT) para codificación conjunta de audio multicanal con o sin transmisión de señales residuales.

15 En MPEG–H, unos elementos de canal cuádruple aplican jerárquicamente cajas estéreo MPS 2–1–2 seguido por predicción compleja/cajas estéreo MS que forman un árbol de remezcla 4x4 fijo (véase, por ejemplo, [1]).

20 AC4 (véase, por ejemplo, [6]) introduce nuevos elementos de 3, 4 y 5 canales que permiten la remezcla de canales transmitidos a través de una matriz de mezcla transmitida y posterior información de codificación estereofónica conjunta. Por otra parte, las publicaciones anteriores sugieren el uso de transformadas ortogonales tales como la transformada de Karhunen–Loeve (KLT) para mayor codificación multicanal de audio (véase, por ejemplo, [7]).

25 Por ejemplo, en el contexto de audio 3D, los canales de autoparlantes se distribuyen en varias capas de altura, que resulta en pares de canales horizontales y verticales. La codificación conjunta de solo dos canales como se definen en USAC no es suficiente para considerar las relaciones espaciales y perceptuales entre los canales. MPEG envolvente se aplica en una etapa adicional de pre–/posprocesamiento, las señales residuales se transmiten en forma individual sin la posibilidad de codificación estereofónica conjunta, por ejemplo, para explotar las dependencias entre las señales residuales verticales izquierdas y derechas. En AC–4, se introducen elementos de canales N dedicados que permiten una codificación eficaz de parámetros de codificación conjunta, pero falla para las configuraciones genéricas de parlantes con mayores canales que los propuestos para nuevos escenarios de reproducción inmersiva (7.1+4, 22.2). El elemento del canal cuádruple MPEG–H también se restringe a solo 4 canales y no se puede aplicar dinámicamente a canales sino solo a un número de canales preconfigurado y fijo.

35 La herramienta de codificación multicanal MPEG–H permite la creación de un árbol arbitrario de cajas estereofónicas discretamente codificadas, es decir, pares de canales conjuntamente codificados, véase [2].

40 Un problema que a menudo surge en la codificación de señales de audio es causado por la cuantificación, por ejemplo, cuantificación espectral. La cuantificación puede dar como resultado, posiblemente huecos espectrales. Por ejemplo, todos los valores espectrales en una banda de frecuencia particular se pueden fijar en cero en el lado del codificador como resultado de cuantificación. Por ejemplo, el valor exacto de tales líneas espectrales antes de la cuantificación puede ser relativamente bajo y la cuantificación luego llevar a una situación en la que los valores espectrales de todas las líneas espectrales, por ejemplo, dentro de una banda de frecuencia particular se fijaron en cero. En el lado del decodificador, cuando se decodifica, esto puede llevar a huecos espectrales no deseados.

45 Los sistemas modernos de lenguaje de dominio de frecuencia/codificación de audio tales como Opus/Celt codec de IETF [9], MPEG–4 (HE–)AAC [10] o, en particular, MPEG–D xHE–AAC (USAC) [11], ofrecen medios para codificación de tramas de audio usando ya sea una transformada larga –un bloque largo– u ocho transformadas cortas secuenciales –bloques cortos– según la temporalidad de la señal. Además, para la codificación de bajas velocidades, estos esquemas proporcionan herramientas para reconstruir coeficientes de frecuencia de un canal usando ruido pseudoaleatorio o coeficientes de baja frecuencia del mismo canal. En xHE–AAC, estas herramientas se conocen como llenado de ruido y replicación de bandas espectrales, respectivamente.

55 Sin embargo, para cada entrada estereofónica muy tonal o transitoria, el llenado de ruido y/o la replicación de bandas espectrales solos limitan la calidad de codificación lograble a muy bajas tasas de bits, mayormente dado que demasiados coeficientes espectrales de ambos canales necesitan ser transmitidos explícitamente.

60 El llenado estéreo MPEG–H es una herramienta paramétrica que se basa en el uso de una mezcla de tramas previa para mejorar el llenado de los huecos espectrales causado por la cuantificación en el dominio de frecuencia. Como el llenado de ruido, el llenado estéreo opera directamente en el dominio MDCT del codificador central MPEG–H, véase [1], [5], [8].

Sin embargo, el uso de MPEG envolvente y llenado estéreo en MPEG–H se restringe a elementos de pares de canales fijos y, por ende, no pueden explotar dependencias intercanales de tiempo variante.

5 La herramienta de codificación multicanal (MCT) en MPEG-H permite adaptar la variación de las dependencias intercanal pero, debido al uso de elementos de un canal en configuraciones operativas típicas, no permite el llenado estéreo. La técnica anterior no revela vías perceptualmente óptimas para generar mezclas previas de tramas en caso de pares de canales conjuntamente codificados de tiempo variante arbitrarios. El uso del llenado de ruido como un sustituto para el llenado estéreo en combinación con MCT para llenar huecos espectrales llevaría a artefactos de ruido, en especial para señales tonales.

10 El objeto de la presente invención consiste en proporcionar conceptos mejorados de codificación de audio. El objeto de la presente invención se resuelve por un aparato para decodificar de acuerdo con la reivindicación 1, por un método para decodificar de acuerdo con la reivindicación 17 y por un programa de computadora de acuerdo con la reivindicación 18.

15 Se proporciona un aparato para decodificar una señal multicanal codificada de una trama actual para obtener tres o más canales de salida de audio actuales. Un procesador multicanal se adapta para seleccionar dos canales decodificados de tres o más canales decodificados según los primeros parámetros multicanal. Más aún, el procesador multicanal se adapta para generar un primer grupo de dos o más canales procesados a base de dichos canales seleccionados. Un módulo de llenado de ruidos se adapta para identificar al menos uno de los canales seleccionados, una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero y para generar un canal de mezcla usando, según la información secundaria, un subgrupo apropiado de tres o más canales de salida de audio previos que fueron decodificados y para llenar las líneas espectrales de bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero, con ruido generado usando líneas espectrales del canal de mezcla.

25 De acuerdo con formas de realización, se proporciona un aparato para decodificar una señal multicanal codificada previa de una trama previa para obtener tres o más canales de salida de audio previos y para decodificar una señal multicanal codificada actual de una trama actual para obtener tres o más canales de salida de audio actuales.

30 El aparato comprende una interfaz, un decodificador de canales, un procesador multicanal para generar los tres o más canales de salida de audio actuales y un módulo de llenado de ruido. La interfaz se adapta para recibir la señal multicanal codificada actual y para recibir información secundaria que comprende primeros parámetros multicanal.

35 El decodificador de canales se adapta para decodificar la señal multicanal codificada actual de la trama actual para obtener un grupo de tres o más canales decodificados de la trama actual.

El procesador multicanal se adapta para seleccionar un primer par seleccionado de dos canales decodificados del grupo de tres o más canales decodificados según los primeros parámetros multicanal.

40 Más aún, el procesador multicanal se adapta para generar un primer grupo de dos o más canales procesados en función de dicho primer par seleccionado de dos canales decodificados para obtener un grupo actualizado de tres o más canales decodificados.

45 Antes de que el procesador multicanal genere el primer par de dos o más canales procesados en función de dicho primer par seleccionado de dos canales decodificados, el módulo de llenado de ruidos se adapta para identificar al menos uno de los dos canales de dicho primer par seleccionado de dos canales decodificados, una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero y para generar un canal de mezcla usando dos o más, pero no todos los tres o más canales de salida de audio previos y para llenar las líneas espectrales de la una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero, con ruido generado usando líneas espectrales del canal de mezcla, en donde el módulo de llenado de ruidos se adapta para seleccionar los dos o más canales de salida de audio previos que se usan para generar el canal de mezcla de los tres o más canales de salida de audio previos según la información secundaria.

55 Un concepto particular de formas de realización que se pueden emplear por el módulo de llenado de ruidos que especifica cómo generar y llenar el ruido se menciona como llenado estéreo.

Más aún, se proporciona un método para decodificar una señal multicanal codificada previa de una trama previa para obtener tres o más canales de salida de audio previos y para decodificar una señal multicanal codificada actual de una trama actual para obtener tres o más canales de salida de audio actuales. El método comprende:

- 60
- Recibir la señal multicanal codificada actual y recibir información secundaria que comprende primeros parámetros multicanal.
  - Decodificar la señal multicanal codificada actual de la trama actual para obtener un grupo de tres o más canales decodificados de la trama actual.

- Seleccionar un primer par seleccionado de dos canales decodificados del grupo de tres o más canales decodificados según los primeros parámetros multicanal.

- 5    – Generar un primer grupo de dos o más canales procesados en función de dicho primer par seleccionado de dos canales decodificados para obtener un grupo actualizado de tres o más canales decodificados.

Antes de generar el primer par de dos o más canales procesados en función de dicho primer par seleccionado de dos canales decodificados, se llevan a cabo las siguientes etapas:

- 10    – Identificar para al menos uno de los dos canales de dicho primer par seleccionado de dos canales decodificados, una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero y generar un canal de mezcla usando dos o más, pero no todos los tres o más canales de salida de audio previos y llenar las líneas espectrales de la una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se
- 15    cuantifican a cero, con ruido generado usando líneas espectrales del canal de mezcla, en donde la selección de los dos o más canales de salida de audio previos que se usan para generar el canal de mezcla de los tres o más canales de salida de audio previos se lleva a cabo según la información secundaria.

20    Más aún, se proporcionan programas de computadora, en donde cada uno de los programas de computadora se configura para implementar uno de los métodos antes descritos cuando se ejecuta en una computadora o procesador de señales, de modo que cada uno de los métodos antes descritos es implementado por uno de los programas de computadora.

25    Todas las siguientes apariciones de la palabra “realización(es)”, si se refieren a combinaciones de características diferentes de aquellas definidas por las reivindicaciones independientes, se refieren a ejemplos que se presentaron originalmente pero que no representan realizaciones de la invención actualmente reivindicada; estos ejemplos todavía se muestran con fines ilustrativos únicamente.

30    A continuación, se describen formas de realización de la presente invención con mayor detalle con referencia a las figuras, en donde:

- Fig. 1a        muestra un aparato para decodificar de acuerdo con una forma de realización;
- Fig. 1b        muestra un aparato para decodificar de acuerdo con otra forma de realización;
- 35    Fig. 2        muestra un diagrama de bloque de un decodificador de dominio–frecuencia paramétrica de acuerdo con una forma de realización de la presente solicitud;
- Fig. 3        muestra un diagrama esquemático que ilustra la secuencia de espectros que forman los espectrogramas de canales de una señal de audio multicanal a fin de facilitar la comprensión de la descripción del decodificador de la Fig. 2;
- 40    Fig. 4        muestra un diagrama esquemático que ilustra espectros actuales fuera de los espectrogramas mostrados en la Fig. 3 a fin de aliviar la comprensión de la descripción de la Fig. 2;
- 45    Fig. 5a y 5b   muestran un diagrama de bloque de un decodificador de audio de dominio de frecuencia paramétrica de acuerdo con una forma de realización alternativa de acuerdo con la que la mezcla de la trama previa se usa como una base para el llenado de ruido intercanal;
- 50    Fig. 6        muestra un diagrama de bloque de un codificador de audio de dominio de frecuencia paramétrica de acuerdo con una forma de realización;
- Fig. 7        muestra un diagrama de bloque esquemático de un aparato para codificar una señal multicanal que tiene al menos tres canales, de acuerdo con una forma de realización;
- 55    Fig. 8        muestra un diagrama de bloque esquemático de un aparato para codificar una señal multicanal que tiene al menos tres canales, de acuerdo con una forma de realización;
- Fig. 9        muestra un diagrama de bloque esquemático de una caja estéreo, de acuerdo con una forma de
- 60    realización;
- Fig. 10       muestra un diagrama de bloque esquemático de un aparato para decodificar una señal multicanal codificada que tiene canales codificados y al menos dos parámetros multicanal, de acuerdo con una forma de realización;

- Fig. 11 muestra un diagrama de flujo de un método para codificar una señal multicanal que tiene al menos tres canales, de acuerdo con una forma de realización;
- 5 Fig. 12 muestra un diagrama de flujo de un método para decodificar una señal multicanal codificada que tiene canales codificados y al menos dos parámetros multicanal, de acuerdo con una forma de realización;
- Fig. 13 muestra un sistema de acuerdo con una forma de realización;
- 10 Fig. 14 muestra en el escenario (a) una generación de canales de combinación para una primera trama en el escenario y en el escenario (b) una generación de canales de combinación para una segunda trama que sucede a la primera trama de acuerdo con una forma de realización; y
- 15 Fig. 15 muestra un esquema de indexación para los parámetros multicanal de acuerdo con formas de realización.

Los elementos iguales o equivalentes o elementos con igual o equivalente funcionalidad se designan en la siguiente descripción por números de referencia iguales o equivalentes.

- 20 En la siguiente descripción, una pluralidad de detalles se establece para proporcionar una explicación más exhaustiva de las formas de realización de la presente invención. Sin embargo, será obvio para los expertos en la técnica que las formas de realización de la presente invención se puedan poner en práctica sin estos detalles específicos. En otras instancias, las estructuras y dispositivos bien conocidos se muestran en una forma de diagrama de bloques más que en detalle, a fin de evitar oscurecer las formas de realización de la presente invención. Además, las características de las distintas formas de realización descritas en la presente más adelante se pueden combinar entre sí, a menos que se denote específicamente otra cosa.
- 25

- Antes de describir el aparato 201 para decodificar de la Fig. 1a, en primer lugar, se describe el llenado de ruido para codificación multicanal de audio. En formas de realización, el módulo de llenado de ruido 220 de la Fig. 1a se puede configurar, por ejemplo, para llevar a cabo una o más de las tecnologías de abajo que se describen respecto del llenado de ruido para codificación multicanal de audio.
- 30

- La Fig. 2 muestra un decodificador de audio de dominio de frecuencia de acuerdo con una forma de realización de la presente solicitud. El decodificador se indica en general usando el signo de referencia 10 y comprende un identificador de la banda de factor de escala 12, un descuantificador 14, un llenador de ruidos 16 y un transformador inverso 18, así como un extractor de línea espectral 20 y un extractor de factor de escala 22. Otros elementos opcionales que pueden estar comprendidos en el decodificador 10 abarcan un predictor estéreo complejo 24, un decodificador MS (mid-side) 26 y una herramienta de filtro de TNS inverso (Temporal Noise Shaping) de los cuales se muestran dos instancias 28a y 28b en la Fig. 2. Además, se muestra un proveedor de mezcla y se destaca con mayor detalle más abajo usando el signo de referencia 30.
- 35
- 40

- El decodificador de audio de dominio de frecuencia 10 de la Fig. 2 es un decodificador paramétrico que soporta el llenado de ruido de acuerdo con el que se llena cierta banda de factor de escala cuantificada a cero con ruido usando el factor de escala de esa banda de factor de escala como un medio para controlar el nivel del ruido llenado en esa banda de factor de escala. Más allá de ello, el decodificador 10 de la Fig. 2 representa un decodificador de audio multicanal configurado para reconstruir una señal de audio multicanal a partir de un flujo de datos de entrada 30. La Fig. 2, sin embargo, se concentra en los elementos del decodificador 10 implicados en la reconstrucción de una de las señales de audio multicanal codificadas en el flujo de datos 30 y emite este canal (de salida) en una salida 32. Un signo de referencia 34 indica que el decodificador 10 puede comprender otros elementos o puede comprender cierto control de la operación de tuberías responsable de la reconstrucción de los otros canales de la señal de audio multicanal en donde la descripción que figura a continuación indica cómo la reconstrucción del decodificador 10 del canal de interés en la salida 32 interactúa con la decodificación de los otros canales.
- 45
- 50

- La señal de audio multicanal representada por el flujo de datos 30 puede comprender dos o más canales. A continuación, la descripción de las formas de realización de la presente solicitud se concentra en el caso estéreo donde la señal de audio multicanal comprende meramente dos canales, pero en principio las formas de realización proporcionadas a continuación se pueden transferir con facilidad a formas de realización alternativas que se refieren a señales de audio multicanal y su codificación que comprende más de dos canales.
- 55

- Como quedará claro de la descripción de la Fig. 2 de abajo, el decodificador 10 de la Fig. 2 es un decodificador de transformada. Es decir, de acuerdo con la técnica de codificación en que se basa el decodificador 10, los canales se codifican en un dominio de transformada de modo de usar una transformada envuelta de los canales. Más aún, según el creador de la señal de audio, hay fases temporales durante las cuales los canales de la señal de audio representan ampliamente el mismo contenido de audio, desviándose de otro meramente por cambios menores o deterministas,
- 60

tales como diferentes amplitudes y/o fase a fin de representar una escena de audio donde las diferencias entre los canales permiten el posicionamiento virtual de una fuente de audio de la escena de audio con respecto a posiciones virtuales del parlante asociadas con los canales de salida de la señal de audio multicanal. En otras fases temporales, sin embargo, los diferentes canales de la señal de audio pueden ser más o menos no correlacionados entre sí e incluso pueden representar, por ejemplo, fuentes de audio completamente diferentes.

A fin de dar cuenta de la relación posiblemente variable en el tiempo entre los canales de la señal de audio, el códec de audio en que se basa el decodificador 10 de la Fig. 2 permite un uso variable en el tiempo de diferentes medidas para explotar redundancias intercanal. Por ejemplo, la codificación MS permite la conmutación entre representar los canales izquierdo y derecho de una señal de audio estéreo como tales o como un par de canales M (mid) y S (side) que representan la mezcla de canales izquierdo y derecho y la mitad de su diferencia, respectivamente. Es decir, continuamente hay –en un sentido espectrot temporal– espectrogramas de dos canales transmitidos por flujo de datos 30, pero los significados de estos canales (transmitidos) pueden cambiar en tiempo y respecto de los canales de salida, respectivamente.

La predicción estéreo compleja –otra herramienta de explotación de redundancia intercanal– permite, en el dominio espectral, predecir coeficientes del dominio de frecuencia del canal o líneas espectrales usando líneas colocadas espectralmente de otro canal. Más abajo se describen mayores detalles respecto de ello.

A fin de facilitar la comprensión de la siguiente descripción de la Fig. 2 y sus componentes mostrados allí, la Fig. 3 muestra, para el caso ilustrativo de una señal de audio estéreo representada por el flujo de datos 30, una vía posible de cómo los valores de muestra para las líneas espectrales de los dos canales se pueden codificar en el flujo de datos 30 de modo de procesarlos por el decodificador 10 de la Fig. 2. En particular, si bien está representado en la mitad de arriba de la Fig. 3 el espectrograma 40 de un primer canal de la señal de audio estéreo, la mitad de abajo de la Fig. 3 ilustra el espectrograma 42 del otro canal de la señal de audio estéreo. Nuevamente, vale la pena notar que el “significado” de los espectrogramas 40 y 42 puede cambiar con el tiempo debido, por ejemplo, a una conmutación variable en el tiempo entre un dominio codificado por MS y un dominio no codificado por MS. En la primera instancia, los espectrogramas 40 y 42 se refieren a un canal M y S, respectivamente, mientras que, en el último caso, los espectrogramas 40 y 42 se refieren a los canales izquierdo o derecho. La conmutación entre el dominio codificado por MS y el dominio no codificado por MS se puede señalar en el flujo de datos 30.

La Fig. 3 muestra que los espectrogramas 40 y 42 se pueden codificar en el flujo de datos 30 a una resolución espectrot temporal variable en el tiempo. Por ejemplo, ambos canales (transmitidos) se pueden subdividir, de una manera alineada en el tiempo, en una secuencia de tramas indicadas usando llaves 44 que pueden ser de igual largo y contactan sin superposición. Como se acaba de mencionar, la resolución espectral en donde los espectrogramas 40 y 42 se representan en el flujo de datos 30 puede cambiar con el tiempo. De modo preliminar, se asume que la resolución espectrot temporal cambia en tiempo igualmente para los espectrogramas 40 y 42, pero también es viable una extensión de esta simplificación como será obvio de la siguiente descripción. El cambio de la resolución espectrot temporal se señala, por ejemplo, en el flujo de datos 30 en unidades de las tramas 44. Es decir, la resolución espectrot temporal cambia en unidades de tramas 44. El cambio en la resolución espectrot temporal de los espectrogramas 40 y 42 se logra conmutando el largo de la transformada y la cantidad de transformadas usadas para describir los espectrogramas 40 y 42 dentro de cada trama 44. En el ejemplo de la Fig. 3, las tramas 44a y 44b ejemplifican tramas donde se usó una transformada larga a fin de muestrear los canales de señal de audio allí, resultando así en la mayor resolución espectral con un valor de muestra de línea espectral por línea espectral para cada una de tales tramas por canal. En la Fig. 3, los valores de muestra de las líneas espectrales se indican usando pequeñas cruces dentro de las cajas, en donde las cajas, a su vez, se disponen en filas y columnas y han de representar una rejilla temporal espectral donde cada fila corresponde a una línea espectral y cada columna corresponde a subintervalos de tramas 44 correspondientes a las transformadas más cortas implicadas en la formación de espectrogramas 40 y 42. En particular, la Fig. 3 ilustra, por ejemplo, para la trama 44d, que una trama puede estar sujeta alternativamente a transformadas consecutivas de longitud más corta, resultando, para tales tramas como trama 44d, en varios espectros que se suceden temporalmente de reducida resolución espectral. Ocho transformadas cortas se usan a modo de ejemplo para la trama 44d, dando como resultado un muestreo espectrot temporal de los espectrogramas 40 y 42 dentro de esa trama 44d, en las líneas espectrales separadas entre sí de modo que estén pobladas meramente cada octava línea espectral, pero con un valor de muestra para cada una de las ocho ventanas de transformada o transformadas de longitud más corta usadas para transformar la trama 44d. Con fines ilustrativos, se muestra en la Fig. 3 que también serían viables otros números de transformadas para una trama, como el uso de dos transformadas de una longitud de transformada que es, por ejemplo, la mitad del largo de la transformada de las transformadas largas para las tramas 44a y 44b, dando como resultado así un muestreo de la grilla espectrot temporal o los espectrogramas 40 y 42, donde dos valores de muestra de líneas espectrales se obtienen para cada segunda línea espectral, una de las que se refiere a la transformada guía, la otra a la transformada de arrastre.

Las ventanas de transformada para las transformadas en las que se subdividen las tramas se ilustran en la Fig. 3 debajo de cada espectrograma usando líneas de tipo ventana de superposición. La superposición temporal sirve, por ejemplo, para fines de TDAC (Time–Domain Aliasing Cancellation).

A pesar de que las formas de realización descritas más abajo también se podrían implementar de otro modo, la Fig. 3 ilustra el caso en que la conmutación entre las diferentes resoluciones espectrotemporales para las tramas 44 individuales se llevan a cabo de una forma tal que, para cada trama 44, la misma cantidad de valores de línea espectral indicada por las pequeñas cruces en la Fig. 3 resulta para el espectrograma 40 y el espectrograma 42, residiendo la diferencia meramente en la forma en que las líneas muestrean espectrotemporalmente el respectivo mosaico espectrotemporal en la respectiva trama 44, extendidas temporalmente en el tiempo de la respectiva trama 44 y extendidas espectralmente de frecuencia cero a la frecuencia máxima  $f_{max}$ .

Usando las flechas en la Fig. 3, la Fig. 3 ilustra con respecto a la trama 44d que se pueden obtener similares espectros para todas las tramas 44 distribuyendo apropiadamente los valores de muestra de línea espectral que pertenecen a la misma línea espectral, pero ventanas de transformada corta dentro de una trama de un canal, en las líneas espectrales no ocupadas (vacías) dentro de esa trama hasta la siguiente línea espectral ocupada de la misma trama. Estos espectros resultantes se denominan a continuación "espectros intercalados". Al intercalar n transformadas de una trama de un canal, por ejemplo, siguen valores de línea espectral espectralmente colocalizadas de las n transformadas cortas a otros antes de que siga el conjunto de n valores de línea espectral espectralmente colocalizadas de las n transformadas cortas de la línea espectral que se sucede espectralmente. Una forma intermedia de intercalación también sería viable: en vez de intercalar todos los coeficientes de líneas espectrales de una trama, sería viable intercalar meramente los coeficientes de líneas espectrales de un subgrupo apropiado de las transformadas cortas de una trama 44d. En todo caso, siempre que se analicen los espectros de tramas de los dos canales correspondientes a los espectrogramas 40 y 42, estos espectros pueden referirse a los intercalados o no intercalados.

A fin de codificar de modo eficaz los coeficientes de líneas espectrales que representan los espectrogramas 40 y 42 a través del flujo de datos 30 que pasan por el decodificador 10, se cuantifican los mismos. A fin de controlar el ruido de cuantificación espectrotemporalmente, el tamaño de la etapa de cuantificación se controla por medio de factores de escala que se establecen en una determinada grilla espectrotemporal. En particular, dentro de cada una de las secuencias de espectros de cada espectrograma, se agrupan las líneas espectrales en grupos de factores de escala no superpuestos espectralmente consecutivos. La Fig. 4 muestra un espectro 46 del espectrograma 40 en su mitad superior y un espectro cotemporal 48 fuera del espectrograma 42. Como se muestra en la presente, los espectros 46 y 48 se subdividen en bandas de factor de escala a lo largo del eje espectral f de modo de agrupar las líneas espectrales en grupos no superpuestos. Las bandas de factor de escala se ilustran en la Fig. 4 usando llaves 50. Por simplicidad, se asume que los límites entre las bandas de factor de escala coinciden entre el espectro 46 y 48, pero esto no necesita ser particularmente el caso.

Es decir, por medio de la codificación en el flujo de datos 30, los espectrogramas 40 y 42 se subdividen cada uno en una secuencia temporal de espectros y cada uno de estos espectros se subdivide espectralmente en bandas de factor de escala y para cada banda de factor de escala, el flujo de datos 30 codifica o lleva información sobre un factor de escala correspondiente a la respectiva banda de factor de escala. Los coeficientes de línea espectral que entran dentro de una respectiva banda de factor de escala 50 se cuantifican usando el respectivo factor de escala o, en lo que se refiere al decodificador 10, se pueden descuantificar usando el factor de escala de la correspondiente banda de factor de escala.

Antes de volver a cambiar a la Fig. 2 y su descripción, se ha de asumir a continuación que el canal específicamente tratado, es decir, una decodificación con la cual están implicados los elementos del decodificador específicos de la Fig. 2 excepto 34, es el canal transmitido de espectrograma 40 que, como ya se estableció con anterioridad, puede representar uno de los canales izquierdo y derecho, un canal M o un canal S asumiendo que la señal de audio multicanal codificada en el flujo de datos 30 es una señal de audio estéreo.

Si bien el extractor de línea espectral 20 se configura para extraer los datos de línea espectral, es decir, los coeficientes de líneas espectrales para tramas 44 del flujo de datos 30, el extractor del factor de escala 22 se configura para extraer para cada trama 44 los correspondientes factores de escala. Para este fin, los extractores 20 y 22 pueden usar decodificación de entropía. De acuerdo con una forma de realización, el extractor del factor de escala 22 se configura para extraer secuencialmente los factores de escala, por ejemplo, del espectro 46 en la Fig. 4, es decir, los factores de escala de bandas de factor de escala 50, del flujo de datos 30 usando decodificación de entropía adaptada al contexto. El orden de la decodificación secuencial puede seguir el orden espectral definido entre las bandas de factor de escala que lleva, por ejemplo, de baja frecuencia a alta frecuencia. El extractor de factor de escala 22 puede usar decodificación de entropía adaptada al contexto y puede determinar el contexto para cada factor de escala según factores de escala ya extraídos en una vecindad espectral de un factor de escala actualmente extraído, dependiendo del factor de escala de la banda de factor de escala inmediatamente precedente. De modo alternativo, el extractor del factor de escala 22 puede decodificar predictivamente los factores de escala del flujo de datos 30 como, por ejemplo, usando decodificación diferencial mientras predice un factor de escala actualmente decodificado en función de cualquiera de los factores de escala previamente decodificados como el inmediatamente precedente. Notablemente, este proceso de extracción de factor de escala es agnóstico con respecto a un factor de escala que pertenece a una

banda de factor de escala poblada por líneas espectrales cuantificadas a cero exclusivamente o poblada por líneas espectrales entre las que al menos una se cuantifica a un valor no cero. Un factor de escala que pertenece a una banda de factor de escala poblada por líneas espectrales cuantificadas a cero solo puede servir tanto como una base de predicción para un posterior factor de escala decodificado que pertenece posiblemente a una banda de factor de escala poblada por líneas espectrales entre los que uno es no cero y para predecir en función de un factor de escala previamente decodificado que posiblemente pertenece a una banda de factor de escala poblada por líneas espectrales entre los que uno es no cero.

Solo por un tema de exhaustividad, se observa que el extractor de línea espectral 20 extrae los coeficientes de líneas espectrales con los que las bandas de factor de escala 50 son pobladas del mismo modo usando, por ejemplo, codificación de entropía y/o codificación predictiva. La codificación de entropía puede usar adaptabilidad al contexto en función de coeficientes de línea espectral en una vecindad espectral de un coeficiente de línea espectral actualmente decodificado y del mismo modo, la predicción puede ser una predicción espectral, una predicción temporal o una predicción espectral que predice un coeficiente de línea espectral actualmente decodificado en función de coeficientes de línea espectral previamente decodificados en una de sus vecindades espectral-temporales. En aras de una eficacia de codificación mayor, el extractor de línea espectral 20 se puede configurar para realizar la decodificación de las líneas espectrales o los coeficientes de líneas en tuplas, que recolectan o agrupan las líneas espectrales a lo largo del eje de frecuencia.

De esta manera, a la salida del extractor de línea espectral 20, se proporcionan los coeficientes de líneas espectrales, por ejemplo, como en unidades de espectros como recolección de espectros 46, por ejemplo, todos los coeficientes de líneas espectrales de una correspondiente trama o alternativamente recolección de todos los coeficientes de líneas espectrales de ciertas transformadas cortas de una correspondiente trama. A la salida del extractor de factor de escala 22, a su vez, se emiten los correspondientes factores de escala de los respectivos espectros.

El identificador de banda de factor de escala 12, así como el descuantificador 14 tienen entradas de líneas espectrales acopladas con la salida del extractor de líneas espectrales 20 y el descuantificador 14 y el llenador de ruidos 16 tienen entradas de factores de escala acopladas con la salida del extractor de factor de escala 22. El identificador de banda de factor de escala 12 se configura para identificar las así llamadas bandas de factor de escala cuantificadas a cero dentro de un espectro actual 46, es decir, bandas de factor de escala dentro de las que todas las líneas espectrales se cuantifican a cero, tales como la banda de factor de escala 50c en la Fig. 4 y las demás bandas de factor de escala del espectro dentro del cual al menos una línea espectral se cuantifica a no cero. En particular, en la Fig. 4, los coeficientes de líneas espectrales se indican usando áreas sombreadas en la Fig. 4. Es visible de allí que en el espectro 46, todas las bandas de factor de escala salvo la banda de factor de escala 50b tienen al menos una línea espectral, cuyo coeficiente de línea espectral se cuantifica a un valor no cero. Más tarde, se aclarará que bandas de factor de escala cuantificadas a cero como 50d forman el sujeto del llenado de ruido intercanal descrito más abajo. Antes de proceder con la descripción, se ha de notar que el identificador de la banda de factor de escala 12 puede restringir su identificación meramente en un subgrupo apropiado de las bandas de factor de escala 50 como en las bandas de factor de escala sobre una determinada frecuencia inicial 52. En la Fig. 4, esto restringiría el procedimiento de identificación en bandas de factor de escala 50d, 50e y 50f.

El identificador de la banda de factor de escala 12 informa al llenador de ruido 16 sobre aquellas bandas de factor de escala que son bandas de factor de escala cuantificadas a cero. El descuantificador 14 usa los factores de escala asociados con un espectro de entrada 46 de modo de descuantificar o escalar los coeficientes de líneas espectrales de las líneas espectrales del espectro 46 de acuerdo con los factores de escala asociados, es decir, los factores de escala asociados con las bandas de factor de escala 50. En particular, el descuantificador 14 descuantifica y escala los coeficientes de línea espectral que entran en una respectiva banda de factor de escala con el factor de escala asociado con la respectiva banda de factor de escala. La Fig. 4 se ha de interpretar como muestra el resultado de la descuantificación de las líneas espectrales.

El llenador de ruidos 16 obtiene la información acerca de las bandas de factor de escala cuantificadas a cero que forman el sujeto del siguiente llenado de ruido, el espectro descuantificado, así como los factores de escala de al menos aquellas bandas de factor de escala identificadas como bandas de factor de escala cuantificadas a cero y una señalización obtenida de flujo de datos 30 para la trama actual que revela si el llenado de ruido intercanal se ha de realizar para la trama actual.

El proceso de llenado de ruido intercanal descrito en el siguiente ejemplo realmente implica dos tipos de llenado de ruido, a saber, la inserción de un ruido de fondo 54 que pertenece a todas las líneas espectrales que fueron cuantificadas a respecto de su relación potencial a cualquier banda de factor de escala cuantificada a cero y el procedimiento actual de llenado de ruido intercanal. A pesar de que esta combinación se describe de ahora en adelante en la presente, se ha de enfatizar que la inserción de ruido de fondo se puede omitir de acuerdo con una forma de realización alternativa. Más aún, la señalización respecto del encendido y apagado de llenado de ruido en lo que se refiere a la trama actual y obtenida del flujo de datos 30 se podría relacionar con el llenado de ruido intercanal solo o podría controlar la combinación de ambos tipos de llenado de ruido.

- En lo que respecta a la inserción del ruido de fondo, el llenador de ruidos 16 podría operar de la siguiente manera. En particular, el llenador de ruidos 16 podría emplear generación de ruido artificial como un generador de números pseudoaleatorios o alguna otra fuente de aleatoriedad a fin de llenar líneas espectrales, cuyos coeficientes de líneas espectrales eran cero. El nivel del ruido de fondo 54 así insertado en las líneas espectrales cuantificadas a cero se podrían fijar de acuerdo con una señalización explícita dentro del flujo de datos 30 para la trama actual o el espectro actual 46. El "nivel" del ruido de fondo 54 se podría determinar usando una media cuadrática (RMS) o medición de energía, por ejemplo.
- La inserción de ruido de fondo representa así un tipo de prellenado para aquellas bandas de factor de escala que fueron identificadas como aquellas cuantificadas a cero como la banda de factor de escala 50d en la Fig. 4. También afecta a otras bandas de factor de escala más allá de las cuantificadas a cero, pero las últimas también son sujeto del siguiente llenado de ruido intercanal. Como se describe más abajo, el proceso de llenado de ruido intercanal consiste en llenar las bandas de factor de escala cuantificadas a cero hasta un nivel que es controlado por medio del factor de escala de la respectiva banda de factor de escala cuantificada a cero. La última se puede usar directamente para este fin debido a todas las líneas espectrales de la respectiva banda de factor de escala cuantificada a cero que se cuantifica a cero. Sin embargo, el flujo de datos 30 puede contener una señalización adicional de un parámetro, para cada trama o cada espectro 46, que comúnmente se aplica a los factores de escala de todas las bandas de factor de escala cuantificadas a cero de la correspondiente trama o espectro 46 y resulta cuando se aplica sobre los factores de escala de las bandas de factor de escala cuantificadas a cero por el llenador de ruido 16, en un nivel de llenado respectivo que es individual para las bandas de factor de escala cuantificadas a cero. Es decir, el llenador de ruidos 16 se puede modificar, usando la misma función de modificación, para cada banda de factor de escala cuantificada a cero del espectro 46, usando el factor de escala de la respectiva banda de factor de escala el parámetro recién mencionado contenido en el flujo de datos 30 para ese espectro 46 de la trama actual de modo de obtener un nivel de destino de llenado para la respectiva banda de factor de escala cuantificada a cero midiendo, en términos de energía o RMS, por ejemplo, el nivel hasta el cual el proceso de llenado de ruido intercanal ha de llenar la respectiva banda de factor de escala cuantificada a cero con (opcionalmente) ruido adicional (además del ruido de fondo 54).
- En particular, a fin de llevar a cabo el llenado de ruido intercanal 56, el llenador de ruidos 16 obtiene una porción espectralmente colocalizada del otro espectro del canal 48, en un estado ya amplia o totalmente decodificado y copia la porción obtenida del espectro 48 en la banda de factor de escala cuantificada a cero donde se colocó esta porción espectralmente, escalada de manera tal que el nivel de ruido general resultante dentro de esa banda de factor de escala cuantificada a cero –derivada por una integración sobre las líneas espectrales de la respectiva banda de factor de escala– iguale el nivel de destino de llenado antes mencionado obtenido de la banda de factor de escala cuantificada a cero. Mediante esta medida, la tonalidad del ruido llenado en la respectiva banda de factor de escala cuantificada a cero se mejora en comparación con el ruido artificialmente generado como el que forma la base del ruido de fondo 54 y también es mejor que el copiado/replicación espectral no controlados de líneas de muy baja frecuencia dentro del mismo espectro 46.
- Para ser aún más precisos, el llenador de ruidos 16 localiza, para una banda actual como 50d, una porción espectralmente colocalizada dentro del espectro 48 del otro canal, escala sus líneas espectrales según el factor de escala de la banda de factor de escala cuantificada a cero 50d de una manera descrita justamente que implica, opcionalmente, cierta compensación adicional o parámetro de factor de ruido contenido en el flujo de datos 30 para la trama actual o espectro 46, de modo que su resultado llene la respectiva banda de factor de escala cuantificada a cero 50d hasta el nivel deseado según se define por el factor de escala de la banda de factor de escala cuantificada a cero 50d. En la presente forma de realización, esto significa que el llenado se realiza de una manera aditiva respecto del ruido de fondo 54.
- De acuerdo con una forma de realización simplificada, el espectro llenado con ruido 46 resultante ingresaría directamente en la entrada del transformador inverso 18 de modo de obtener, para cada ventana de transformada a la que pertenecen los coeficientes de líneas espectrales de espectro 46, una porción de dominio de tiempo de la respectiva señal de tiempo de audio del canal, después de lo cual (no mostrado en la Fig. 2) un proceso de adición–superposición puede combinar estas porciones de dominio de tiempo. Es decir, si el espectro 46 es un espectro no intercalado, los coeficientes de líneas espectrales de este meramente pertenecen a una transformada, entonces el transformador inverso 18 somete a esa transformada de modo de dar como resultado una porción de dominio de tiempo y los extremos precedentes y de arrastre que estarán sujetos a un proceso de superposición–adición con porciones de dominio de tiempo precedentes y de arrastre obtenidas por transformación inversa de transformadas inversas precedentes y posteriores de modo de realizar, por ejemplo, una cancelación de solapamiento de dominio de tiempo. Si, sin embargo, el espectro 46 tiene coeficientes de línea espectral intercalados de más de una transformada consecutiva, entonces el transformador inverso 18 se someterá a transformaciones inversas separadas para obtener una porción de dominio de tiempo por transformación inversa y de acuerdo con el orden temporal definido, estas porciones de dominio de tiempo se someterían a un proceso de superposición–adición entre medio, así como con respecto a porciones de dominio de tiempo precedentes y posteriores de los otros espectros o tramas.

Sin embargo, en aras de la exhaustividad, se debe observar que se puede llevar a cabo un posterior procesamiento en el espectro lleno de ruidos. Como se muestra en la Fig. 2, el filtro TNS inverso puede llevar a cabo una filtración de TNS inversa en el espectro lleno de ruidos. Es decir, controlado por medio de coeficientes de filtro de TNS para la trama o el espectro 46 actual, el espectro obtenido se somete a una filtración lineal a lo largo de la dirección espectral.

5 Con o sin filtración de TNS inversa, el predictor estéreo complejo 24 podría tratar luego al espectro como una predicción residual de una predicción intercanal. Más específicamente, el predictor intercanal 24 podría usar una porción espectralmente colocalizada del otro canal para predecir el espectro 46 o al menos un subgrupo de sus bandas de factor de escala 50. El proceso de predicción compleja se ilustra en la Fig. 4 con recuadro discontinuo 58 en relación con la banda de factor de escala 50b. Es decir, el flujo de datos 30 puede contener parámetros de predicción intercanal de control, por ejemplo, cuyas bandas de factor de escala 50 se han de precedir intercanal y que no se ha de predecir de esa manera. Por otra parte, los parámetros de predicción intercanal en el flujo de datos 30 también pueden comprender factores de predicción intercanal complejos aplicados por el predictor intercanal 24 de modo de obtener el resultado de predicción intercanal. Estos factores pueden estar contenidos en el flujo de datos 30 en forma individual para cada banda de factor de escala o alternativamente, cada grupo de uno o más bandas de factor de escala, para los que se activa la predicción intercanal o se señala para ser activados en el flujo de datos 30.

20 La fuente de predicción intercanal puede ser, como se indica en la Fig. 4, el espectro 48 del otro canal. Para ser más precisos, la fuente de predicción intercanal puede ser la porción espectralmente colocalizada del espectro 48, colocalizada en la banda de factor de escala 50b por predecir intercanal, extendida por una estimulación de esta parte imaginaria. La estimación de la parte imaginaria se puede llevar a cabo en función de la porción espectralmente colocalizada 60 del espectro 48 propiamente dicho, y/o puede usar una mezcla de los canales ya decodificados de la trama previa, es decir, la trama inmediatamente precedente a cuyo espectro 46 pertenece la trama actualmente decodificada. En efecto, el predictor intercanal 24 añade a las bandas de factor de escala por predecir intercanal como la banda de factor de escala 50b en la Fig. 4, la señal de predicción obtenida como recién se describió.

30 Como se acaba de observar en la descripción precedente, el canal al que pertenece el espectro 46 puede ser un canal codificado por MS o puede ser un canal relacionado con un altoparlante, como un canal izquierdo o derecho de una señal de audio estéreo. Por consiguiente, opcionalmente un MS decodificador 26 somete opcionalmente el espectro predicho intercanal 46 a decodificación de MS, para que realice, por línea espectral o espectro 46, una adición o sustracción con líneas espectrales espectralmente correspondientes del otro canal correspondiente a un espectro 48. Por ejemplo, a pesar de que no se muestra en la Fig. 2, el espectro 48 como se exhibe en la Fig. 4 se obtuvo por medio de la porción 34 del decodificador 10 de una forma análoga a la descripción indicada con anterioridad con respecto al canal al que pertenece el espectro 46 y el módulo de decodificación de MS 26, al realizar la decodificación de MS, somete los espectros 46 y 48 a adición en línea espectral o sustracción en línea espectral con ambos espectros 46 y 48 en la misma etapa dentro de la línea de procesamiento, lo que significa que ambos fueron obtenidos por predicción intercanal, por ejemplo, o ambos fueron obtenidos por llenado de ruido o filtración de TNS inversa.

40 Se observa que, opcionalmente, la decodificación de MS se puede llevar a cabo de una manera globalmente referida al espectro completo 46 o en forma individual activable por el flujo de datos 30 en unidades de, por ejemplo, bandas de factor de escala 50. En otras palabras, la decodificación de MS se puede encender o apagar usando la respectiva señalización en el flujo de datos 30 en unidades de, por ejemplo, tramas o alguna resolución espectrotemporal más fina como, por ejemplo, en forma individual para las bandas de factor de escala de los espectros 46 y/o 48 de los espectrogramas 40 y/o 42, en donde se asume que se definen los límites idénticos de ambas bandas de factor de escala de los canales.

50 Como se ilustra en la Fig. 2, la filtración de TNS inversa por filtro de TNS inverso 28 también se podría llevar a cabo después de cualquier procesamiento intercanal como predicción intercanal 58 o la decodificación de MS por el decodificador MS 26. El rendimiento en frente o corriente abajo del procesamiento intercanal se podría fijar o se podría controlar por medio de la respectiva señalización para cada trama en el flujo de datos 30 o a otro nivel de granularidad. Cuando se lleva a cabo la filtración de TNS inversa, los respectivos coeficientes del filtro de TNS presentes en el flujo de datos para el espectro actual 46 controlan un filtro de TNS, es decir, un filtro de predicción lineal que corre a lo largo de la dirección espectral para filtrar de modo lineal el espectro ingresado en el respectivo módulo de filtro de TNS inverso 28a y/o 28b.

55 De esta manera, el espectro 46 que llega a la entrada del transformador inverso 18 puede ser sujeto de ulterior procesamiento como se acaba de describir. Nuevamente, la anterior descripción no pretende entenderse de manera tal que todas estas herramientas opcionales estén presentes ya sea concurrentemente o no. Estas herramientas pueden estar presentes en el decodificador 10 parcial o colectivamente.

60 En cualquier caso, el espectro resultante en la entrada del transformador inverso representa la reconstrucción final de la señal de salida del canal y forma la base de la mezcla antes mencionada para la trama actual que sirve, como se describió con respecto a la predicción compleja 58, como la base para la estimación de la parte imaginaria potencial para la siguiente trama por decodificar. También puede servir como la reconstrucción final para la predicción intercanal

de otro canal que aquel al que se refieren los elementos excepto 34 en la Fig. 2.

La respectiva mezcla se forma por el proveedor de mezcla 31 combinando este espectro final 46 con la respectiva versión final del espectro 48. La última entidad, es decir, la respectiva versión final del espectro 48, formó la base para la predicción intercanal compleja en el predictor 24.

La Fig. 5 muestra una alternativa relativa a la Fig. 2 en la medida en que la base para el llenado de ruido intercanal esté representada por la mezcla de líneas espectrales espectralmente colocadas de una trama previa de modo que, en el caso opcional de uso de una predicción intercanal compleja, la fuente de esta predicción intercanal compleja se usa dos veces, como una fuente para el llenado de ruido intercanal así como una fuente para la estimación de la parte imaginaria en la predicción intercanal compleja. La Fig. 5 muestra un decodificador 10 que incluye la porción 70 que pertenece a la decodificación del primer canal al que pertenece el espectro 46, así como la estructura interna de la otra porción 34 antes mencionada que está implicada en la decodificación del otro canal que comprende el espectro 48. El mismo signo de referencia se usó para los elementos internos de la porción 70 en un lado y 34 en el otro lado. Como puede verse, la construcción es la misma. En la salida 32, un canal de la señal de audio estéreo se emite y a la salida del transformador inverso 18 de la segunda porción del decodificador 34, resulta el otro canal (salida) de la señal de audio estéreo, indicando esta emisión con el signo de referencia 74. Nuevamente, las formas de realización descritas con anterioridad se pueden transferir simplemente a un caso de uso de más de dos canales.

El proveedor de la mezcla 31 es cutilizado por ambas porciones 70 y 34 y recibe espectros temporalmente colocados 48 y 46 de espectrogramas 40 y 42 para formar una mezcla en función de ellos sumando estos espectros en una línea espectral por base de línea espectral, potencialmente formando el promedio de ello al dividir la suma en cada línea espectral por el número de canales mezclados, es decir, dos en el caso de la Fig. 5. En la salida del proveedor de la mezcla 31, la mezcla de la trama previa resulta por esta medición. Se observa en este sentido que, en el caso de la trama previa que contiene más de un espectro en uno de los espectrogramas 40 y 42, existen diferentes posibilidades de cómo el proveedor de la mezcla 31 opera en ese caso. Por ejemplo, en ese caso, el proveedor de la mezcla 31 puede usar el espectro de las transformadas de arrastre de la trama actual o puede usar un resultado intercalado de intercalación de todos los coeficientes de línea espectral de la trama actual de espectrograma 40 y 42. El elemento de demora 74 mostrado en la Fig. 5 cuando se conecta a la salida del proveedor de la mezcla 31, indica que la mezcla así proporcionada en la salida de proveedor de la mezcla 31 forma la mezcla de la trama previa 76 (véase la Fig. 4 con respecto al llenado de ruido intercanal 56 y la predicción compleja 58, respectivamente). De esta manera, la salida del elemento de demora 74 se conecta con las entradas de los predictores intercanal 24 de las porciones de decodificador 34 y 70, por un lado, y las entradas de llenadores de ruido 16 de porciones de decodificador 70 y 34, por el otro lado.

Es decir, si bien en la Fig. 2, el llenador de ruidos 16 recibe el otro espectro 48 finalmente reconstruido temporalmente colocado del canal de la misma trama actual como una base del llenado de ruido intercanal, en la Fig. 5, el llenado de ruido intercanal se lleva a cabo en vez de en función de la mezcla de la trama previa según se provee por el proveedor de la mezcla 31. La vía en la que se lleva a cabo el llenado de ruido intercanal, queda igual. Es decir, el llenador de ruidos intercanal 16 agarra una porción espectralmente colocada fuera del espectro de la trama actual del respectivo espectro del otro canal, en el caso de la Fig. 2 y el espectro final amplia o completamente decodificado según se obtiene de la trama previa que representa la mezcla de la trama previa, en el caso de la Fig. 5 y añade cierta porción de "fuente" a las líneas espectrales dentro de la banda de factor de escala para llenar con ruido, como 50d en la Fig. 4, escalada de acuerdo con un nivel de ruido blanco determinado por el factor de escala de la respectiva banda de factor de escala.

Concluyendo el análisis anterior de las formas de realización que describen el llenado de ruido intercanal en un decodificador de audio, será evidente para los expertos en la técnica que antes de añadir la porción espectral o temporalmente colocada agarrada del espectro de la "fuente" a las líneas espectrales de la banda de factor de escala "destino", se puede aplicar cierto preprocesamiento a las líneas espectrales de "fuente" sin apartarse del concepto general del llenado intercanal. En particular, puede ser beneficioso aplicar una operación de filtrado como, por ejemplo, un aplanamiento espectral o eliminación de la inclinación, a las líneas espectrales de la región de "fuente" por añadir a la banda de factor de escala "destino", como 50d en la Fig. 4, a fin de mejorar la calidad de audio del proceso de llenado de ruido intercanal. Del mismo modo y como un ejemplo de un espectro ampliamente decodificado (en vez de completamente), la porción de "fuente" antes mencionada se puede obtener de un espectro que no fue filtrado por un filtro de TNS inverso disponible (es decir, síntesis).

De esta manera, las formas de realización anteriores se referían a un concepto de un llenado de ruido intercanal. A continuación, se describe una posibilidad de cómo se puede formar el concepto anterior de llenado de ruido intercanal en un códec existente, a saber, xHE-AAC, en una manera compatible cuasi hacia atrás. En particular, de ahora en adelante, se describe una implementación preferida de las formas de realización anteriores, de acuerdo con la que una herramienta de llenado estéreo se forma en un códec de audio a base de xHE-AAC en una manera de señalización compatible cuasi hacia atrás. Al usar la implementación descrita luego más abajo, para ciertas señales estéreo, es viable un llenado estéreo de los coeficientes de transformada en uno de los dos canales en un códec de

audio a base de un MPEG–D xHE–AAC (USAC), mejorando así la calidad de codificación de ciertas señales de audio en especial a bajas tasas de bits. La herramienta de llenado estéreo se señala de modo compatible cuasi hacia atrás de modo que los decodificadores xHE–AAC heredados puedan analizar y decodificar los flujos de bits sin errores obvios de audio o abandonos. Como ya se describió con anterioridad, se puede obtener una mejor calidad general si un codificador de audio puede usar una combinación de los coeficientes previamente decodificados/cuantificados de dos canales estéreo para reconstruir coeficientes cuantificados a cero (no transmitidos) de uno de los canales actualmente decodificados. Por ello, es deseable permitir tal llenado estéreo (de los coeficientes de canales previos a presentes) además de la replicación de bandas espectrales (de coeficientes de canales de baja a alta frecuencia) y llenado de ruido (de una fuente pseudoaleatoria no correlacionada) en codificadores de audio, en especial xHE–AAC o codificadores a base de ellos.

Para permitir flujos de bits codificados con llenado estéreo por leer y analizar por decodificadores xHE–AAC heredados, la herramienta de llenado estéreo deseada se ha de usar en una forma compatible cuasi hacia atrás: su presencia no debería causar que los decodificadores heredados paren –o incluso no inicien– la decodificación. La capacidad de lectura del flujo de bits por la infraestructura de xHE–AAC también puede facilitar la adopción en el mercado.

Para lograr el deseo antes mencionado para la compatibilidad cuasi hacia atrás para una herramienta de llenado estéreo en el contexto de xHE–AAC o sus derivados potenciales, la siguiente implementación implica la funcionalidad de llenado estéreo, así como la capacidad de señalar la misma por sintaxis en el flujo de datos realmente implicado con el llenado de ruido. La herramienta de llenado estéreo trabajaría en línea con la descripción anterior. En un par de canales con configuración de ventana común, un coeficiente de una banda de factor de escala cuantificada a cero es, cuando se activa la herramienta de llenado estéreo, como una alternativa (o, según se describe, adicionalmente) al llenado de ruido, reconstruida por una suma o diferencia de los coeficientes de la trama previa en uno de los dos canales, con preferencia, el canal derecho. El llenado estéreo se lleva a cabo similarmente al llenado de ruido. La señalización se haría por medio de la señalización del llenado de ruido de xHE–AAC. El llenado estéreo se transporta por medio de la información secundaria del llenado de ruido de 8 bits. Esto es viable porque el estándar MPEG–D USAC [3] establece que todos los 8 bits se transmiten incluso si el nivel de ruido por aplicar es cero. En esa situación, algunos de los bits de llenado de ruidos se pueden reutilizar para la herramienta del llenado estéreo.

La compatibilidad cuasi hacia atrás respecto del análisis y la reproducción del flujo de bits por decodificadores xHE–AAC heredados se asegura de la siguiente manera. El llenado estéreo se señala por medio de un nivel de ruido de cero (es decir, los primeros tres bits de llenado de ruido que tienen todos, un valor de cero) seguido por los cinco bits no cero (que tradicionalmente representan una compensación de ruido) que contiene información secundaria para la herramienta de llenado estéreo, así como el nivel de ruido faltante. Si bien un decodificador xHE–AAC heredado no tiene en cuenta el valor de la compensación de ruido de 5 bits si el nivel de ruido de 3 bits es cero, la presencia de la herramienta de señalización de llenado estéreo solo tiene un efecto sobre el llenado de ruido en el decodificador heredado: el llenado de ruido se apaga ya que los primeros tres bits son cero y el resto de la operación de decodificación corre como se pretende. En particular, el llenado estéreo no se lleva a cabo debido al hecho de que se opera como el proceso de llenado de ruido, que se desactiva. Así, un decodificador heredado aún ofrece una decodificación “gentil” del flujo de bits mejorado 30 porque no necesita silenciar la señal de salida o incluso abortar la decodificación después de alcanzar una trama con llenado estéreo encendido. Naturalmente, sin embargo, es incapaz de proporcionar una reconstrucción pretendida correcta de coeficientes de línea llenados estéreo, que lleva a una calidad deteriorada en tramas afectadas en comparación con la decodificación por un decodificador apropiado capaz de tratar apropiadamente con la nueva herramienta de llenado estéreo. No obstante, asumiendo que la herramienta de llenado estéreo se usa como se pretende, es decir, solo en la entrada estéreo a bajas tasas de bits, la calidad a través de los decodificadores xHE–AAC debería ser mejor que si las tramas afectadas cayeran debido al silenciamiento o llevaran a otros errores obvios de reproducción.

A continuación, se presenta una descripción detallada de cómo se puede formar una herramienta de llenado estéreo, como una extensión, en el codec xHE–AAC.

Cuando se forma en el estándar, la herramienta de llenado estéreo se podría describir de la siguiente manera. En particular, tal herramienta de llenado estéreo (SF) representaría una nueva herramienta en la parte del dominio de frecuencia (FD) de audio 3 D MPEG–H. En línea con el análisis anterior, el objeto de tal herramienta de llenado estéreo sería la reconstrucción paramétrica de coeficientes espectrales MDCT a bajas tasas de bits, similar a lo que ya se puede lograr con llenado de ruido de acuerdo con la sección 7.2 del estándar descrito en el punto [3]. Sin embargo, a diferencia del llenado de ruido, que emplea una fuente de ruido pseudoaleatoria para generar valores espectrales MDCT de cualquier canal FD, SF estará disponible también para reconstruir los valores MDCT del canal derecho de un par estéreo conjuntamente codificado de canales usando una mezcla de los espectros de la trama previa izquierdo y derecho MDCT. SF, de acuerdo con la implementación establecida más abajo, se señala de modo compatible cuasi hacia atrás por medio de la información secundaria de llenado de ruido que se puede analizar correctamente por un decodificador MPEG–D USAC heredado.

La descripción de la herramienta podría ser la siguiente. Cuando SF es activo en una trama FD estéreo conjunta, los coeficientes MDCT de bandas de factor de escala vacías (es decir, completamente cuantificados a cero) del canal derecho (segundo), como 50d, son reemplazados por una suma o diferencia de los coeficientes MDCT de la trama previa de los correspondientes canales izquierdos y derechos decodificados (si FD). Si el llenado de ruido heredado es activo para el segundo canal, los valores pseudoaleatorios también se añaden a cada coeficiente. Los coeficientes resultantes de cada banda de factor de escala se escalan luego de modo que RMS (raíz del cuadrado de coeficiente medio) de cada banda coincida con el valor transmitido por el factor de escala de banda. Véase la sección 7.3 del estándar en [3].

Se pueden proporcionar algunos límites operativos para el uso de la nueva herramienta SF en el estándar MPEG-D USAC. Por ejemplo, la herramienta SF se puede poner a disposición para usar solo en el canal derecho FD de un par de canales FD comunes, es decir, un elemento de par de canales que transmiten un StereoCoreToolInfo( ) con common\_window == 1. Además, debido a la señalización compatible cuasi hacia atrás, la herramienta SF puede estar disponible para usar solo cuando noiseFilling == 1 en el contenedor de sintaxis UsacCoreConfig( ). Si cualquiera de los canales en el par está en LPD core\_mode, la herramienta SF puede no usarse, incluso si el canal derecho está en el modo FD.

Los siguientes términos y definiciones se usan más adelante en la presente a fin de describir más claramente la extensión del estándar como se describió en el punto [3].

En particular, en lo que respecta a los elementos de datos, se introducen de nuevo los siguientes elementos de datos:

stereo\_filling      bandera binaria que indica si se utiliza SF en la trama actual y el canal

Por otra parte, se introducen nuevos elementos de ayuda:

noise\_offset      compensación del llenado de ruidos para modificar las bandas de factores de escala cuantificados a cero (sección 7.2)

noise\_level      nivel de llenado de ruido que representa la amplitud de espectro de ruido añadido (sección 7.2)

downmix\_prev[ ]    mezcla (es decir, suma o diferencia) de los canales izquierdo y derecho de la trama previa

sf\_index[g][sfb]   índice de factor de escala (es decir, número entero transmitido) para el grupo de ventana g y sfb de banda

El proceso de decodificación del estándar se extenderá de la siguiente manera. En particular, la decodificación de un canal FD codificado estéreo conjunto con la herramienta SF activada se ejecuta en tres etapas secuenciales de la siguiente manera:

En primer lugar, la decodificación de la bandera stereo\_filling tendría lugar.

stereo\_filling no representa un elemento de flujo de bits independiente pero se deriva de los elementos de llenado de ruido, noise\_offset y noise\_level, en una bandera UsacChannelPairElement() y common\_window en StereoCoreToolInfo(). Si noiseFilling == 0 o common\_window == 0 o el canal actual es el canal izquierdo (primero) en el elemento, stereo\_filling es 0 y el proceso de llenado estéreo termina. Por otra parte,

```
si ((noiseFilling != 0) && (common_window != 0) && (noise_level == 0)) {
stereo_filling = (noise_offset & 16) / 16;
noise_level = (noise_offset & 14) / 2;
noise_offset = (noise_offset & 1) * 16;
}
```

```
de lo contrario {
stereo_filling = 0;
}
```

En otras palabras, si noise\_level == 0, noise\_offset contiene la bandera stereo\_filling seguido por datos de 4 bits de llenado de ruido, que luego se reordenan. Si bien esta operación altera los valores de noise\_level y noise\_offset, necesita realizarse antes del proceso de llenado de ruido de la sección 7.2. Más aún, el pseudocódigo anterior no se ejecuta en el canal izquierdo (primero) de un UsacChannelPairElement( ) o cualquier otro elemento.

A continuación, el cálculo de downmix\_prev tendría lugar.

downmix\_prev[ ], la mezcla espectral que se ha de usar para el llenado estéreo, es igual a dmx\_re\_prev[ ] usado para la estimación del espectro MDST en la predicción estéreo compleja (sección 7.7.2.3). Esto significa que

- Todos los coeficientes de downmix\_prev[ ] deben ser cero si cualquiera de los canales de la trama y el elemento

con el que la mezcla se lleva a cabo – es decir, la trama antes de la actualmente decodificada –usan `core_mode == 1` (LPD) o los canales usan largos de transformadas desiguales (`split_transform == 1` o conmutación de bloque a `window_sequence == EIGHT_SHORT_SEQUENCE` en solo un canal) o `useIndependencyFlag == 1`.

- Todos los coeficientes de `downmix_prev[ ]` deben ser cero durante el proceso de llenado estéreo si el largo de la transformada del canal cambiado de la última a la trama actual (es decir, `split_transform == 1` precedido por `split_transform == 0` o `window_sequence == EIGHT_SHORT_SEQUENCE` precedido por `window_sequence != EIGHT_SHORT_SEQUENCE` o viceversa resp.) en el elemento actual.
- Si se aplica división de transformada en los canales de la trama previa o actual, `downmix_prev[ ]` representa una mezcla de espectros intercalados línea por línea. Véase la herramienta de división de transformada para detalles.
- Si la predicción estéreo compleja no se utiliza en la trama y elemento actuales, `pred_dir` es igual a 0.

En consecuencia, la mezcla previa solo debe ser computada una vez para ambas herramientas, salvando la complejidad. La única diferencia entre `downmix_prev[ ]` y `dmx_re_prev[ ]` en la sección 7.7.2 es el comportamiento cuando la predicción estéreo compleja no se usa actualmente o cuando está activa pero `use_prev_frame == 0`. En ese caso, `downmix_prev[ ]` se computa para la codificación del llenado estéreo de acuerdo con la sección 7.7.2.3 incluso a pesar de que no se necesita `dmx_re_prev[ ]` para la decodificación de la predicción estéreo compleja y, en consecuencia, es indefinido/cero.

De ahora en adelante en la presente, se podría llevar a cabo el llenado estéreo de bandas de factor de escala vacías.

Si `stereo_filling == 1`, el siguiente procedimiento se lleva a cabo después del proceso de llenado de ruido en todas las bandas de factor de escala inicialmente vacías `sfb[ ]` debajo de `max_sfb_ste`, es decir, todas las bandas en las que todas las líneas MDCT se cuantificaron a cero. En primer lugar, las energías de `sfb[ ]` dado y las correspondientes líneas en `downmix_prev[ ]` se computan por medio de las sumas de los cuadrados lineales. Después, dado `sfbWidth` que contiene el número de líneas por `sfb[ ]`,

```

si (energy[sfb] < sfbWidth[sfb]) { /* el nivel del ruido no es el máximo o la banda
comienza por debajo de la región de llenado de ruido */
facDmx = sqrt((sfbWidth[sfb] - energy[sfb]) / energy_dmx[sfb]);
factor = 0.0;
/* si la mezcla previa no está vacía, añadir las líneas de mezcla escaladas de modo tal
que esa banda alcanza la energía de la unidad */
para (índice = swb_offset[sfb]; index < swb_offset[sfb+1]; index++) {
spectrum>window][index] += downmix_prev>window][index] * facDmx;
factor += spectrum>window][index] * spectrum>window][index];
}
si ((factor != sfbWidth[sfb]) && (factor > 0)) { /* la energía de unidad no se alcanza,
entonces modificar la banda */
factor = sqrt(sfbWidth[sfb] / (factor + 1e-8));
para (índice = swb_offset[sfb]; index < swb_offset[sfb+1]; index++) {
spectrum>window][index] *= factor;
}
}
}

```

para el espectro de cada ventana de grupo. Luego, los factores de escala se aplican al espectro resultante como en la sección 7.3, procesando los factores de escala de las bandas vacías como factores de escala regulares.

Una alternativa a la extensión anterior del estándar xHE–AAC usaría un método de señalización implícito compatible casi hacia atrás.

La implementación anterior en el marco del código xHE–AAC describe un enfoque que emplea un bit en un flujo de bits para señalar el uso de una nueva herramienta de llenado estéreo, contenida en `stereo_filling`, en un decodificador de acuerdo con la Fig. 2. Más precisamente, esta señalización (llámese explícitamente señalización compatible casi hacia atrás) permite usar los siguientes datos del flujo de bits heredados –aquí, la información secundaria de llenado de ruido– independientemente de la señalización SF: en la presente forma de realización, los datos de llenado de ruido no dependen de la información de llenado estéreo y viceversa. Por ejemplo, los datos del llenado de ruido que consisten todos en ceros (`noise_level = noise_offset = 0`) se pueden transmitir mientras que `stereo_filling` puede señalar cualquier valor posible (siendo una bandera binaria, ya sea 0 o 1).

En casos en los que no se requiere una estricta independencia entre el legado y los datos del flujo de bits de la invención y la señal de la invención es una decisión binaria, la transmisión de un bit de señalización explícita se puede evitar y dicha decisión binaria se puede señalar por la presencia o ausencia de lo que se llama señalización implícita compatible casi hacia atrás. Tomando nuevamente la forma de realización anterior como un ejemplo, el uso de

llenado estéreo se podría transmitir empleando simplemente la nueva señalización: si `noise_level` es cero y, al mismo tiempo, `noise_offset` no es cero, la bandera `stereo_filling` se fija igual a 1. Si tanto `noise_level` como `noise_offset` no son cero, `stereo_filling` es igual a 0. Una dependiente de esta señal implícita de la señal de llenado de ruido heredada se produce cuando tanto `noise_level` como `noise_offset` son cero. En este caso, no queda claro si se usa señalización SF de legado o nueva SF implícita. Para evitar tal ambigüedad, el valor de `stereo_filling` se debe definir por adelantado. En el presente ejemplo, es apropiado definir `stereo_filling = 0` si los datos de llenado de ruido consisten en todos ceros, ya que esto es lo que los codificadores heredados sin señal de capacidad de llenado estéreo con llenado de ruido no han de aplicar en una trama.

El tema que queda por resolver en el caso de señalización compatible implícita cuasi hacia atrás es cómo señalar `stereo_filling == 1` y ningún llenado de ruido al mismo tiempo. Como se explicó, los datos de llenado de ruido no deben ser todos cero y si se requiere una magnitud de ruido de cero, `noise_level ((noise_offset & 14)/2` como se mencionó con anterioridad) debe ser igual a 0. Esto deja solo un `noise_offset ((noise_offset & 1)*16` como se mencionó con anterioridad) mayor que 0 como una solución. El `noise_offset`, sin embargo, se considera en caso de llenado estéreo cuando se aplican los factores de escala, incluso si `noise_level` es cero. Afortunadamente, un codificador puede compensar el hecho de que un `noise_offset` de cero puede no ser transmisible al alterar los factores de escala afectados de modo que, después de escribir el flujo de bits, contienen una compensación que se deshace en el decodificador por medio de `noise_offset`. Esto permite dicha señalización implícita en la forma de realización anterior al costo de un aumento potencial en la tasa de datos de factor de escala. Así, la señalización de llenado estéreo en un pseudocódigo de la descripción anterior se podría cambiar de la siguiente manera, usando el bit de señalización SF salvado para transmitir `noise_offset` con 2 bits (4 valores) en vez de 1 bit:

```

si ((noiseFilling) && (common_window) && (noise_level == 0) && (noise_offset > 0)) {
stereo_filling = 1;
25 noise_level = (noise_offset & 28) / 4;
noise_offset = (noise_offset & 3) * 8;
}
de lo contrario {
30 stereo_filling = 0;
}

```

Con fines de exhaustividad, la Fig. 6 muestra un codificador de audio paramétrico de acuerdo con una forma de realización de la presente solicitud. En primer lugar, el codificador de la Fig. 6 que se indica en general usando el signo de referencia 90 comprende un transformador 92 para realizar la transformación de la versión de la señal de audio original, no distorsionada reconstruida a la salida 32 de la Fig. 2. Como se describe con respecto a la Fig. 3, se puede usar una transformada superpuesta con una conmutación entre diferentes longitudes de transformada con las correspondientes ventanas de transformada en unidades de tramas 44. La diferente longitud de transformada y las correspondientes ventanas de transformada se ilustran en la Fig. 3 usando el signo de referencia 104. De una manera similar a la Fig. 2, la Fig. 6 se concentra en una porción del codificador 90 responsable de codificar un canal de la señal de audio multicanal, mientras que otra porción de decodificador del dominio de canal 90 se indica en general usando el signo de referencia 96 en la Fig. 6.

En la salida del transformador 92, las líneas espectrales y los factores de escala no están cuantificados y sustancialmente no se produjo aún una pérdida de codificación. La emisión del espectrograma por el transformador 92 ingresa en un cuantificador 98, que se configura para cuantificar las líneas espectrales de la salida del espectrograma por el transformador 92, espectro por espectro, fijando y usando factores de escala preliminares de las bandas de factor de escala. Es decir, a la salida del cuantificador 98, resultan los factores de escala preliminares y los correspondientes coeficientes de línea espectral y una secuencia de un llenador de ruidos 16', un filtro de TNS inverso opcional 28a', predictor intercanal 24', decodificador MS 26' y filtro de TNS inverso 28b' se conectan secuencialmente para proporcionar al codificador 90 de la Fig. 6 la capacidad de obtener una versión reconstruida final del espectro actual cuando se obtiene en el lado del decodificador en la entrada del proveedor de la mezcla (véase la Fig. 2). En caso de usar la predicción intercanal 24' y/o usar el llenado de ruido intercanal en la versión que forma el ruido intercanal usando la mezcla de la trama previa, el codificador 90 también comprende un proveedor de mezcla 31' para formar una mezcla de las versiones reconstruidas finales de los espectros de los canales de la señal de audio multicanal. De hecho, para salvar los cálculos, en vez de las versiones finales, se pueden usar las versiones originales, no cuantificadas de dichos espectros de los canales por el proveedor de la mezcla 31' en la formación de la mezcla.

El codificador 90 puede usar la información acerca de la versión reconstruida final disponible de los espectros a fin de realizar la predicción espectral intratrama como la versión posible antes mencionada de realización de la predicción intercanal usando una estimación de la parte imaginaria, y/o a fin de realizar el control de la tasa, es decir, a fin de determinar, dentro de la pendiente de control de la tasa, que los posibles parámetros finalmente codificados en el flujo de datos 30 por el codificador 90 se fijan en un sentido óptimo de tasa / distorsión.

Por ejemplo, uno de tales parámetros fijados en tal curva de predicción y/o curva de control de la tasa del codificador 90 es, para cada banda de factor de escala cuantificada a cero identificada por el identificador 12', el factor de escala de la respectiva banda de factor de escala que meramente fue fijado preliminarmente por el cuantificador 98. En una curva de predicción y/o de control de la tasa del codificador 90, el factor de escala de las bandas de factor de escala cuantificadas a cero se fija en cierto sentido psicoacústicamente óptimo o de tasa/distorsión óptima de modo de determinar el nivel de ruido de destino antes mencionado, como se describió con anterioridad, junto con un parámetro de modificación opcional también transportado por el flujo de datos para la correspondiente trama al lado del decodificador. Se ha de notar que este factor de escala se puede computar usando solo las líneas espectrales del espectro y el canal al que pertenece (es decir, el espectro "destino", como se describió con anterioridad) o, de modo alternativo, se puede determinar usando tanto las líneas espectrales del espectro del canal "destino" como, además, las líneas espectrales del otro espectro del canal o el espectro de la mezcla de la trama previa (es decir, el espectro "fuente", como se introdujo más temprano) obtenido del proveedor de la mezcla 31'. En particular, para estabilizar el nivel de ruido de destino y para reducir las fluctuaciones temporales del nivel en los canales de audio decodificados en donde se aplica el llenado de ruido intercanal, el factor de escala de destino se puede computar usando una relación entre una medida de energía de las líneas espectrales en la banda de factor de escala "destino" y una medida de energía de las líneas espectrales colocalizadas en la región "fuente" correspondiente. Finalmente, como se observó con anterioridad, esta región "fuente" se puede originar a partir de una versión reconstruida final de otro canal o la mezcla de la trama previa o si la complejidad del codificador se ha de reducir, la versión original, no cuantificada del mismo canal o la mezcla de versiones originales, no cuantificadas de los espectros de la trama previa.

A continuación, se explica la codificación multicanal y decodificación multicanal de acuerdo con las formas de realización. En las formas de realización, el procesador multicanal 204 del aparato 201 para decodificar de la Fig. 1a se puede configurar, por ejemplo, para realizar una o más de las tecnologías de abajo que se describen respecto de la decodificación del ruido multicanal.

En primer lugar, sin embargo, antes de describir la decodificación multicanal, la codificación multicanal de acuerdo con formas de realización se explica con referencia a la Fig. 7 a la Fig. 9 y luego, la codificación multicanal se explica con referencia a las Fig. 10 y Fig. 12.

Ahora, la codificación multicanal de acuerdo con formas de realización se explica con referencia a las Fig. 7 a Fig. 9 y Fig. 11:

La Fig. 7 muestra un diagrama de bloque esquemático de un aparato (codificador) 100 para codificar una señal multicanal 101 que tiene al menos tres canales CH1 a CH3.

El aparato 100 comprende un procesador de iteración 102, un codificador de canales 104 y una interfaz de salida 106.

El procesador de iteración 102 se configura para calcular, en una primera etapa de iteración, valores de correlación intercanal entre cada par de los al menos tres canales CH1 a CH3 para seleccionar, en la primera etapa de iteración, un par que tiene un valor máximo o que tiene un valor por encima del umbral y para procesar el par seleccionado usando una operación de procesamiento multicanal para derivar parámetros multicanal MCH\_PAR1 para el par seleccionado y para derivar primeros canales procesados P1 y P2. A continuación, tales canales procesados P1 y tal canal procesado P2 también se pueden referir a un canal de combinación P1 y un canal de combinación P2, respectivamente. Por otra parte, el procesador de iteración 102 se configura para llevar a cabo el cálculo, la selección y el procesamiento en una segunda etapa de iteración usando al menos uno de los canales procesados P1 o P2 para derivar parámetros multicanal MCH\_PAR2 y segundos canales procesados P3 y P4.

Por ejemplo, como se indica en la Fig. 7, el procesador de iteración 102 puede calcular en la primera etapa de iteración un valor de correlación intercanal entre un primer par de los al menos tres canales CH1 a CH3, donde el primer par consiste en un primer canal CH1 y un segundo canal CH2, un valor de correlación intercanal entre un segundo par de los al menos tres canales CH1 a CH3, donde el segundo par consiste en el segundo canal CH2 y un tercer canal CH3 y un valor de correlación intercanal entre un tercer par de los al menos tres canales CH1 a CH3, donde el tercer par consiste en el primer canal CH1 y el tercer canal CH3.

En la Fig. 7, se asume que, en la primera etapa de iteración, el tercer par que consiste en el primer canal CH1 y el tercer canal CH3 comprende el máximo valor de correlación intercanal, de modo que el procesador de iteración 102 seleccione en la primera etapa de iteración el tercer par que tiene el máximo valor de correlación intercanal y procesa el par seleccionado, es decir, el tercer par, usando una operación de procesamiento multicanal para derivar parámetros multicanal MCH\_PAR1 para el par seleccionado y para derivar primeros canales procesados P1 y P2.

Por otra parte, el procesador de iteración 102 se puede configurar para calcular, en la segunda etapa de iteración, valores de correlación intercanal entre cada par de los al menos tres canales CH1 a CH3 y los canales procesados P1 y P2, para seleccionar, en la segunda etapa de iteración, un par que tiene un máximo valor de correlación intercanal o que tiene un valor por encima del umbral. En este caso, el procesador de iteración 102 se puede configurar para no

seleccionar el par seleccionado de la primera etapa de iteración en la segunda etapa de iteración (o en cualquier otra etapa de iteración).

5 Haciendo referencia al ejemplo mostrado en la Fig. 7, el procesador de iteración 102 también puede calcular un valor de correlación intercanal entre un cuarto par de canales que consiste en el primer canal CH1 y el primer canal procesado P1, un valor de correlación intercanal entre un quinto par que consiste en el primer canal CH1 y el segundo canal procesado P2, un valor de correlación intercanal entre un sexto par que consiste en el segundo canal CH2 y el primer canal procesado P1, un valor de correlación intercanal entre un séptimo par que consiste en el segundo canal CH2 y el segundo canal procesado P2, un valor de correlación intercanal entre un octavo par que consiste en el tercer canal CH3 y el primer canal procesado P1, un valor de intercorrelación entre un noveno par que consiste en el tercer canal CH3 y el segundo canal procesado P2 y un valor de correlación intercanal entre un décimo par que consiste en el primer canal procesado P1 y el segundo canal procesado P2.

15 En la Fig. 7, se asume que en la segunda etapa de iteración el sexto par que consiste en el segundo canal CH2 y el primer canal procesado P1 comprende el máximo valor de correlación intercanal, de modo tal que el procesador de iteración 102 seleccione en la segunda etapa de iteración el sexto par y procese el par seleccionado, es decir, el sexto par, usando una operación de procesamiento multicanal para derivar parámetros multicanal MCH\_PAR2 para el par seleccionado y para derivar segundos canales procesados P3 y P4.

20 El procesador de iteración 102 se puede configurar solo para seleccionar un par cuando la diferencia de nivel del par es menor que un umbral, donde el umbral es inferior a 40 dB, 25 dB, 12 dB o inferior a 6 dB. En este caso, los umbrales de 25 o 40 dB corresponden a los ángulos de rotación de 3 o 0,5 grados.

25 El procesador de iteración 102 se puede configurar para calcular valores de correlación enteros normalizados, en donde el procesador de iteración 102 se puede configurar para seleccionar un par, cuando el valor de correlación entero es mayor que, por ejemplo, 0,2 o preferentemente 0,3.

30 Por otra parte, el procesador de iteración 102 puede proporcionar los canales que resultan del procesamiento multicanal en el codificador de canales 104. Por ejemplo, haciendo referencia a la Fig. 7, el procesador de iteración 102 puede proporcionar el tercer canal procesado P3 y el cuarto canal procesado P4 que resultan del procesamiento multicanal realizado en la segunda etapa de iteración y el segundo canal procesado P2 que resultan del procesamiento multicanal realizado en la primera etapa de iteración en el codificador de canales 104. En este caso, el procesador de iteración 102 solo puede proporcionar aquellos canales procesados en el codificador de canales 104 que no se procesan (luego) en una posterior etapa de iteración. Como se muestra en la Fig. 7, el primer canal procesado P1 no se proporciona en el codificador de canales 104 ya que luego se procesa en la segunda etapa de iteración.

40 El codificador de canales 104 se puede configurar para codificar los canales P2 a P4 que resultan del procesamiento de iteración (o procesamiento multicanal) realizado por el procesador de iteración 102 para obtener canales codificados E1 a E3.

45 Por ejemplo, el codificador de canales 104 se puede configurar para usar monocodificadores (o monocabajas o monoherramientas) 120\_1 a 120\_3 para codificar los canales P2 a P4 que resultan del procesamiento de iteración (o procesamiento multicanal). Las monocabajas se pueden configurar para codificar los canales de modo tal que se requieran menos bits para codificar un canal que tiene menos energía (o una menor amplitud) que para codificar un canal que tiene más energía (o una mayor amplitud). Las monocabajas 120\_1 a 120\_3 pueden ser, por ejemplo, codificadores de audio a base de transformación. Por otra parte, el codificador de canales 104 se puede configurar para usar codificadores estéreo (por ejemplo, codificadores estéreo paramétricos o codificadores estéreo con pérdida) para codificar los canales P2 a P4 que resultan del procesamiento de iteración (o procesamiento multicanal).

50 La interfaz de salida 106 se puede configurar para generar la señal multicanal codificada 107 que tiene los canales codificados E1 a E3 y los parámetros multicanal MCH\_PAR1 y MCH\_PAR2.

55 Por ejemplo, la interfaz de salida 106 se puede configurar para generar la señal multicanal codificada 107 como una señal serial o flujo de bits serial y de modo que los parámetros multicanal MCH\_PAR2 estén en la señal codificada 107 antes de los parámetros multicanal MCH\_PAR1. De esta manera, un decodificador, una forma de realización del cual se describirá más abajo con respecto a la Fig. 10, recibirá los parámetros multicanal MCH\_PAR2 antes de los parámetros multicanal MCH\_PAR1.

60 En la Fig. 7, el procesador de iteración 102 realiza a modo de ejemplo dos operaciones de procesamiento multicanal, una operación de procesamiento multicanal en la primera etapa de iteración y una operación de procesamiento multicanal en la segunda etapa de iteración. Naturalmente, el procesador de iteración 102 también puede llevar a cabo otras operaciones de procesamiento multicanal en posteriores etapas de iteración. En este caso, el procesador de iteración 102 se puede configurar para realizar etapas de iteración hasta alcanzar un criterio de terminación de iteración. El criterio de terminación de iteración puede ser que un número máximo de etapas de iteración es igual o

mayor que un número total de canales de la señal multicanal 101 en dos o en donde el criterio de terminación de iteración es cuando los valores de correlación intercanal no tienen un valor mayor que el umbral, siendo el umbral preferentemente mayor que 0,2 o siendo el umbral preferentemente de 0,3. En otras formas de realización, el criterio de terminación de iteración puede ser que un número máximo de etapas de iteración es igual o mayor que un número total de canales de la señal multicanal 101 o en donde el criterio de terminación de iteración es cuando los valores de correlación intercanal no tienen un valor mayor que el umbral, siendo el umbral preferentemente mayor que 0,2 o siendo el umbral preferentemente 0,3.

Con fines ilustrativos, las operaciones de procesamiento multicanal realizadas por el procesador de iteración 102 en la primera etapa de iteración y la segunda etapa de iteración se ilustran a modo de ejemplo en la Fig. 7 por cajas de procesamiento 110 y 112. Las cajas de procesamiento 110 y 112 se pueden implementar en hardware o software. Las cajas de procesamiento 110 y 112 pueden ser cajas estéreo, por ejemplo.

En este caso, la dependencia de señales intercanal se puede explotar aplicando jerárquicamente herramientas conocidas de codificación estereofónica conjunta. Contrariamente a los enfoques previos de MPEG, los pares de señales para procesar no se predeterminan por una vía de señales fijas (por ejemplo, árbol de codificación estéreo) pero se pueden cambiar dinámicamente para adaptarse para ingresar características de señales. Las entradas de la caja estéreo actual pueden ser (1) canales no procesados, tales como los canales CH1 a CH3, (2) salidas de una caja estéreo precedente, tales como las señales procesadas P1 a P4 o (3) un canal de combinación de un canal no procesado y una salida de una caja estéreo precedente.

El procesamiento dentro de la caja estéreo 110 y 112 puede ser en función de predicción (como caja de predicción compleja en USAC) o en función de KLT/PCA (los canales de entrada se rotan (por ejemplo, por una matriz de rotación de  $2 \times 2$ ) en el codificador para maximizar la compactación de energía, es decir, concentrar la energía de señal en un canal, en el decodificador, se retransformarán las señales rotadas en las direcciones de señales de entrada originales).

En una posible implementación del codificador 100, (1) el codificador calcula una correlación intercanal entre cada par de canales y selecciona un par de señales apropiadas fuera de las señales de entrada y aplica la herramienta estereofónica a los canales seleccionados; (2) el codificador recalcula la correlación intercanal entre todos los canales (los canales no procesados así como los canales de salida intermedios procesados) y selecciona un par de señales apropiadas fuera de las señales de entrada y aplica la herramienta estereofónica a los canales seleccionados; y (3) el codificador repite la etapa (2) hasta que toda la correlación intercanal esté por debajo de un umbral o si se aplica un número máximo de transformaciones.

Como ya se mencionó, los pares de señales para procesar por el codificador 100 o más precisamente, el procesador de iteración 102, no se predeterminan por una vía de señales fijas (por ejemplo, árbol de codificación estéreo) pero se pueden cambiar dinámicamente para adaptar para ingresar características de señales. En este caso, el codificador 100 (o el procesador de iteración 102) se puede configurar para construir el árbol estéreo en dependencia de los al menos tres canales CH1 a CH3 de la señal multicanal (entrada) 101. En otras palabras, el codificador 100 (o el procesador de iteración 102) se puede configurar para construir el árbol estéreo en función de una correlación intercanal (por ejemplo, calculando, en la primera etapa de iteración, valores de correlación intercanal entre cada par de los al menos tres canales CH1 a CH3, para seleccionar, en la primera etapa de iteración, un par que tiene el valor máximo o un valor por encima del umbral y calculando, en una segunda etapa de iteración, valores de correlación intercanal entre cada par de los al menos tres canales y canales previamente procesados, para seleccionar, en la segunda etapa de iteración, un par que tiene el valor máximo o un valor por encima del umbral). De acuerdo con un enfoque de etapa, una matriz de correlación se puede calcular posiblemente para cada iteración que contiene las correlaciones de todos los canales procesados posiblemente en iteraciones previas.

Como se indicó con anterioridad, el procesador de iteración 102 se puede configurar para derivar parámetros multicanal MCH\_PAR1 para el par seleccionado en la primera etapa de iteración y para derivar parámetros multicanal MCH\_PAR2 para el par seleccionado en la segunda etapa de iteración. Los parámetros multicanal MCH\_PAR1 pueden comprender una primera identificación de pares de canales (o índice) que identifica (o señala) el par de canales seleccionados en la primera etapa de iteración, en donde los parámetros multicanal MCH\_PAR2 pueden comprender una segunda identificación de pares de canales (o índice) que identifica (o señala) el par de canales seleccionados en la segunda etapa de iteración.

A continuación, se describa una indexación eficaz de señales de entrada. Por ejemplo, los pares de canales se pueden señalar eficazmente usando un único índice para cada par, dependiendo de la cantidad total de canales. Por ejemplo, la indexación de pares para seis canales se puede mostrar en la siguiente tabla:

	0	1	2	3	4	5
0		0	1	2	3	4
1			5	6	7	8
2				9	10	11
3					12	13
4						14
5						

Por ejemplo, en la tabla anterior, el índice 5 puede señalar el par que consiste en el primer canal y el segundo canal. De modo similar, el índice 6 puede señalar el par que consiste en el primer canal y el tercer canal.

5 La cantidad total de posibles índices de pares de canales para n canales se pueden calcular para:

$$\text{numPairs} = \text{numChannels} * (\text{numChannels} - 1) / 2$$

Así, la cantidad de bits necesarios para la señalización de un par de canales da:

10

$$\text{numBits} = \text{floor}(\log_2(\text{numPairs} - 1)) + 1$$

Por otra parte, el codificador 100 puede usar una máscara de canal. La configuración de la herramienta multicanal puede contener una máscara de canal que indica para qué canales la herramienta está activa. De esta manera, LFEs (LFE = bajos efectos de frecuencia / canales de mejora) se pueden remover de la indexación de pares de canales, que permite una codificación más eficaz. Por ejemplo, para una configuración 11.1, esto reduce la cantidad de índices de pares de canales de  $12 * 11 / 2 = 66$  a  $11 * 10 / 2 = 55$ , permitiendo la señalización con 6 en vez de 7 bits. Este mecanismo también se puede usar para excluir canales que pretenden ser monoobjetos (por ejemplo, pistas de lenguajes múltiples). Al decodificar la máscara del canal (channelMask), se puede generar un mapa de canal (channelMap) para permitir un remapeo de índices de pares de canales para decodificar canales.

15

20

Más aún, el procesador de iteración 102 se puede configurar para derivar, para una primera trama, una pluralidad de indicaciones de pares seleccionadas, en donde la interfaz de salida 106 se puede configurar para incluir, en la señal multicanal 107, para una segunda trama, después de la primera trama, un indicador de mantenimiento, indicando que la segunda trama tiene la misma pluralidad de indicaciones de pares seleccionadas como la primera trama.

25

El indicador de mantenimiento o la bandera de árbol de mantenimiento se puede usar para señalar que no se transmite ningún árbol nuevo, pero se ha de usar el último árbol estéreo. Esto se puede usar para evitar una múltiple transmisión de la misma configuración de árbol estéreo si las propiedades de correlación de los canales se quedan permanentemente durante un período más prolongado.

30

La Fig. 8 muestra un diagrama de bloque esquemático de una caja estéreo 110, 112. La caja estéreo 110, 112 comprende entradas para una primera señal de entrada I1 y una segunda señal de entrada I2 y salidas para una primera señal de salida O1 y una segunda señal de salida O2. Como se indica en la Fig. 8, las dependencias de las señales de salida O1 y O2 de las señales de entrada I1 y I2 se pueden describir por los parámetros s S1 a S4.

35

El procesador de iteración 102 puede usar (o comprende) cajas estéreo 110, 112 a fin de realizar las operaciones de procesamiento multicanal en los canales de entrada y/o canales procesados a fin de derivar (luego) los canales procesados. Por ejemplo, el procesador de iteración 102 se puede configurar para usar cajas estéreo 110, 112 en función de predicción genética o de rotación en función de KLT (Karhunen-Loève-Transformation).

40

Un codificador genérico (o caja estéreo del lado del codificador) se puede configurar para codificar las señales de entrada I1 y I2 para obtener las señales de salida O1 y O2 en función de la ecuación:

45

$$\begin{bmatrix} O_1 \\ O_2 \end{bmatrix} = \begin{bmatrix} s_1 & s_2 \\ s_3 & s_4 \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}.$$

Un decodificador genérico (o caja estéreo del lado del decodificador) se puede configurar para decodificar las señales de entrada I1 y I2 para obtener las señales de salida O1 y O2 en función de la ecuación:

50

$$\begin{bmatrix} O_1 \\ O_2 \end{bmatrix} = \begin{bmatrix} s_1 & s_2 \\ s_3 & s_4 \end{bmatrix}^{-1} \cdot \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}.$$

Un codificador a base de predicciones (o caja estéreo del lado del codificador) se puede configurar para codificar las señales de entrada I1 y I2 para obtener las señales de salida O1 y O2 en función de la ecuación

$$\begin{bmatrix} O_1 \\ O_2 \end{bmatrix} = 0,5 \cdot \begin{bmatrix} 1 & 1 \\ 1-p & -(1+p) \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ I_2 \end{bmatrix},$$

en donde p es el coeficiente de predicción.

- 5 Un decodificador a base de predicciones (o caja estéreo del lado del decodificador) se puede configurar para decodificar las señales de entrada I1 y I2 para obtener las señales de salida O1 y O2 en función de la ecuación:

$$\begin{bmatrix} O_1 \\ O_2 \end{bmatrix} = \begin{bmatrix} 1+p & 1 \\ 1-p & -1 \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}.$$

- 10 Un codificador de rotación a base de KLT (o caja estéreo del lado del codificador) se puede configurar para codificar las señales de entrada I1 a I2 para obtener las señales de salida O1 y O2 en función de la ecuación:

$$\begin{bmatrix} O_1 \\ O_2 \end{bmatrix} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}.$$

- 15 Un decodificador de rotación a base de KLT (o caja estéreo del lado del decodificador) se puede configurar para decodificar las señales de entrada I1 y I2 para obtener las señales de salida O1 y O2 en función de la ecuación (rotación inversa):

$$\begin{bmatrix} O_1 \\ O_2 \end{bmatrix} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}.$$

- 20 A continuación, se describe un cálculo del ángulo de rotación  $\alpha$  para la rotación a base de KLT.

El ángulo de rotación  $\alpha$  para la rotación a base de KLT se puede definir como:

$$25 \quad \alpha = \frac{1}{2} \tan^{-1} \left( \frac{2c_{12}}{c_{11} - c_{22}} \right)$$

siendo  $c_{xy}$  las entradas de una matriz de correlación no normalizada, en donde  $c_{11}$ ,  $c_{22}$  son las energías de los canales.

- 30 Esto se puede implementar usando la función atan2 para permitir la diferenciación entre correlaciones negativas en el numerador y la diferencia de energía negativa en el denominador:

$$\text{alfa} = 0,5 * \text{atan2}(2 * \text{correlación}[\text{ch1}][\text{ch2}], (\text{correlación}[\text{ch1}][\text{ch1}] - \text{correlación}[\text{ch2}][\text{ch2}]));$$

- 35 Por otra parte, el procesador de iteración 102 se puede configurar para calcular una correlación intercanal usando una trama de cada canal que comprende una pluralidad de bandas de modo de obtener un valor de correlación intercanal simple para la pluralidad de bandas, en donde el procesador de iteración 102 se puede configurar para realizar el procesamiento multicanal para cada una de la pluralidad de bandas de modo que los parámetros multicanal se obtengan de cada una de la pluralidad de bandas.

- 40 En este caso, el procesador de iteración 102 se puede configurar para calcular parámetros estéreo en el procesamiento multicanal, en donde el procesador de iteración 102 se puede configurar solo para realizar un procesamiento estéreo en bandas, en donde un parámetro estéreo es mayor que un umbral cuantificado a cero definido por un cuantificador estéreo (por ejemplo, codificación de rotación a base de KLT). Los parámetros estéreo pueden ser, por ejemplo, MS On/Off o ángulos de rotación o coeficientes de predicción.

- 45 Por ejemplo, el procesador de iteración 102 se puede configurar para calcular ángulos de rotación en el procesamiento multicanal, en donde el procesador de iteración 102 se puede configurar solo para realizar un procesamiento de rotación en bandas, en donde un ángulo de rotación es mayor que un umbral cuantificado a cero definido por un ángulo de rotación de cuantificador (por ejemplo, codificador de rotación a base de KLT).

- 50 De esta manera, el codificador 100 (o interfaz de salida 106) se puede configurar para transmitir la información de transformación / rotación ya sea como un parámetro para el espectro completo (caja de banda completa) o como parámetros dependientes de frecuencia múltiples para partes del espectro.

- 55 El codificador 100 se puede configurar para generar el flujo de bits 107 en función de las siguientes tablas:

Tabla 1 — Sintaxis de mpeg3daExtElementConfig()

Sintaxis	N.º de bits	Mnemotécnico
<pre> mpegh3daExtElementConfig() {     usacExtElementType      = escapedValue(4, 8, 16);     usacExtElementConfigLength = escapedValue(4, 8, 16);      si (<b>usacExtElementDefaultLengthPresent</b>) {         usacExtElementDefaultLength = escapedValue(8, 16, 0) + 1;     } de lo contrario {         usacExtElementDefaultLength = 0;     }      <b>usacExtElementPayloadFrag;</b>      switch (usacExtElementType) {     case ID_EXT_ELE_FILL:         /* No configuration element */         break;     case ID_EXT_ELE_MPEGS:         SpatialSpecificConfig();         break;     case ID_EXT_ELE_SAOC:         SAOCSpecificConfig();         break;     case ID_EXT_ELE_AUDIOPREROLL:         /* No configuration element */ </pre>	<p>1</p> <p>1</p>	<p><b>uimsbf</b></p> <p><b>uimsbf</b></p>

(continuación)

Sintaxis	N.º de bits	Mnemotécnico
<pre> break; case ID_EXT_ELE_UNI_DRC:     mpeg3daUniDrcConfig();     break; case ID_EXT_ELE_OBJ_METADATA:     ObjectMetadataConfig();     break; case ID_EXT_ELE_SAOC_3D:     SAOC3DSpecificConfig();     break; case ID_EXT_ELE_HOA:     HOAConfig();     break; case ID_EXT_ELE_MCC: /* multi channel coding */     MCCConfig(grp);     break; case ID_EXT_ELE_FMT_CNRTR     /* No configuration element */     break; default:     while (usacExtElementConfigLength—) {         tmp;     }     break; } </pre>	<p>NOTA</p> <p><b>8</b></p>	<p><b>uimbsf</b></p>
<p>NOTA: La entrada por defecto para usacExtElementType se usa para extElementTypes desconocidos, de modo que los decodificadores heredados pueden lidiar con futuras extensiones.</p>		

Tabla 21 — Sintaxis de MCCConfig(),

Sintaxis	N.º de bits	Mnemotéc.
<pre> MCCConfig(grp) {     nChannels = 0     para(chan=0;chan &lt; bsNumberOfSignals[grp]; chan++)         <b>chanMask[chan]</b>         si(chanMask[chan] &gt; 0) {             mctChannelMap[nChannels]=chan;             nChannels++;         } } </pre>	<p><b>1</b></p>	
<p>NOTA: El correspondiente elemento ID_USAC_EXT debe ser anterior a cualquier elemento de audio de cierto grupo de señales grp.</p>		

Tabla 32 — Sintaxis de MultichannelCodingBoxBandWise()

Sintaxis	N.º de bits	Mnemotéc.
<pre> MultichannelCodingBoxBandWise() {     para(pair=0; pair&lt;numPairs;pair++) {         si (keepTree == 0) {             <b>channelPairIndex[pair]</b>         }         de lo contrario {             channelPairIndex[pair]=                 lastChannelPairIndex[pair];         }     } } </pre>	<p><b>nBits</b> NOTA 1)</p>	
<p><b>hasMctMask</b></p> <p><b>hasBandwiseAngles</b></p>	<p><b>1</b></p> <p><b>1</b></p>	

(continuación)

Sintaxis	N.º de bits	Mnemotéc.
<pre> si (hasMctMask    hasBandwiseAngles) {     <b>isShort</b>     <b>numMaskBands;</b>     <b>si (isShort) {</b>         <b>numMaskBands</b>     <b>}</b> } de lo contrario {     numMaskBands MAX_NUM_MC_BANDS; } si (hasMctMask) {     para(j=0;j&lt;numMaskBands;j++) {         <b>msMask[pair][j];</b>     } de lo contrario {         para(j=0;j&lt;numMaskBands;j++) {             msMask[pair][j] = 1;         }     } } si(indepFlag &gt; 0) {     delta_code_time = 0; } de lo contrario {     <b>delta_code_time;</b> } si (hasBandwiseAngles == 0) {     <b>hcod_angle[dpcm_alpha[pair][0]];</b> } de lo contrario {     para(j=0;j&lt; numMaskBands;j++) {         si (msMask[pair][j] ==1) {             <b>hcod_angle[dpcm_alpha[pair][j]];</b>         }     } } } } </pre>	<p>1</p> <p>5</p> <p>=</p> <p>NOTA 2)</p> <p>=</p> <p>1</p> <p>1</p> <p>1..10</p> <p>1..10</p>	<p></p> <p></p> <p></p> <p></p> <p></p> <p>vclbf</p> <p>vclbf</p>
<p>NOTA 1) nBits = floor(log2(nChannels*(nChannels-1)/2 - 1))+1</p>		

Tabla 4 — Sintaxis de MultichannelCodingBoxFullband()

Sintaxis	N.º de bits	Mnemotéc.
<pre> MultichannelCodingBoxFullband() {     para (pair=0; pair&lt;numPairs; pair++) {         si(keepTree == 0) {             <b>channelPairIndex[pair]</b>         }         de lo contrario {             numPairs = lastNumPairs;         }     }     <b>alpha;</b> } </pre>	<p>nBits</p> <p>NOTA 1)</p> <p>8</p>	<p></p>
<p>NOTA: 1) nBits = floor(log2(nChannels*(nChannels-1)/2 - 1))+1</p>		

Tabla 5 — Sintaxis de MultichannelCodingFrame()

Sintaxis	N.º de bits	Mnemotéc.
MultichannelCodingFrame() { <b>MCCSignalingType</b>	<b>2</b>	

(Continuación)

Sintaxis	N.º de bits	Mnemotéc.
<b>keepTree</b>	<b>1</b>	
si(keepTree==0) { <b>numPairs</b>	<b>5</b>	
}		
de lo contrario { numPairs=lastNumPairs;		
}		
si(MCCSignalingType == 0) { /* tree of standard stereo boxes */ para(i=0;i<numPairs;i++) { MCCBox[i] = StereoCoreToolInfo(0); }		
}		
si(MCCSignalingType == 1) { /* arbitrary mct trees */ MultichannelCodingBoxBandWise();		
}		
si(MCCSignalingType == 2) { /* transmitted trees */		
}		
si(MCCSignalingType == 3) { /* simple fullband tree */ MultichannelCodingBoxFullband();		
}		
}		

Tabla 6 — Valor de usacExtElementType

usacExtElementType	Valor
ID_EXT_ELE_FILL	0
ID_EXT_ELE_MPEGS	1
ID_EXT_ELE_SAOC	2
ID_EXT_ELE_AUDIOPREROLL	3
ID_EXT_ELE_UNI_DRC	4
ID_EXT_ELE_OBJ_METADATA	5
ID_EXT_ELE_SAOC_3D	6
ID_EXT_ELE_HOA	7
ID_EXT_ELE_FMT_CNRTR	8
ID_EXT_ELE_MCC	9 o 10
/* reservado para uso ISO */	10–127
/* reservado para usar fuera del alcance de ISO */	128 y mayor
NOTA: Los valores de usacExtElementType específicos de la aplicación están obligados a estar en los espacios reservados para usar fuera del alcance de ISO. Se omiten por un decodificador ya que se requiere un mínimo de estructura por el decodificador para omitir estas extensiones.	

Tabla 7 — Interpretación de bloques de datos para extensión de la decodificación de la carga útil

<b>usacExtElementType</b>	<b>usacExtElementSegmentData concatenado representa:</b>
ID_EXT_ELE_FILL	Serie de <b>fill_byte</b>
ID_EXT_ELE_MPEGS	SpatialFrame()
ID_EXT_ELE_SAOC	SaocFrame()
ID_EXT_ELE_AUDIOPREROLL	AudioPreRoll()
ID_EXT_ELE_UNI_DRC	uniDrcGain() como se define en ISO/IEC 23003-4
ID_EXT_ELE_OBJ_METADATA	object_metadata()
ID_EXT_ELE_SAOC_3D	Saoc3DFrame()
ID_EXT_ELE_HOA	HOAFrame()
ID_EXT_ELE_FMT_CNVTR	FormatConverterFrame()
ID_EXT_ELE_MCC	MultichannelCodingFrame()
desconocido	Datos desconocidos. El bloque de datos debe ser descartado.

La Fig. 9 muestra un diagrama de bloque esquemático de un procesador de iteración 102, de acuerdo con una forma de realización. En la forma de realización mostrada en la Fig. 9, la señal multicanal 101 es una señal de canal 5.1 que tiene seis canales: un canal izquierdo L, un canal derecho R, un canal envolvente izquierdo Ls, un canal envolvente derecho Rs, un canal central C y un canal de efectos de baja frecuencia LFE.

Como se indica en la Fig. 9, el canal LFE no es procesado por el procesador de iteración 102. Esto puede ser el caso ya que los valores de correlación intercanal entre el canal LFE y cada uno de los otros cinco canales L, R, Ls, Rs y C son demasiado pequeños o dado que la máscara del canal no indica procesar el canal LFE, que se asumirá a continuación.

En una primera etapa iteración, el procesador de iteración 102 calcula los valores de correlación intercanal entre cada par de los cinco canales L, R, Ls, Rs y C, para seleccionar, en la primera etapa de iteración, un par que tiene un valor máximo o que tiene un valor por encima del umbral. En la Fig. 9 se asume que el canal izquierdo L y el canal derecho R tienen el valor más alto, de modo que el procesador de iteración 102 procese el canal izquierdo L y el canal derecho R usando una caja estéreo (o herramienta estéreo) 110, que realiza la operación de procesamiento multicanal, para derivar los primeros y segundos canales procesados P1 y P2.

En una segunda etapa de iteración, el procesador de iteración 102 calcula valores de correlación intercanal entre cada par de los cinco canales L, R, Ls, Rs y C y los canales procesados P1 y P2, para seleccionar, en la segunda etapa de iteración, un par que tiene un valor máximo o que tiene un valor por encima del umbral. En la Fig. 9 se asume que el canal envolvente izquierdo Ls y el canal envolvente derecho Rs tienen el valor más alto, de modo tal que el procesador de iteración 102 procese el canal envolvente izquierdo Ls y el canal envolvente derecho Rs usando la caja estéreo (o herramienta estéreo) 112, para derivar el tercer y el cuarto canal procesado P3 y P4.

En una tercera etapa de iteración, el procesador de iteración 102 calcula valores de correlación intercanal entre cada par de los cinco canales L, R, Ls, Rs y C y los canales procesados P1 a P4, para seleccionar, en la tercera etapa de iteración, un par que tiene un valor máximo o que tiene un valor por encima del umbral. En la Fig. 9 se asume que el primer canal procesado P1 y el tercer canal procesado P3 tienen el valor más alto, de modo tal que el procesador de iteración 102 procese el primer canal procesado P1 y el tercer canal procesado P3 usando la caja estéreo (o herramienta estéreo) 114, para derivar el quinto y sexto canal procesado P5 y P6.

En una cuarta etapa de iteración, el procesador de iteración 102 calcula valores de correlación intercanal entre cada par de los cinco canales L, R, Ls, Rs y C y los canales procesados P1 a P6, para seleccionar, en la cuarta etapa de iteración, un par que tiene un valor máximo o que tiene un valor por encima del umbral. En la Fig. 9 se asume que el quinto canal procesado P5 y el canal central C tienen el valor más alto, de modo tal que el procesador de iteración 102 procese el quinto canal procesado P5 y el canal central C usando la caja estéreo (o herramienta estéreo) 115, para derivar el séptimo y octavo canal procesado P7 y P8.

Las cajas estéreo 110 a 116 pueden ser cajas estéreo MS, es decir, cajas estereofónicas mid/side configuradas para proporcionar un canal medio y un canal lateral. El canal medio puede ser la suma de los canales de entrada de la caja estéreo, en donde el canal lateral puede ser la diferencia entre los canales de entrada de la caja estéreo. Por otra parte, las cajas estéreo 110 y 116 pueden ser cajas de rotación o cajas de predicción estéreo.

En la Fig. 9, el primer canal procesado P1, el tercer canal procesado P3 y el quinto canal procesado P5 pueden ser canales medios, en donde el segundo canal procesado P2, el cuarto canal procesado P4 y el sexto canal procesado P6 pueden ser canales laterales.

Por otra parte, como se indica en la Fig. 9, el procesador de iteración 102 se puede configurar para realizar el cálculo,

la selección y el procesamiento en la segunda etapa de iteración y, de ser aplicable, en cualquier otra etapa de iteración usando los canales de entrada L, R, Ls, Rs y C y (solo) los canales medios P1, P3 y P5 de los canales procesados. En otras palabras, el procesador de iteración 102 se puede configurar para no usar los canales laterales P1, P3 y P5 de los canales procesados en el cálculo, la selección y el procesamiento en la segunda etapa de iteración y, de ser aplicable, en cualquier otra etapa de iteración.

La Fig. 11 muestra un diagrama de flujo de un método 300 para codificar una señal multicanal que tiene al menos tres canales. El método 300 comprende una etapa 302 de cálculo, en una primera etapa de iteración, de valores de correlación intercanal entre cada par de los al menos tres canales, seleccionando, en la primera etapa de iteración, un par que tiene un valor máximo o que tiene un valor por encima del umbral y procesamiento del par seleccionado usando una operación de procesamiento multicanal para derivar parámetros multicanal MCH\_PAR1 para el par seleccionado y para derivar los primeros canales procesados; una etapa 304 de realización del cálculo, la selección y el procesamiento en una segunda etapa de iteración usando al menos uno de los canales procesados para derivar parámetros multicanal MCH\_PAR2 y segundos canales procesados; una etapa 306 de codificación de canales que resultan de un procesamiento de iteración realizado por el procesador de iteración para obtener canales codificados; y una etapa 308 de generación de una señal multicanal codificada que tiene los canales codificados y los primeros parámetros y los parámetros multicanal MCH\_PAR2.

A continuación, se explica la codificación multicanal.

La Fig. 10 muestra un diagrama de bloque esquemático de un aparato (decodificador) 200 para decodificar una señal multicanal codificada 107 que tiene canales codificados E1 a E3 y al menos dos parámetros multicanal MCH\_PAR1 y MCH\_PAR2.

El aparato 200 comprende un decodificador de canales 202 y un procesador multicanal 204.

El decodificador de canales 202 se configura para decodificar los canales codificados E1 a E3 para obtener canales decodificados en D1 a D3.

Por ejemplo, el decodificador de canales 202 puede comprender al menos tres monodecodificadores (o monocabajas o monoherramientas) 206\_1 a 206\_3, en donde cada uno de los monodecodificadores 206\_1 a 206\_3 se puede configurar para decodificar uno de los al menos tres canales codificados E1 a E3, para obtener el respectivo canal decodificado E1 a E3. Los monodecodificadores 206\_1 a 206\_3 pueden ser, por ejemplo, decodificadores de audio a base de transformación.

El procesador multicanal 204 se configura para realizar un procesamiento multicanal usando un segundo par de los canales decodificados identificados por los parámetros multicanal MCH\_PAR2 y usando los parámetros multicanal MCH\_PAR2 para obtener canales procesados y para realizar otro procesamiento multicanal usando un primer par de canales identificados por los parámetros multicanal MCH\_PAR1 y usando los parámetros multicanal MCH\_PAR1, donde el primer par de canales comprende al menos un canal procesado.

Como se indica en la Fig. 10 a modo de ejemplo, los parámetros multicanal MCH\_PAR2 pueden indicar (o señalar) que el segundo par de canales decodificados consisten en el primer canal decodificado D1 y el segundo canal decodificado D2. De esta manera, el procesador multicanal 204 realiza un procesamiento multicanal usando el segundo par de los canales decodificados que consisten en el primer canal decodificado D1 y el segundo canal decodificado D2 (identificado por los parámetros multicanal MCH\_PAR2) y usando los parámetros multicanal MCH\_PAR2, para obtener canales procesados P1\* y P2\*. Los parámetros multicanal MCH\_PAR1 pueden indicar que el primer par de canales decodificados consiste en el primer canal procesado P1\* y el tercer canal decodificado D3. De esta manera, el procesador multicanal 204 realiza el posterior procesamiento multicanal usando el primer par de canales decodificados que consiste en el primer canal procesado P1\* y el tercer canal decodificado D3 (identificado por los parámetros multicanal MCH\_PAR1) y usando los parámetros multicanal MCH\_PAR1, para obtener canales procesados P3\* y P4\*.

Por otra parte, el procesador multicanal 204 puede proporcionar el tercer canal procesado P3\* como primer canal CH1, el cuarto canal procesado P4\* como tercer canal CH3 y el segundo canal procesado P2\* como segundo canal CH2.

Asumiendo que el decodificador 200 mostrado en la Fig. 10 recibe la señal multicanal codificada 107 del codificador 100 mostrado en la Fig. 7, el primer canal decodificado D1 del decodificador 200 puede ser equivalente al tercer canal procesado P3 del codificador 100, en donde el segundo canal decodificado D2 del decodificador 200 puede ser equivalente al cuarto canal procesado P4 del codificador 100 y en donde el tercer canal decodificado D3 del decodificador 200 puede ser equivalente al segundo canal procesado P2 del codificador 100. Por otra parte, el primer canal procesado P1\* del decodificador 200 puede ser equivalente al primer canal procesado P1 del codificador 100.

Por otra parte, la señal multicanal codificada 107 puede ser una señal serial, en donde se reciben los parámetros multicanal MCH\_PAR2, en el decodificador 200, antes de los parámetros multicanal MCH\_PAR1. En este caso, el procesador multicanal 204 se puede configurar para procesar los canales decodificados en un orden en el que los parámetros multicanal MCH\_PAR1 y MCH\_PAR2 son recibidos por el decodificador. En el ejemplo mostrado en la Fig. 10, el decodificador recibe los parámetros multicanal MCH\_PAR2 antes de los parámetros multicanal MCH\_PAR1 y así, se realiza el procesamiento multicanal usando el segundo par de los canales decodificados (que consiste en el primer y el segundo canal decodificados D1 y D2) identificados por los parámetros multicanal MCH\_PAR2 antes de realizar el procesamiento multicanal usando el primer par de los canales decodificados (que consiste en el primer canal procesado P1\* y el tercer canal decodificado D3) identificados por el parámetro multicanal MCH\_PAR1.

En la Fig. 10, el procesador multicanal 204 realiza a modo de ejemplo dos operaciones de procesamiento multicanal. Con fines ilustrativos, las operaciones de procesamiento multicanal realizadas por el procesador multicanal 204 se ilustran en la Fig. 10 por cajas de procesamiento 208 y 210. Las cajas de procesamiento 208 y 210 se pueden implementar en hardware o software. Las cajas de procesamiento 208 y 210 pueden ser, por ejemplo, cajas estéreo, como se analizó con anterioridad con referencia al codificador 100, como decodificadores genéricos (o cajas estéreo del lado de los decodificadores), decodificadores a base de predicciones (o cajas estéreo del lado de los decodificadores) o decodificadores a base de rotación KLT (o cajas estéreo del lado de los decodificadores).

Por ejemplo, el codificador 100 puede usar codificadores de rotación a base de KLT (o cajas estéreo del lado de los codificadores). En ese caso, el codificador 100 puede derivar los parámetros multicanal MCH\_PAR1 y MCH\_PAR2 de modo tal que los parámetros multicanal MCH\_PAR1 y MCH\_PAR2 comprendan ángulos de rotación. Los ángulos de rotación se pueden codificar de modo diferencial. En consecuencia, el procesador multicanal 204 del decodificador 200 puede comprender un decodificador diferencial que decodifica los ángulos de rotación codificados de modo diferencial.

El aparato 200 también puede comprender una interfaz de entrada 212 configurada para recibir y procesar la señal multicanal codificada 107, para proporcionar los canales codificados E1 a E3 al decodificador de canales 202 y los parámetros multicanal MCH\_PAR1 y MCH\_PAR2 al procesador multicanal 204.

Como ya se mencionó, un indicador de mantenimiento (o bandera de árbol de mantenimiento) se puede usar para señalar que no se transmite un nuevo árbol, pero se ha de usar el último árbol estéreo. Esto se puede usar para evitar múltiple transmisión de la misma configuración de árbol estéreo si las propiedades de correlación del canal se quedan fijas durante un tiempo más prolongado.

En consecuencia, cuando la señal multicanal codificada 107 comprende, para una primera trama, los parámetros multicanal MCH\_PAR1 y MCH\_PAR2 y, para una segunda trama, después de la primera trama, el indicador de mantenimiento, el procesador multicanal 204 se puede configurar para realizar el procesamiento multicanal o el ulterior procesamiento multicanal en la segunda trama en el mismo segundo par o el mismo primer par de canales como se usan en la primera trama.

El procesamiento multicanal y el ulterior procesamiento multicanal pueden comprender un procesamiento estéreo usando un parámetro estéreo, en donde para las bandas de factor de escala individual o grupos de bandas de factor de escala de los canales decodificados D1 a D3, se incluye un primer parámetro estéreo en el parámetro multicanal MCH\_PAR1 y se incluye un segundo parámetro estéreo en el parámetro multicanal MCH\_PAR2. En este caso, el primer parámetro estéreo y el segundo parámetro estéreo pueden ser del mismo tipo, tales como ángulos de rotación o coeficientes de predicción. Naturalmente, el primer parámetro estéreo y el segundo parámetro estéreo pueden ser de diferente tipo. Por ejemplo, el primer parámetro estéreo puede ser un ángulo de rotación, en donde el segundo parámetro estéreo puede ser un coeficiente de predicción o viceversa.

Por otra parte, los parámetros multicanal MCH\_PAR1 y MCH\_PAR2 pueden comprender una máscara de procesamiento multicanal que indica qué bandas de factor de escala se procesan por multicanal y qué bandas de factor de escala no se procesan por multicanal. En este caso, el procesador multicanal 204 se puede configurar para no realizar el procesamiento multicanal en las bandas de factor de escala indicadas por la máscara de procesamiento multicanal.

Los parámetros multicanal MCH\_PAR1 y MCH\_PAR2 pueden incluir cada uno una identificación de par de canales (o índice), en donde el procesador multicanal 204 se puede configurar para decodificar las identificaciones de pares de canales (o índices) usando una regla de decodificación predefinida o una regla de decodificación indicada en la señal multicanal codificada.

Por ejemplo, los pares de canales se pueden señalar eficazmente usando un único índice para cada par, según la cantidad total de canales, como se describió con anterioridad con referencia al codificador 100.

Por otra parte, la regla de decodificación puede ser una regla de decodificación de Huffman, en donde el procesador

multicanal 204 se puede configurar para realizar una decodificación de Huffman de las identificaciones del par de canales.

La señal multicanal codificada 107 también puede comprender un indicador que permite el procesamiento multicanal que indica solo un subgrupo de los canales decodificados, para los que se permite el procesamiento multicanal y que indica al menos un canal decodificado para el que no se permite el procesamiento multicanal. En este caso, el procesador multicanal 204 se puede configurar no para realizar cualquier procesamiento multicanal para el al menos un canal decodificado, para el que no se permite el procesamiento multicanal como se indica por el indicador que permite el procesamiento multicanal.

Por ejemplo, cuando la señal multicanal es una señal de canal 5.1, el indicador que permite el procesamiento multicanal puede indicar que el procesamiento multicanal solo se permite para los 5 canales, es decir, derecho R, izquierdo L, envolvente derecho Rs, envolvente izquierdo LS y central C, en donde el procesamiento multicanal no se permite para el canal LFE.

Para el proceso de decodificación (decodificación de índices de pares de canales) se puede usar el siguiente código c. En este caso, para todos los pares de canales, se necesita la cantidad de canales con procesamiento KLT activo (nChannels) así como la cantidad de pares de canales (numPairs) de la trama actual.

```

20  maxNumPairIdx          =          nChannels*(nChannels-1)/2          -          1;
    numBits = floor(log2(maxNumPairIdx)+1);
    pairCounter = 0;

    para (chan1=1; chan1 < nChannels; chan1++) {
25     para (chan0=0; chan0 < chan1; chan0++) {
        si (pairCounter == pairIdx) {
            channelPair[0] = chan0;
            channelPair[1] = chan1;
            return;
30     }
        de lo contrario
            pairCounter++;
        }
    }
35 }

```

Para la decodificación de los coeficientes de predicción para ángulos no de banda, se puede usar el siguiente código c.

```

40  para(pair=0; pair<numPairs; pair++) {
        mctBandsPerWindow = numMaskBands[pair]/windowsPerFrame;

        si(delta_code_time[pair] > 0) {
45         lastVal = alpha_prev_fullband[pair];
        } de lo contrario
        {
            lastVal = DEFAULT_ALPHA;
        }

50     newAlpha = lastVal + dpcm_alpha[pair][0];
        si(newAlpha >= 64) {
            newAlpha -= 64;
        }

55     para (band=0; band < numMaskBands; band++){
        /* set all angles to fullband angle */
        pairAlpha[pair][band] = newAlpha;

        /* set previous angles de acuerdo con mctMask */
60     si(mctMask[pair][band] > 0) {
            alpha_prev_frame[pair][band%mctBandsPerWindow] = newAlpha;
        }
        de lo contrario
        {
65     alpha_prev_frame[pair][band%mctBandsPerWindow] = DEFAULT_ALPHA;

```

```

    }
    }
    alpha_prev_fullband[pair] = newAlpha;
    para(band=bandsPerWindow ; band<MAX_NUM_MC_BANDS; band++) {
5     alpha_prev_frame[pair][band] = DEFAULT_ALPHA;
    }
}

```

10 Para decodificar los coeficientes de predicción para los ángulos KLT no de banda, se puede usar el siguiente código c.

```

para(pair=0; pair<numPairs; pair++) {
    mctBandsPerWindow = numMaskBands[pair]/windowsPerFrame;
    para(band=0; band<numMaskBands[pair]; band++) {
15     si(delta_code_time[pair] > 0) {
        lastVal = alpha_prev_frame[pair][band%mctBandsPerWindow];
    }
    de lo contrario
    {
20     si ((band % mctBandsPerWindow) == 0) {
        lastVal = DEFAULT_ALPHA;
    }
    }
    si (msMask[pair][band] > 0 ) {
25     newAlpha = lastVal + dpcm_alpha[pair][band];
    si(newAlpha >= 64) {
        newAlpha -= 64;
    }
    pairAlpha[pair][band] = newAlpha;
    alpha_prev_frame[pair][band%mctBandsPerWindow] = newAlpha;
    lastVal = newAlpha;
    }
    de lo contrario
35     {
        alpha_prev_frame[pair][band%mctBandsPerWindow] = DEFAULT_ALPHA; /* -45° */
    }

    /* reset fullband angle */
40     alpha_prev_fullband[pair] = DEFAULT_ALPHA;
    }
    para(band=bandsPerWindow ; band<MAX_NUM_MC_BANDS; band++) {
        alpha_prev_frame[pair][band] = DEFAULT_ALPHA;
    }
45 }

```

Para evitar diferencias de puntos de flotación de funciones trigonométricas en diferentes plataformas, se han de usar las siguientes tablas de búsqueda para los índices de los ángulos de conversión directamente para sen/cos:

```

50 tabIndexToSinAlpha[64] = {
    -1.000000f,-0.998795f,-0.995185f,-0.989177f,-0.980785f,-0.970031f,-0.956940f,-
    0.941544f,
    -0.923880f,-0.903989f,-0.881921f,-0.857729f,-0.831470f,-0.803208f,-0.773010f,-
    0.740951f,
55 -0.707107f,-0.671559f,-0.634393f,-0.595699f,-0.555570f,-0.514103f,-0.471397f,-
    0.427555f,
    -0.382683f,-0.336890f,-0.290285f,-0.242980f,-0.195090f,-0.146730f,-0.098017f,-
    0.049068f,
    0.000000f, 0.049068f, 0.098017f, 0.146730f, 0.195090f, 0.242980f, 0.290285f, 0.336890f,
60 0.382683f, 0.427555f, 0.471397f, 0.514103f, 0.555570f, 0.595699f, 0.634393f, 0.671559f,
    0.707107f, 0.740951f, 0.773010f, 0.803208f, 0.831470f, 0.857729f, 0.881921f, 0.903989f,
    0.923880f, 0.941544f, 0.956940f, 0.970031f, 0.980785f, 0.989177f, 0.995185f, 0.998795f
};
tabIndexToCosAlpha[64] = {
65 0.000000f, 0.049068f, 0.098017f, 0.146730f, 0.195090f, 0.242980f, 0.290285f, 0.336890f,

```

```

0.382683f, 0.427555f, 0.471397f, 0.514103f, 0.555570f, 0.595699f, 0.634393f, 0.671559f,
0.707107f, 0.740951f, 0.773010f, 0.803208f, 0.831470f, 0.857729f, 0.881921f, 0.903989f,
0.923880f, 0.941544f, 0.956940f, 0.970031f, 0.980785f, 0.989177f, 0.995185f, 0.998795f,
1.000000f, 0.998795f, 0.995185f, 0.989177f, 0.980785f, 0.970031f, 0.956940f, 0.941544f,
5 0.923880f, 0.903989f, 0.881921f, 0.857729f, 0.831470f, 0.803208f, 0.773010f, 0.740951f,
0.707107f, 0.671559f, 0.634393f, 0.595699f, 0.555570f, 0.514103f, 0.471397f, 0.427555f,
0.382683f, 0.336890f, 0.290285f, 0.242980f, 0.195090f, 0.146730f, 0.098017f, 0.049068f
};

10 Para decodificar la codificación multicanal, se puede usar el siguiente código c para el enfoque a base de rotación
KLT.

decode_mct_rotation()
{
15 para (pair=0; pair < self->numPairs; pair++) {

    mctBandOffset = 0;

    /* inverse MCT rotation */
20 para (win = 0, group = 0; group < num_window_groups; group++) {

        para (groupwin = 0; groupwin < window_group_length[group]; groupwin++, win++) {
            *dmx = spectral_data[ch1][win];
            *res = spectral_data[ch2][win];
25 apply_mct_rotation_wrapper(self, dmx, res, &alphaSfb[mctBandOffset],
                &mctMask[mctBandOffset], mctBandsPerWindow, alpha,
                totalSfb, pair, nSamples);
        }

30 mctBandOffset += mctBandsPerWindow;
    }
}

35 Para el procesamiento de bandas, se puede usar el siguiente código c. apply_mct_rotation_wrapper(self, *dmx, *res,
*alphaSfb,
                *mctMask,
                mctBandsPerWindow,

    alpha, totalSfb, pair, nSamples)
{
40 sfb = 0;

    si (self->MCCSignalingType == 0) {
    }
    de lo contrario
45 si (self->MCCSignalingType == 1) {

        /* apply fullband box */
        si (!self->bHasBandwiseAngles[pair] && !self->bHasMctMask[pair]) {
            apply_mct_rotation(dmx, res, alphaSfb[0], nSamples);
50 }
        de lo contrario
        {
            /* apply bandwise procesamiento */
            para (i = 0; i < mctBandsPerWindow; i++) {
55 si (mctMask[i] == 1) {
                startLine = swb_offset [sfb];
                stopLine = (sfb+2<totalSfb)? swb_offset [sfb+2] : swb_offset [sfb+1];
                nSamples = stopLine-startLine;

60 apply_mct_rotation(&dmx[startLine], &res[startLine], alphaSfb[i], nSamples);
            }
            sfb += 2;

            /* break condition */
65 si (sfb >= totalSfb) {

```

```

    break;
    }
    }
    }
5   }
    de lo contrario
    si (self->MCCSignalingType == 2) {
    }
    de lo contrario
10  si (self->MCCSignalingType == 3) {
    apply_mct_rotation(dmx, res, alpha, nSamples);
    }
    }

```

15 Para una aplicación de la rotación de KLT, se puede usar el siguiente código c.

```

apply_mct_rotation(*dmx, *res, alpha, nSamples)
{
    para (n=0;n<nSamples;n++) {
20
        L = dmx[n] * tabIndexToCosAlpha [alphaIdx] - res[n] * tabIndexToSinAlpha [alphaIdx];
        R = dmx[n] * tabIndexToSinAlpha [alphaIdx] + res[n] * tabIndexToCosAlpha [alphaIdx];

        dmx[n] = L;
25  res[n] = R;
    }
}

```

30 La Fig. 12 muestra un diagrama de flujo de un método 400 para decodificar una señal multicanal codificada que tiene canales codificados y al menos dos parámetros multicanal MCH\_PAR1, MCH\_PAR2. El método 400 comprende una etapa 402 de decodificación de los canales codificados para obtener canales decodificados; y una etapa 404 de realización de un procesamiento multicanal usando un segundo par de los canales decodificados identificados por los parámetros multicanal MCH\_PAR2 y usando los parámetros multicanal MCH\_PAR2 para obtener canales procesados y realización de otro procesamiento multicanal usando un primer par de canales identificados por los parámetros multicanal MCH\_PAR1 y usando los parámetros multicanal MCH\_PAR1, en donde el primer par de canales comprende al menos un canal procesado.

A continuación, se explica el llenado estéreo en la codificación multicanal de acuerdo con formas de realización:

40 Como ya se destacó, un efecto no deseado de cuantificación espectral puede ser que la cuantificación puede resultar posiblemente en huecos espectrales. Por ejemplo, todos los valores espectrales en una banda de frecuencia particular se pueden fijar en cero en el lado del codificador como resultado de cuantificación. Por ejemplo, el valor exacto de tales líneas espectrales antes de la cuantificación puede ser relativamente bajo y la cuantificación luego se lleva a una situación en la que los valores espectrales de todas las líneas espectrales, por ejemplo, dentro de una banda de frecuencia particular se fijaron en cero. En el lado del decodificador, cuando se decodifica, esto puede llevar a huecos espectrales no deseados.

La herramienta de codificación multicanal (MCT) en MPEG-H permite la adaptación a varias dependencias intercanal pero, debido al uso de elementos de canal únicos en configuraciones de operación típicas, no permite el llenado estéreo.

50 Como puede verse en la Fig. 14, la herramienta de codificación multicanal combina los tres o más canales que se codifican en una manera jerárquica. Sin embargo, la forma de cómo la herramienta de codificación multicanal (MCT) combina los diferentes canales durante la codificación varía de trama en trama según las propiedades actuales de señales de los canales.

55 Por ejemplo, en la Fig. 14, el escenario (a), para generar una primera trama de señal de audio codificada, la herramienta de codificación multicanal (MCT) puede combinar un primer canal Ch1 y un segundo canal CH2 para obtener un primer canal de combinación (canal procesado) P1 y un segundo canal de combinación P2. A continuación, la herramienta de codificación multicanal (MCT) puede combinar el primer canal de combinación P1 y el tercer canal CH3 para obtener un tercer canal de combinación P3 y un cuarto canal de combinación P4. La herramienta de codificación multicanal (MCT) luego puede codificar el segundo canal de combinación P2, el tercer canal de combinación P3 y el cuarto canal de combinación P4 para generar la primera trama.

65 A continuación, por ejemplo, en la Fig. 14, el escenario (b), para generar una segunda trama de señal de audio codificada (temporalmente) que sucede a la primera trama de señal de audio codificada, la herramienta de codificación

multicanal (MCT) puede combinar el primer canal CH1' y el tercer canal CH3' para obtener un primer canal de combinación P1' y un segundo canal de combinación P2'. A continuación, la herramienta de codificación multicanal (MCT) puede combinar el primer canal de combinación P1' y el segundo canal CH2' para obtener un tercer canal de combinación P3' y un cuarto canal de combinación P4'. La herramienta de codificación multicanal (MCT) luego puede codificar el segundo canal de combinación P2', el tercer canal de combinación P3' y el cuarto canal de combinación P4' para generar la segunda trama.

Como puede verse de la Fig. 14, la forma en que el segundo, tercero y cuarto canal de combinación de la primera trama se generó en el escenario de la Fig. 14 (a) difiere significativamente del modo en que se generó el segundo, tercero y cuarto canal de combinación de la segunda trama, respectivamente, en el escenario de la Fig. 14 (b), ya que se usaron diferentes combinaciones de canales para generar los respectivos canales de combinación P2, P3 y P4 y P2', P3', P4', respectivamente.

Inter alia, las formas de realización de la presente invención se basan en los siguientes hallazgos:

Como puede verse en la Fig. 7 y Fig. 14, los canales de combinación P3, P4 y P2 (o P2', P3' y P4' en el escenario (b) de la Fig. 14) se alimentan en el codificador de canales 104. Inter alia, el codificador de canales 104 puede realizar, por ejemplo, una cuantificación, de modo que valores espectrales de los canales P2, P3 y P4 se pueden fijar en cero debido a la cuantificación. Las muestras espectrales en vecindad espectral se pueden codificar como una banda espectral, en donde cada banda espectral puede comprender una cantidad de muestras espectrales.

La cantidad de muestras espectrales de una banda de frecuencia puede ser diferente para distintas bandas de frecuencia. Por ejemplo, las bandas de frecuencia con un rango a menor frecuencia pueden comprender, por ejemplo, menos muestras espectrales (por ejemplo, 4 muestras espectrales) que las bandas de frecuencia en un rango de mayor frecuencia, que pueden comprender, por ejemplo, 16 muestras de frecuencia. Por ejemplo, las bandas críticas de escala Bark pueden definir las bandas de frecuencia usadas.

Puede surgir una situación particularmente indeseada cuando todas las muestras espectrales de una banda de frecuencia se fijaron en cero después de la cuantificación. Si se produce tal situación, de acuerdo con la presente invención, es aconsejable llevar a cabo un llenado estéreo. La presente invención se basa más aún en el hallazgo de que al menos se debería generar no solo ruido (pseudo)aleatorio.

En vez o además de añadir ruido (pseudo)aleatorio, de acuerdo con formas de realización de la presente invención, si, por ejemplo, en la Fig. 14, escenario (b), todos los valores espectrales de una banda de frecuencia del canal P4' se fijaron en cero, un canal de combinación que se habría generado de la misma forma o de una forma similar al canal P3' sería una base muy apropiada para generar ruido para el llenado en la banda de frecuencia que se cuantificó a cero.

Sin embargo, de acuerdo con formas de realización de la presente invención, es preferible no usar los valores espectrales del canal de combinación P3' de la trama actual / del punto en el tiempo actual como una base para el llenado de una banda de frecuencia del canal de combinación P4', que comprende solo valores espectrales que son cero, porque tanto el canal de combinación P3' como el canal de combinación P4' fueron generados en función del canal P1' y P2' y, de esta manera, usar el canal de combinación P3' del punto en el tiempo actual resultaría en un mero paneo.

Por ejemplo, si P3' es un canal medio de P1' y P2' (por ejemplo,  $P3' = 0,5 * (P1' + P2')$ ) y P4' es un canal lateral de P1' y P2' (por ejemplo,  $P4' = 0,5 * (P1' - P2')$ ), la introducción, por ejemplo, atenuada, de valores espectrales de P3' en una banda de frecuencia de P4' resultaría meramente en un paneo.

En vez de ello, se preferirá el uso de canales de un previo punto en el tiempo para generar valores espectrales para el llenado de los huecos espectrales en el canal de combinación actual P4'. De acuerdo con los hallazgos de la presente invención, una combinación de canales de una trama previa que corresponde al canal de combinación P3' de la trama actual sería una base deseable para generar muestras espectrales para el llenado de los huecos espectrales de P4'.

Sin embargo, el canal de combinación P3 que se generó en el escenario de la Fig. 10 (a) para la trama previa no corresponde al canal de combinación P3' de la trama actual, ya que el canal de combinación P3 de la trama previa fue generado de un modo diferente al canal de combinación P3' de la trama actual.

De acuerdo con los hallazgos de formas de realización de la presente invención, una aproximación del canal de combinación P3' se debería generar sobre la base de los canales reconstruidos de una trama previa en el lado del decodificador.

La Fig. 10 (a) ilustra un escenario de codificador donde los canales CH1, CH2 y CH3 se codifican para una trama previa generando E1, E2 y E3. El decodificador recibe los canales E1, E2 y E3 y reconstruye los canales CH1, CH2 y

CH3 que fueron codificados. Se puede haber producido cierta pérdida de codificación pero aún los canales generados CH1\*, CH2\* y CH3\* que se aproximan a CH1, CH2 y CH3 serán casi similares a los canales originales CH1, CH2 y CH3, de modo que  $CH1^* \approx CH1$ ;  $CH2^* \approx CH2$  y  $CH3^* \approx CH3$ . De acuerdo con formas de realización, el decodificador mantiene los canales CH1\*, CH2\* y CH3\*, generados para una trama previa en una memoria intermedia para usarlos para el llenado de ruido en una trama actual.

La Fig. 1a, que ilustra un aparato 201 para decodificar de acuerdo con formas de realización, se describe ahora con mayor detalle:

El aparato 201 de la Fig. 1a se adapta para decodificar una señal multicanal codificada previa de una trama previa para obtener tres o más canales de salida de audio previos y se configura para decodificar una señal multicanal codificada actual 107 de una trama actual para obtener tres o más canales de salida de audio actuales.

El aparato comprende una interfaz 212, un decodificador de canales 202, un procesador multicanal 204 para generar los tres o más canales de salida de audio actuales CH1, CH2, CH3 y un módulo de llenado de ruidos 220.

La interfaz 212 se adapta para recibir la señal multicanal codificada actual 107 y para recibir información secundaria que comprende primeros parámetros multicanal MCH\_PAR2.

El decodificador de canales 202 se adapta para decodificar la señal multicanal codificada actual de la trama actual para obtener un grupo de tres o más canales decodificados D1, D2, D3 de la trama actual.

El procesador multicanal 204 se adapta para seleccionar un primer par seleccionado de dos canales decodificados D1, D2 del grupo de tres o más canales decodificados D1, D2, D3 según los primeros parámetros multicanal MCH\_PAR2.

Como un ejemplo, esto se ilustra en la Fig. 1a por los dos canales D1, D2 que se alimentan en una caja (opcional) de procesamiento 208.

Más aún, el procesador multicanal 204 se adapta para generar un primer grupo de dos o más canales procesados P1\*, P2\* en función de dicho primer par seleccionado de dos canales decodificados D1, D2 para obtener un grupo actualizado de tres o más canales decodificados D3, P1\*, P2\*.

En el ejemplo, donde los dos canales D1 y D2 se alimentan en la caja (opcional) 208, dos canales procesados P1\* y P2\* se generan de los dos canales seleccionados D1 y D2. La configuración actualizada de los tres o más canales decodificados comprende entonces el canal D3 que se dejó y no se modificó y también comprende P1\* y P2\* que fueron generados a partir de D1 y D2.

Antes de que el procesador multicanal 204 genere el primer par de dos o más canales procesados P1\*, P2\* en función de dicho primer par seleccionado de dos canales decodificados D1, D2, el módulo de llenado de ruidos 220 se adapta para identificar al menos uno de los dos canales de dicho primer par seleccionado de dos canales decodificados D1, D2, una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero y para generar un canal de mezcla usando dos o más, pero no todos los tres o más canales de salida de audio previos y para llenar las líneas espectrales de la una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero, con ruido generado usando líneas espectrales del canal de mezcla, en donde el módulo de llenado de ruidos 220 se adapta para seleccionar los dos o más canales de salida de audio previos que se usan para generar el canal de mezcla de los tres o más canales de salida de audio previos según la información secundaria.

De esta manera, el módulo de llenado de ruidos 220 analiza si hay bandas de frecuencia que solo tienen valores espectrales que son cero y además llena las bandas de frecuencia vacías halladas con ruido generado. Por ejemplo, una banda de frecuencia puede tener, por ejemplo, 4 u 8 o 16 líneas espectrales y cuando todas las líneas espectrales de una banda de frecuencia se cuantificaron a cero, entonces el módulo de llenado de ruidos 220 llena el ruido generado.

Un concepto particular de formas de realización que se pueden emplear por el módulo de llenado de ruidos 220 que especifica cómo generar y llenar ruido se menciona como llenado estéreo.

En las formas de realización de la Fig. 1a, el módulo de llenado de ruidos 220 interactúa con el procesador multicanal 204. Por ejemplo, en una forma de realización, cuando el módulo de llenado de ruidos pretende procesar dos canales, por ejemplo, por una caja de procesamiento, alimenta estos canales al módulo de llenado de ruidos 220 y el módulo de llenado de ruidos 220 controla si las bandas de frecuencia se cuantificaron a cero y llena tales bandas de frecuencia, si se detectan.

En otras formas de realización ilustradas por la Fig. 1b, el módulo de llenado de ruidos 220 interactúa con el decodificador de canales 202. Por ejemplo, ya cuando el decodificador de canales decodifica la señal multicanal

codificada para obtener los tres o más canales decodificados D1, D2 y D3, el módulo de llenado de ruidos puede controlar, por ejemplo, si las bandas de frecuencia fueron cuantificadas a cero, y, por ejemplo, llenar tales bandas de frecuencia, si se detecta. En tal forma de realización, el procesador multicanal 204 puede asegurar que todos los huecos espectrales fueron cerrados antes por el llenado de ruido.

5 En otras formas de realización (no mostradas), el módulo de llenado de ruidos 220 puede interactuar con el decodificador de canales y el procesador multicanal. Por ejemplo, cuando el decodificador de canales 202 genera los canales decodificados D1, D2 y D3, el módulo de llenado de ruidos 220 puede controlar ya si las bandas de frecuencia fueron cuantificadas a cero, justo después de que el decodificador de canales 202 las haya generado, pero solo puede  
10 generar el ruido y llenar las respectivas bandas de frecuencia, cuando el procesador multicanal 204 procesa realmente estos canales.

Por ejemplo, ruido aleatorio, una operación económica computacional se puede insertar en cualquiera de las bandas de frecuencia cuantificadas a cero, pero el módulo de llenado de ruidos puede llenar el ruido que se generó de los  
15 canales de salida de audio previamente generados solo si se procesan realmente por el procesador multicanal 204. En tales formas de realización, sin embargo, antes de insertar ruido aleatorio, se debería hacer una detección de si existen huecos espectrales antes de insertar ruido aleatorio y esa información se debería mantener en la memoria, porque, después de insertar ruido aleatorio, las respectivas bandas de frecuencia tienen valores espectrales diferentes de cero, porque se insertó el ruido aleatorio.

20 En formas de realización, se inserta ruido aleatorio en las bandas de frecuencia que fueron cuantificadas a cero además del ruido generado en función de las señales de salida de audio previas.

En algunas formas de realización, la interfaz 212 se puede adaptar, por ejemplo, para recibir la señal multicanal codificada actual 107 y para recibir la información secundaria que comprende los primeros parámetros multicanal  
25 MCH\_PAR2 y los segundos parámetros multicanal MCH\_PAR1.

El procesador multicanal 204 se puede adaptar, por ejemplo, para seleccionar un segundo par seleccionado de dos canales decodificados P1\*, D3 del grupo actualizado de tres o más canales decodificados D3, P1\*, P2\* según los  
30 segundos parámetros multicanal MCH\_PAR1, en donde al menos un canal P1\* del segundo par seleccionado de dos canales decodificados (P1\*, D3) es un canal del primer par de dos o más canales procesados P1\*, P2\* y el procesador multicanal 204 se puede adaptar, por ejemplo, para generar un segundo grupo de dos o más canales procesados P3\*, P4\* en función de dicho segundo par seleccionado de dos canales decodificados P1\*, D3 para actualizar luego el grupo actualizado de tres o más canales decodificados.

35 Un ejemplo para tal forma de realización puede verse en las Figs. 1a y 1b, donde la caja (opcional) de procesamiento 210 recibe el canal D3 y el canal procesado P1\* y los procesa para obtener canales procesados P3\* y P4\* de modo que la configuración luego actualizada de los tres canales decodificados comprende P2\*, que no fue modificado por la caja de procesamiento 210 y los P3\* y P4\* generados.

40 Las cajas de procesamiento 208 y 210 fueron marcadas en la Fig. 1a y Fig. 1b como opcionales. Esto es para mostrar que, a pesar de que es una posibilidad usar las cajas de procesamiento 208 y 210 para implementar el procesador multicanal 204, existen varias otras posibilidades de cómo implementar exactamente el procesador multicanal 204. Por ejemplo, en vez de usar una caja de procesamiento diferente 208, 210 para cada canal diferente de procesamiento de dos (o más) canales, se puede reutilizar la misma caja de procesamiento o el procesador multicanal 204 puede  
45 implementar el procesamiento de dos canales sin usar cajas de procesamiento 208, 210 (como subunidades del procesador multicanal 204) en absoluto.

De acuerdo con otra forma de realización, el procesador multicanal 204 se puede adaptar, por ejemplo, para generar  
50 el primer grupo de dos o más canales procesados P1\*, P2\* generando un primer grupo de exactamente dos canales procesados P1\*, P2\* en función de dicho primer par seleccionado de dos canales decodificados D1, D2. El procesador multicanal 204 se puede adaptar, por ejemplo, para reemplazar dicho primer par seleccionado de dos canales decodificados D1, D2 en el grupo de tres de más canales decodificados D1, D2, D3 por el primer grupo de exactamente dos canales procesados P1\*, P2\* para obtener el grupo actualizado de tres o más canales decodificados D3, P1\*, P2\*.  
55 El procesador multicanal 204 se puede adaptar, por ejemplo, para generar el segundo grupo de dos o más canales procesados P3\*, P4\* generando un segundo grupo de exactamente dos canales procesados P3\*, P4\* en función de dicho segundo par seleccionado de dos canales decodificados P1\*, D3. Por otra parte, el procesador multicanal 204 se puede adaptar, por ejemplo, para reemplazar dicho segundo par seleccionado de dos canales decodificados P1\*, D3 en el grupo actualizado de tres de más canales decodificados D3, P1\*, P2\* por el segundo grupo de exactamente  
60 dos canales procesados P3\*, P4\* para actualizar luego el grupo actualizado de tres o más canales decodificados.

En tal forma de realización, de los dos canales seleccionados (por ejemplo, los dos canales de entrada de una caja de procesamiento 208 o 210) se generan exactamente dos canales procesados y estos exactamente dos canales procesados reemplazan a los canales seleccionados en el grupo de los tres o más canales decodificados. Por ejemplo,

la caja de procesamiento 208 del procesador multicanal 204 reemplaza los canales seleccionados D1 y D2 por P1\* y P2\*.

5 Sin embargo, en otras formas de realización, puede tener lugar una mezcla en el aparato 201 para decodificar y más de dos canales procesados se pueden generar de los dos canales seleccionados o no todos los canales seleccionados se pueden suprimir del grupo actualizado de canales decodificados.

10 Otro tema es cómo generar el canal de mezcla que se usa para generar el ruido siendo generada por el módulo de llenado de ruidos 220.

De acuerdo con algunas formas de realización, el módulo de llenado de ruidos 220 se puede adaptar, por ejemplo, para generar el canal de mezcla usando exactamente dos de los tres o más canales de salida de audio previos como los dos o más de los tres o más canales de salida de audio previos; en donde el módulo de llenado de ruidos 220 se puede adaptar, por ejemplo, para seleccionar los exactamente dos canales de salida de audio previos de los tres o más canales de salida de audio previos según la información secundaria.

Usar solo dos de los tres o más canales de salida previos ayuda a reducir la complejidad computacional de cálculo del canal de mezcla.

20 Sin embargo, en otras formas de realización, más de dos canales de los canales de salida de audio previos se usan para generar un canal de mezcla, pero la cantidad de canales de salida de audio previos que se tienen en cuenta es menor que la cantidad total de los tres o más canales de salida de audio previos.

25 En formas de realización, donde solo dos de los canales de salida previos se tienen en cuenta, el canal de mezcla se puede calcular, por ejemplo, de la siguiente manera:

En una forma de realización, el módulo de llenado de ruidos 220 se adapta para generar el canal de mezcla usando exactamente dos canales de salida de audio previos en función de la fórmula

$$30 \quad D_{ch} = (\hat{O}_1 + \hat{O}_2) \cdot d$$

o en función de la fórmula

$$35 \quad D_{ch} = (\hat{O}_1 - \hat{O}_2) \cdot d$$

en donde  $D_{ch}$  es el canal de mezcla; en donde  $\hat{O}_1$  es un primero de los exactamente dos canales de salida de audio previos; en donde  $\hat{O}_2$  es un segundo de los exactamente dos canales de salida de audio previos, siendo diferente del primero de los exactamente dos canales de salida de audio previos y en donde  $d$  es un escalar positivo real.

40 En situaciones típicas, un canal medio  $D_{ch} = (\hat{O}_1 + \hat{O}_2) \cdot d$  puede ser un canal de mezcla apropiado. Tal método calcula el canal de mezcla como un canal medio de los dos canales de salida de audio previos que se tienen en cuenta.

Sin embargo, en algunos escenarios, un canal de mezcla cercano a cero puede ocurrir cuando se aplica  $D_{ch} = (\hat{O}_1 + \hat{O}_2) \cdot d$ , por ejemplo, cuando  $\hat{O}_1 \approx -\hat{O}_2$ . A continuación, puede ser preferible usar, por ejemplo,

45  $D_{ch} = (\hat{O}_1 - \hat{O}_2) \cdot d$  como la señal de mezcla. De esta manera, entonces, se usa un canal lateral (para canales de entrada fuera de fase).

De acuerdo con un método alternativo, el módulo de llenado de ruidos 220 se adapta para generar el canal de mezcla usando exactamente dos canales de salida de audio previos en función de la fórmula

$$50 \quad \hat{I}_{ch} = (\cos \alpha \cdot \hat{O}_1 + \sin \alpha \cdot \hat{O}_2) \cdot d$$

o en función de la fórmula

$$55 \quad \hat{I}_{ch} = (-\sin \alpha \cdot \hat{O}_1 + \cos \alpha \cdot \hat{O}_2) \cdot d$$

en donde  $\hat{I}_{ch}$  es el canal de mezcla, en donde  $\hat{O}_1$  es un primero de los exactamente dos canales de salida de audio previos, en donde  $\hat{O}_2$  es un segundo de los exactamente dos canales de salida de audio previos, siendo diferente del primero de los exactamente dos canales de salida de audio previos y en donde  $\alpha$  es un ángulo de rotación.

5 Este enfoque calcula el canal de mezcla realizando una rotación de los dos canales de salida de audio previos que se tienen en cuenta.

El ángulo de rotación  $\alpha$  puede estar, por ejemplo, en el rango de:  $-90^\circ < \alpha < 90^\circ$ .

10 En una forma de realización, el ángulo de rotación puede estar, por ejemplo, en el rango de:  $30^\circ < \alpha < 60^\circ$ .

Nuevamente, en situaciones típicas, un canal  $\hat{I}_{ch} = (\cos \alpha \cdot \hat{O}_1 + \sin \alpha \cdot \hat{O}_2) \cdot d$  puede ser un canal de mezcla apropiado. Tal enfoque calcula el canal de mezcla como un canal medio de los dos canales de salida de audio previos que se tienen en cuenta.

15 Sin embargo, en algunos escenarios, un canal de mezcla cercano a cero puede ocurrir cuando se aplica

$\hat{I}_{ch} = (\cos \alpha \cdot \hat{O}_1 + \sin \alpha \cdot \hat{O}_2) \cdot d$ , por ejemplo, cuando  $\cos \alpha \cdot \hat{O}_1 \approx -\sin \alpha \cdot \hat{O}_2$ . A continuación, puede

20 ser preferible usar, por ejemplo,  $\hat{I}_{ch} = (-\sin \alpha \cdot \hat{O}_1 + \cos \alpha \cdot \hat{O}_2) \cdot d$  como la señal de mezcla.

De acuerdo con una forma de realización particular, la información secundaria puede ser, por ejemplo, información secundaria actual que se asigna a la trama actual, en donde la interfaz 212 se puede adaptar, por ejemplo, para recibir información secundaria previa que se asigna a la trama previa, en donde la información secundaria previa comprende un ángulo previo; en donde la interfaz 212 se puede adaptar, por ejemplo, para recibir la información secundaria actual que comprende un ángulo actual y en donde el módulo de llenado de ruidos 220 se puede adaptar, por ejemplo, para usar el ángulo actual de la información secundaria actual como el ángulo de rotación  $\alpha$  y se adapta para no usar el ángulo previo de la información secundaria previa como el ángulo de rotación  $\alpha$ .

25 De esta manera, en tal forma de realización, incluso si el canal de mezcla se calcula en función de canales de salida de audio previos, aún, el ángulo actual que se transmite en la información secundaria se usa como el ángulo de rotación y no un ángulo de rotación previamente recibido, a pesar de que se calcula el canal de mezcla en función de canales de salida de audio previos que fueron generados en función de una trama previa.

Otro aspecto de algunas formas de realización de la presente invención se refiere a factores de escala.

35 Las bandas de frecuencia pueden ser, por ejemplo, bandas de factor de escala.

De acuerdo con algunas formas de realización, antes de que el procesador multicanal 204 genere el primer par de dos o más canales procesados P1\*,P2\* en función de dicho primer par seleccionado de dos canales decodificados (D1, D2), el módulo de llenado de ruidos (220) se puede adaptar, por ejemplo, para identificar al menos uno de los dos canales de dicho primer par seleccionado de dos canales decodificados D1, D2, siendo una o más bandas de factor de escala las una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero y se puede adaptar, por ejemplo, para generar el canal de mezcla usando dichos dos o más, pero no todos los tres o más canales de salida de audio previos y para llenar las líneas espectrales de las una o más bandas de factor de escala, dentro de las que todas las líneas espectrales se cuantifican a cero, con el ruido generado usando las líneas espectrales del canal de mezcla según un factor de escala de cada una de las una o más bandas de factor de escala dentro de las que todas las líneas espectrales se cuantifican a cero.

45 En tales formas de realización, un factor de escala se puede asignar, por ejemplo, a cada una de las bandas de factor de escala y el factor de escala se toma en cuenta cuando se genera el ruido usando el canal de mezcla.

50 En una forma de realización particular, la interfaz de recepción 212 se puede configurar, por ejemplo, para recibir el factor de escala de cada una de dichas una o más bandas de factor de escala y el factor de escala de cada una de dichas una o más bandas de factor de escala indica una energía de las líneas espectrales de dicha banda de factor de escala antes de la cuantificación. El módulo de llenado de ruidos 220 se puede adaptar, por ejemplo, para generar el ruido para cada de las una o más bandas de factor de escala, dentro de las que todas las líneas espectrales se cuantifican a cero, de modo que una energía de las líneas espectrales después de añadir el ruido en una de las bandas de frecuencia corresponde a la energía indicada por el factor de escala para dicha banda de factor de escala.

Por ejemplo, un canal de mezcla puede indicar valores espectrales para cuatro líneas espectrales de una banda de factor de escala en donde se ha de insertar ruido y estos valores espectrales pueden ser, por ejemplo: 0,2; 0,3; 0,5; 0,1.

5 Una energía de esa banda de factor de escala del canal de mezcla se puede calcular, por ejemplo, de la siguiente manera:

$$(0,2)^2 + (0,3)^2 + (0,5)^2 + (0,1)^2 = 0,39$$

10 Sin embargo, el factor de escala para esa banda de factor de escala del canal en donde se ha de llenar ruido, por ejemplo, puede ser solo de 0,0039.

15 Un factor de atenuación se puede calcular, por ejemplo, de la siguiente manera:

$$\text{factor de atenuación} = \frac{\text{energía indicada por factor de escala}}{\text{energía de canal de mezcla}}$$

20 De esta manera, en el ejemplo anterior,

$$\text{factor de atenuación} = \frac{0,0039}{0,39} = 0,01$$

En una forma de realización, cada uno de los valores espectrales de la banda del factor de escala del canal de mezcla que se ha de usar como ruido, se multiplica con el factor de atenuación:

25 De esta manera, cada uno de los cuatro valores espectrales de la banda del factor de escala del ejemplo anterior se multiplica por el factor de atenuación y resulta en valores espectrales atenuados:

$$\begin{aligned} 0,2 \cdot 0,01 &= 0,002 \\ 0,3 \cdot 0,01 &= 0,003 \\ 0,5 \cdot 0,01 &= 0,005 \\ 0,1 \cdot 0,01 &= 0,001 \end{aligned}$$

30 Estos valores espectrales atenuados, por ejemplo, se pueden insertar luego en la banda del factor de escala del canal en el que se ha de llenar ruido.

35 El ejemplo anterior es igualmente aplicable sobre valores logarítmicos al reemplazar las operaciones anteriores por sus correspondientes operaciones logarítmicas, por ejemplo, reemplazando la multiplicación por la adición, etc.

40 Más aún, además de la descripción de formas de realización particulares proporcionadas con anterioridad, otras formas de realización del módulo de llenado de ruidos 220 aplican uno, varios o todos los conceptos descritos con referencia a la Fig. 2 a Fig. 6.

45 Otro aspecto de formas de realización de la presente invención se refiere a la cuestión en función de la que los canales de información de los canales de salida de audio previos se seleccionan para su uso para generar el canal de mezcla para obtener el ruido por insertar.

De acuerdo con una forma de realización, el aparato de acuerdo con el módulo de llenado de ruidos 220 se puede adaptar, por ejemplo, para seleccionar los exactamente dos canales de salida de audio previos de los tres o más canales de salida de audio previos según los primeros parámetros multicanal MCH\_PAR2.

50 De esta manera, en tal forma de realización, los primeros parámetros multicanal que dominan qué canales se han de seleccionar para ser procesados, también dominan cuáles de los canales de salida de audio previos se han de usar para generar el canal de mezcla para generar el ruido por insertar.

55 En una forma de realización, los primeros parámetros multicanal MCH\_PAR2 pueden indicar, por ejemplo, dos canales decodificados D1, D2 del grupo de tres o más canales decodificados; y el procesador multicanal 204 se adapta para seleccionar el primer par seleccionado de dos canales decodificados D1, D2 del grupo de tres o más canales decodificados D1, D2, D3 seleccionando los dos canales decodificados D1, D2 indicados por los primeros parámetros multicanal MCH\_PAR2. Más aún, los segundos parámetros multicanal MCH\_PAR1 pueden indicar, por ejemplo, dos

canales decodificados P1\*, D3 del grupo actualizado de tres o más canales decodificados. El procesador multicanal 204 se puede adaptar, por ejemplo, para seleccionar el segundo par seleccionado de dos canales decodificados P1\*, D3 del grupo actualizado de tres o más canales decodificados D3, P1\*, P2\* seleccionando los dos canales decodificados P1\*, D3 indicados por los segundos parámetros multicanal MCH\_PAR1.

5 De esta manera, en tal forma de realización, los canales que se seleccionan para el primer procesamiento, por ejemplo, el procesamiento de la caja de procesamiento 208 en la Fig. 1a o Fig. 1b no solo dependen de los primeros parámetros multicanal MCH\_PAR2. Más que eso, estos dos canales seleccionados se especifican explícitamente en los primeros parámetros multicanal MCH\_PAR2.

10 Del mismo modo, en tal forma de realización, los canales que se seleccionan para el segundo procesamiento, por ejemplo, el procesamiento de la caja de procesamiento 210 en la Fig. 1a o Fig. 1b no solo dependen de los segundos parámetros multicanal MCH\_PAR1. Más que eso, estos dos canales seleccionados se especifican explícitamente en los segundos parámetros multicanal MCH\_PAR1.

15 Las formas de realización de la presente invención introducen un esquema de indexación sofisticado para los parámetros multicanal que se explican con referencia a la Fig. 15.

20 La Fig. 15 (a) muestra una codificación de cinco canales, a saber, los canales izquierdo, derecho, central, envolvente izquierdo y envolvente derecho, en un lado del codificador. La Fig. 15 (b) muestra una decodificación de los canales codificados E0, E1, E2, E3, E4 para reconstruir los canales izquierdo, derecho, central, envolvente izquierdo y envolvente derecho.

25 Se asume que un índice se asigna a cada uno de los cinco canales izquierdo, derecho, central, envolvente izquierdo y envolvente derecho, a saber,

	Índice	Nombre del canal
	0	Izquierdo
	1	Derecho
30	2	Central
	3	Izquierdo envolvente
	4	Derecho envolvente

35 En la Fig. 15 (a), en el lado del codificador, la primera operación que se lleva a cabo puede ser, por ejemplo, la mezcla del canal 0 (izquierdo) y el canal 3 (izquierdo envolvente) en la caja de procesamiento 192 para obtener dos canales procesados. Se puede asumir que uno de los canales procesados es un canal medio y el otro canal es un canal lateral. Sin embargo, también se pueden aplicar otros conceptos de formación de dos canales procesados, por ejemplo, determinando los dos canales procesados al realizar una operación de rotación.

40 Ahora, los dos canales procesados generados obtienen los mismos índices que los índices de los canales que se usaron para el procesamiento. A saber, un primero de los canales procesados tiene índice 0 y un segundo de los canales procesados tiene índice 3. Los parámetros multicanal determinados para este procesamiento pueden ser, por ejemplo, (0; 3).

45 La segunda operación en el lado del codificador que se lleva a cabo puede ser, por ejemplo, la mezcla del canal 1 (Derecho) y el canal 4 (Derecho envolvente) en la caja de procesamiento 194 para obtener otros dos canales procesados. Nuevamente, los otros dos canales procesados generados obtienen los mismos índices que los índices de los canales que se usaron para el procesamiento. A saber, un primer de los demás canales procesados tiene índice 1 y un segundo de los canales procesados tiene índice 4. Los parámetros multicanal determinados para este procesamiento pueden ser, por ejemplo, (1; 4).

50 La tercera operación en el lado del codificador que se lleva a cabo puede ser, por ejemplo, la mezcla del canal procesado 0 y el canal procesado 1 en la caja de procesamiento 196 para obtener otros dos canales procesados. Nuevamente, estos dos canales procesados generados obtienen los mismos índices que los índices de los canales que se usaron para el procesamiento. A saber, un primer de los demás canales procesados tiene índice 0 y un segundo de los canales procesados tiene índice 1. Los parámetros multicanal determinados para este procesamiento pueden ser, por ejemplo, (0; 1).

60 Los canales codificados E0, E1, E2, E3 y E4 se distinguen por sus índices, a saber, E0 tiene índice 0, E1 tiene índice 1, E2 tiene índice 2, etc.

Las tres operaciones en el lado del codificador resultan en los tres parámetros multicanal: (0; 3), (1; 4), (0; 1).

Como el aparato para decodificar debe realizar operaciones de codificador en orden inverso, el orden de los parámetros multicanal puede ser invertido, por ejemplo, cuando se transmiten al aparato para decodificar, dando como resultado los parámetros multicanal:  
(0; 1), (1; 4), (0; 3).

5 Para el aparato para decodificar, (0; 1) se puede referir a los primeros parámetros multicanal, (1; 4) se puede referir a los segundos parámetros multicanal y (0; 3) se puede referir a los terceros parámetros multicanal.

10 En el lado del decodificador mostrado en la Fig. 15 (b), a partir de la recepción de los primeros parámetros multicanal (0; 1), el aparato para decodificar concluye que, como una primera operación de procesamiento en el lado del decodificador, se han de procesar los canales 0 (E0) y 1 (E1). Esto se lleva a cabo en la caja 296 de la Fig. 15 (b). Ambos canales procesados generados heredan los índices de los canales E0 y E1 que fueron usados para generarlos y, de esta manera, los canales procesados generados también tienen los índices 0 y 1.

15 A partir de la recepción de los segundos parámetros multicanal (1; 4), el aparato para decodificar concluye que, como una segunda operación de procesamiento en el lado del decodificador, se han de procesar el canal procesado 1 y el canal 4 (E4). Esto se lleva a cabo en la caja 294 de la Fig. 15 (b). Ambos canales procesados generados heredan los índices de los canales 1 y 4 que fueron usados para generarlos y, de esta manera, los canales procesados generados también tienen los índices 1 y 4.

20 A partir de la recepción de los terceros parámetros multicanal (0; 3), el aparato para decodificar concluye que, como una tercera operación de procesamiento en el lado del decodificador, se han de procesar el canal procesado 0 y el canal 3 (E3). Esto se lleva a cabo en la caja 292 de la Fig. 15 (b). Ambos canales procesados generados heredan los índices de los canales 0 y 3 que fueron usados para generarlos y, de esta manera, los canales procesados generados también tienen los índices 0 y 3.

Como resultado del procesamiento del aparato para decodificar, se reconstruyen los canales izquierdo (índice 0), derecho (índice 1), central (índice 2), izquierdo envolvente (índice 3) y derecho envolvente (índice 4).

30 Supongamos que, en el lado del decodificador, debido a la cuantificación, todos los valores del canal E1 (índice 1) dentro de cierta banda de factor de escala fueron cuantificados a cero. Cuando el aparato para decodificar quiere realizar el procesamiento en la caja 296, se desea un canal lleno de ruido 1 (canal E1).

35 Como ya se destacó, ahora las formas de realización usan dos señales de salida de audio previas para el llenado de ruido del hueco espectral del canal 1.

40 En una forma de realización particular, si un canal con el que se ha de realizar una operación tiene bandas de factor de escala que se cuantifican a cero, entonces los dos canales de salida de audio previos se usan para generar el ruido que tienen el mismo número de índice que los dos canales con los que se ha de llevar a cabo el procesamiento. En el ejemplo, si un hueco espectral del canal 1 se detecta antes del procesamiento en la caja de procesamiento 296, entonces los canales de salida de audio previos que tienen el índice 0 (canal izquierdo previo) y que tienen el índice 1 (canal derecho previo) se usan para generar ruido para llenar el hueco espectral del canal 1 en el lado del decodificador.

45 Como los índices se heredan de modo consistente por los canales procesados que resultan de un procesamiento, se puede asumir que los canales de salida previos podrían desempeñar un papel para generar los canales que son parte del procesamiento real del lado del decodificador, si los canales de salida de audio previos pueden ser los canales de salida de audio actuales. De esta manera, se puede lograr una buena estimación para la banda de factor de escala que se cuantificó a cero.

50 De acuerdo con formas de realización el aparato se puede adaptar, por ejemplo, para asignar un identificador de un grupo de identificadores a cada canal de salida de audio previo de los tres o más canales de salida de audio previos, de modo que cada canal de salida de audio previo de los tres o más canales de salida de audio previos se asigne exactamente a un identificador del grupo de identificadores y de modo que cada identificador del grupo de  
55 identificadores se asigne a exactamente un canal de salida de audio previo de los tres o más canales de salida de audio previos. Más aún, el aparato se puede adaptar, por ejemplo, para asignar un identificador de dicho grupo de identificadores a cada canal del grupo de los tres o más canales decodificados, de modo que cada canal del grupo de los tres o más canales decodificados se asigne exactamente a un identificador del grupo de identificadores y de modo que cada identificador del grupo de identificadores se asigne a exactamente un canal del grupo de los tres o más  
60 canales decodificados.

Por otra parte, los primeros parámetros multicanal MCH\_PAR2 pueden indicar, por ejemplo, un primer par de dos identificadores del grupo de los tres o más identificadores. El procesador multicanal 204 se puede adaptar, por ejemplo, para seleccionar el primer par seleccionado de dos canales decodificados D1, D2 del grupo de tres o más canales

decodificados D1, D2, D3 seleccionando los dos canales decodificados D1, D2 asignados a los dos identificadores del primer par de dos identificadores.

5 El aparato se puede adaptar, por ejemplo, para asignar un primero de los dos identificadores del primer par de dos identificadores a un primer canal procesado del primer grupo de exactamente dos canales procesados P1\*,P2\*. Más aún, el aparato se puede adaptar, por ejemplo, para asignar un segundo de los dos identificadores del primer par de dos identificadores a un segundo canal procesado del primer grupo de exactamente dos canales procesados P1\*,P2\*.

10 El grupo de identificadores pueden ser, por ejemplo, un grupo de índices, por ejemplo, un grupo de números enteros no negativos (por ejemplo, un grupo que comprende los identificadores 0; 1; 2; 3 y 4).

15 En formas de realización particulares, los segundos parámetros multicanal MCH\_PAR1 pueden indicar, por ejemplo, un segundo par de dos identificadores del grupo de los tres o más identificadores. El procesador multicanal 204 se puede adaptar, por ejemplo, para seleccionar el segundo par seleccionado de dos canales decodificados P1\*, D3 del grupo actualizado de tres o más canales decodificados D3, P1\*, P2\* seleccionando los dos canales decodificados (D3, P1\*) asignados a los dos identificadores del segundo par de dos identificadores. Más aún, el aparato se puede adaptar, por ejemplo, para asignar un primero de los dos identificadores del segundo par de dos identificadores a un primer canal procesado del segundo grupo de exactamente dos canales procesados P3\*, P4\*. Por otra parte, el aparato se puede adaptar, por ejemplo, para asignar un segundo de los dos identificadores del segundo par de dos identificadores a un segundo canal procesado del segundo grupo de exactamente dos canales procesados P3\*, P4\*.

25 En una forma de realización particular, los primeros parámetros multicanal MCH\_PAR2 pueden indicar, por ejemplo, dicho primer par de dos identificadores del grupo de los tres o más identificadores. El módulo de llenado de ruidos 220 se puede adaptar, por ejemplo, para seleccionar los exactamente dos canales de salida de audio previos de los tres o más canales de salida de audio previos seleccionando los dos canales de salida de audio previos asignados a los dos identificadores de dicho primer par de dos identificadores.

30 Como ya se destacó, la Fig. 7 ilustra un aparato 100 para codificar una señal multicanal 101 que tiene al menos tres canales (CH1:CH3) de acuerdo con una forma de realización.

35 El aparato comprende un procesador de iteración 102 que se adapta para calcular, en una primera etapa de iteración, valores de correlación intercanal entre cada par de los al menos tres canales (CH:CH3), para seleccionar, en la primera etapa de iteración, un par que tiene un valor máximo o que tiene un valor por encima del umbral y para procesar el par seleccionado usando una operación de procesamiento multicanal 110,112 para derivar parámetros multicanal iniciales MCH\_PAR1 para el par seleccionado y para derivar primeros canales procesados P1, P2.

40 El procesador de iteración 102 se adapta para realizar el cálculo, la selección y el procesamiento en una segunda etapa de iteración usando al menos uno de los canales procesados P1 para derivar otros parámetros multicanal MCH\_PAR2 y segundos canales procesados P3, P4.

Más aún, el aparato comprende un codificador de canales que se adapta para codificar canales (P2:P4) que resultan de un procesamiento de iteración realizado por el procesador de iteración 104 para obtener canales codificados (E1:E3).

45 Por otra parte, el aparato comprende una interfaz de salida 106 que se adapta para generar una señal multicanal codificada 107 que tiene los canales codificados (E1:E3), los parámetros multicanal iniciales y los otros parámetros multicanal MCH\_PAR1, MCH\_PAR2.

50 Más aún, el aparato comprende una interfaz de salida 106 que se adapta para generar la señal multicanal codificada 107 para comprender una información que indica si un aparato para decodificar ha de rellenar o no líneas espectrales de una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero, con ruido generado en función de canales de salida de audio previamente decodificados que fueron previamente decodificados por el aparato para decodificar.

55 De esta manera, el aparato para codificar es capaz de señalar si un aparato para decodificar ha de llenar o no líneas espectrales de una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero, con ruido generado en función de canales de salida de audio previamente decodificados que fueron previamente decodificados por el aparato para decodificar.

60 De acuerdo con una forma de realización, cada uno de los parámetros multicanal iniciales y los otros parámetros multicanal MCH\_PAR1, MCH\_PAR2 indican exactamente dos canales, donde cada uno de los exactamente dos canales es uno de los canales codificados (E1:E3) o es uno de los primeros o los segundos canales procesados P1, P2, P3, P4 o es uno de los al menos tres canales (CH1:CH3).

La interfaz de salida 106 se puede adaptar, por ejemplo, para generar la señal multicanal codificada 107, de modo que la información que indica si un aparato para decodificar ha de rellenar o no líneas espectrales de una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero, comprende información que indica para cada uno de los parámetros iniciales y multicanal MCH\_PAR1, MCH\_PAR2, ya sea o no para al menos un canal de los exactamente dos canales que se indican por dichos uno de los parámetros iniciales y los otros parámetros multicanal MCH\_PAR1, MCH\_PAR2, si el aparato para decodificar ha de llenar líneas espectrales de una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero, de dicho al menos un canal, con los datos espectrales generados a base de los canales de salida de audio previamente decodificados que fueron previamente decodificados por el aparato para decodificar.

Por otra parte, más abajo, se describen formas de realización particulares donde tal información se transmite usando un valor hasStereoFilling[pair] que indica si el llenado estéreo se ha de aplicar o no en un par de canales MCT actualmente procesados.

La Fig. 13 ilustra un sistema de acuerdo con formas de realización.

El sistema comprende un aparato 100 para codificar como se describió con anterioridad y un aparato 201 para decodificar de acuerdo con una de las formas de realización antes descritas.

El aparato 201 para decodificar se configura para recibir la señal multicanal codificada 107, siendo generada por el aparato 100 para codificar, desde el aparato 100 para codificar.

Por otra parte, se proporciona una señal multicanal codificada 107.

La señal multicanal codificada comprende

- canales codificados (E1:E3), y
- parámetros multicanal MCH\_PAR1, MCH\_PAR2, e
- información que indica si un aparato para decodificar ha de rellenar o no líneas espectrales de una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero, con datos espectrales generados en función de canales de salida de audio previamente decodificados que fueron previamente decodificados por el aparato para decodificar.

De acuerdo con una forma de realización, la señal multicanal codificada puede comprender, por ejemplo, como los parámetros multicanal MCH\_PAR1, MCH\_PAR2 dos o más parámetros multicanal.

Cada uno de los dos o más parámetros multicanal MCH\_PAR1, MCH\_PAR2 pueden indicar, por ejemplo, exactamente dos canales, donde cada uno de los exactamente dos canales es uno de los canales codificados (E1:E3) o es uno de una pluralidad de canales procesados P1, P2, P3, P4 o uno de los al menos tres canales originales (por ejemplo, no procesados) (CH:CH3).

La información que indica si un aparato para decodificar ha de rellenar o no líneas espectrales de una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero, puede comprender, por ejemplo, información que indica para cada uno de los dos o más parámetros multicanal MCH\_PAR1, MCH\_PAR2, si para al menos un canal de los exactamente dos canales que se indican por dicho uno de los dos o más parámetros multicanal, el aparato para decodificar ha de llenar líneas espectrales de una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero, de dicho al menos un canal, con los datos espectrales generados a base de los canales de salida de audio previamente decodificados que fueron previamente decodificados por el aparato para decodificar.

Como ya se destacó, también más abajo, se describen formas de realización particulares, donde tal información se transmite usando un valor hasStereoFilling[pair] que indica si el llenado estéreo se ha de aplicar o no en el par de canales MCT actualmente procesados.

A continuación, los conceptos generales y las formas de realización particulares se describen con mayor detalle.

Las formas de realización realizan un modo de codificación de baja tasa de bits paramétrica con la flexibilidad del uso de árboles estéreo arbitrarios de la combinación de llenado estéreo y MCT.

Las dependencias de señal intercanal se explotan por aplicación jerárquica de herramientas conocidas de codificación estereofónica conjunta. Para menores tasas de bits, las formas de realización extienden MCT para usar una combinación de cajas de codificación estéreo discretas y llenado de cajas estéreo. De esta manera, la codificación

semiparamétrica se puede aplicar, por ejemplo, para canales con similar contenido, es decir, pares de canales con la máxima correlación, mientras que diferentes canales se pueden codificar independientemente o por medio de una representación no paramétrica. En consecuencia, la sintaxis de flujo de bits MCT se extiende para ser capaz de señalar si se permite el llenado estéreo y dónde se activa.

5 Las formas de realización realizan una generación de una mezcla previa para pares arbitrarios de llenado estéreo. El llenado estéreo se basa en el uso de la mezcla de la trama previa para mejorar el llenado de huecos espectrales causados por cuantificación en el dominio de frecuencia. Sin embargo, en combinación con MCT, el grupo de pares estéreo conjuntamente codificados permite ahora ser variable en el tiempo. En consecuencia, dos canales codificados  
10 conjuntamente pueden no codificarse en conjunto en la trama previa, es decir, cuando cambió la configuración del árbol.

Para estimar una mezcla previa, los canales de salida previamente decodificados se guardan y se procesan con una operación estéreo inversa. Para una caja estéreo dada, esto se hace usando los parámetros de la trama actual y los canales de salida decodificados de la trama previa correspondientes a los índices del canal de la caja estéreo procesada.

Si una señal del canal de salida previo no está disponible, por ejemplo, debido a una trama independiente (una trama que se puede decodificar sin tener en cuenta los datos de trama previos) o un cambio de la longitud de transformada, la memoria intermedia del canal previo del correspondiente canal se fija en cero. De esta manera, una mezcla previa no cero incluso se puede computar, siempre que al menos una de las señales del canal previas esté disponible.

Si MCT se configura para usar cajas estéreo a base de predicciones, la mezcla previa se calcula con una operación de MS inversa según se especifica para los pares de llenado estéreo, con preferencia, usando una de las dos siguientes ecuaciones a base de una bandera de dirección de predicción (*pred\_dir* en la sintaxis MPEG-H).

$$D_1 = (\widehat{O}_1 + \widehat{O}_2) \cdot d$$

$$D_2 = (\widehat{O}_1 - \widehat{O}_2) \cdot d,$$

30 donde  $d$  es un escalar real y positivo arbitrario.

Si MCT se configura para usar cajas estéreo a base de rotación, la mezcla previa se calcula usando una rotación con el ángulo de rotación negado.

De esta manera, para una rotación dada como:

$$\begin{bmatrix} O_1 \\ O_2 \end{bmatrix} = \begin{bmatrix} \cos \alpha & -\text{sen } \alpha \\ \text{sen } \alpha & \cos \alpha \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}$$

40 la rotación inversa se calcula como:

$$\begin{bmatrix} \widehat{I}_1 \\ \widehat{I}_2 \end{bmatrix} = \begin{bmatrix} \cos \alpha & \text{sen } \alpha \\ -\text{sen } \alpha & \cos \alpha \end{bmatrix} \cdot \begin{bmatrix} \widehat{O}_1 \\ \widehat{O}_2 \end{bmatrix}$$

45 siendo  $\widehat{I}_1$  la mezcla previa deseada de los canales de salida  $\widehat{O}_1$  y  $\widehat{O}_2$  previos. Las formas de realización realizan una aplicación de llenado estéreo en MCT.

La aplicación de llenado estéreo para una caja estéreo simple se describe en [1], [5]. Según una caja estéreo simple, el llenado estéreo se aplica al segundo canal de un par de canales MCT dado.

50 Inter alia, las diferencias de llenado estéreo en combinación con MCT son de la siguiente manera:  
La configuración de árbol MCT se extiende por un bit de señalización por trama para ser capaz de señalar si se permite el llenado estéreo en la trama actual.

55 En la forma de realización preferida, si se permite el llenado estéreo en la trama actual, se transmite un bit adicional para activar el llenado estéreo en una caja estéreo para cada caja estéreo. Esta es la forma de realización preferida dado que permite el control del lado del codificador sobre qué cajas tendrán aplicado un llenado estéreo en el decodificador.

60 En una segunda forma de realización, si se permite el llenado estéreo en la trama actual, el llenado estéreo se permite en todas las cajas estéreo y no se transmite un bit adicional para cada caja estéreo individual. En este caso, la aplicación selectiva de llenado estéreo en las cajas individuales de MCT se controla por el decodificador.

Por otra parte, los conceptos y formas de realización detalladas se describen a continuación:  
Las formas de realización mejoran la calidad para puntos operativos multicanal de baja tasa de bits.

- 5 En un elemento de par de canales codificados de dominio de frecuencia (FD) (CPE), el estándar de audio MPEG–H 3D permite el uso de una herramienta de llenado estéreo, descrito en la subcláusula 5.5.5.4.9 de [1], para mejorar perceptualmente el llenado de huecos espectrales causados por una cuantificación muy burda en el codificador. Esta herramienta mostró ser beneficiosa en especial para estéreo de dos canales codificado a tasas de bits medias y bajas.
- 10 La herramienta de codificación multicanal (MCT), descrita en la sección 7 de [2], se introdujo, que permite definiciones flexibles de adaptación de señales de pares de canales conjuntamente codificadas sobre una base por trama para explotar las dependencias variables en el tiempo en una configuración multicanal. El mérito de MCT es particularmente significativo cuando se usa para la codificación conjunta dinámica eficaz de configuraciones multicanal donde cada canal reside en el elemento de canal simple individual (SCE) dado que, a diferencia de las configuraciones
- 15 tradicionales de CPE + SCE (+ LFE) que se pueden establecer a priori, permite la codificación del canal conjunto para poner en cascada y/o reconfigurar de una trama a la siguiente.

La codificación del sonido envolvente multicanal sin usar CPEs tiene actualmente la desventaja de que las herramientas estéreo conjuntas solo disponibles en CPEs –codificación predictiva de M/S y llenado estéreo– no se pueden explotar, lo que es en especial desventajoso a tasas de bits medias y bajas. MCT puede actuar como un sustituto de la herramienta M/S, pero un sustituto de la herramienta de llenado estéreo actualmente no está disponible.

20

Las formas de realización permiten el uso de la herramienta de llenado estéreo también dentro de los pares de canales de MCT por extensión de la sintaxis de corriente de bits MCT con un respectivo bit de señalización y por generalización de la aplicación de llenado estéreo a pares de canales arbitrarios sin tener en cuenta sus tipos de elementos de canal.

25

Algunas formas de realización se pueden realizar, por ejemplo, para señalar el llenado estéreo MCT de la siguiente manera:

- 30 En un CPE, el uso de la herramienta de llenado estéreo se señala dentro de la información de llenado de ruido FD para el segundo canal, como se describió en la subcláusula 5.5.5.4.9.4 de [1]. Cuando se utiliza MCT, cada canal es potencialmente un “segundo canal” (debido a la posibilidad de pares de canales de elementos cruzados). Por ello, se propone señalar explícitamente el llenado estéreo por medio de un bit adicional por par de canales codificados por MCT. Para evitar la necesidad de este bit adicional cuando no se emplea el llenado estéreo en cualquier par de canales de una instancia de “árbol” específica de MCT, las dos entradas actualmente reservadas del elemento
- 35 *MCTSignalingType* en *MultichannelCodingFrame()* [2] se utilizan para señalar la presencia del bit adicional antes mencionado por par de canales.

Abajo se proporciona una descripción detallada.

- 40 Algunas formas de realización pueden realizar, por ejemplo, el cálculo de la mezcla previa de la siguiente manera: El llenado estéreo en un CPE llena ciertas bandas de factor de escala “vacías” del segundo canal por adición de los respectivos coeficientes MDCT de la mezcla de la trama previa, escalada de acuerdo con los correspondientes factores de escala transmitidos de bandas (que no se usan dado que dichas bandas se cuantifican completamente a cero). El proceso de adición ponderada, controlado usando las bandas de factor de escala del canal de destino, se puede
- 45 emplear de modo idéntico en el contexto de MCT. El espectro fuente para llenado estéreo, es decir, la mezcla de la trama previa, sin embargo, se debe computar de una manera diferente que dentro de CPEs, en particular dado que la configuración de “árbol” de MCT puede ser variable en el tiempo.

En MCT, la mezcla previa se puede derivar de los últimos canales de salida decodificados de la trama (que se almacenan después de la decodificación de MCT) usando los parámetros de MCT de la trama actual para el par de canal conjunto dado. Para un par que aplica codificación conjunta a base de M/S predictiva, la mezcla previa iguala, como en llenado estéreo de CPE, ya sea la suma o la diferencia de los espectros de canales apropiados según el indicador de dirección de la trama actual. Para un par estéreo usando codificación conjunta a base de rotación de Karhunen–Loève, la mezcla previa representa una rotación inversa computada con los ángulos de rotación de la trama

50 actual. Nuevamente, se proporciona una descripción detallada más abajo.

55

Una evaluación de la complejidad muestra que no se espera que el llenado estéreo en MCT, que es una herramienta de tasa de bits media y baja, incremente la complejidad en el peor caso cuando se mide respecto de tasas de bits tanto bajas/medias como altas. Más aún, usando llenado estéreo, normalmente coincide con más coeficientes espectrales que son cuantificados a cero, reduciendo así la complejidad algorítmica del decodificador aritmético en función del

60 contexto. Asumiendo el uso de a lo sumo canales de llenado estéreo N/3 en una configuración envolvente de canal N y 0,2 WMOPS adicional por ejecución de llenado estéreo, la complejidad pico se incrementa en solo 0,4 WMOPS para 5.1 y 0.8 WMOPS para 11,1 canales cuando la tasa de muestreo de codificador es de 48 kHz y la herramienta de IGF opera solo a más de 12 kHz. Esto da menos del 2 % de la complejidad total del decodificador.

Las formas de realización implementan un elemento MultichannelCodingFrame() de la siguiente manera:

Sintaxis	N.º de bits	Mnemotécn.
<pre> MultichannelCodingFrame() {     <b>MCTSignalingType;</b>     <b>keepTree;</b>     si(keepTree==0) {         <b>numPairs=escapedValue(5,8,16);</b>     }     de lo contrario {         numPairs=lastNumPairs;     }     MCTStereoFilling = 0;     si (MCTSignalingType &gt; 1) {         MCTSignalingType = MCTSignalingType - 2;         MCTStereoFilling = 1;     }     para(pair=0; pair&lt;numPairs;pair++) {         hasStereoFilling[pair] = 0;         si(MCTStereoFilling == 1) {             <b>hasStereoFilling[pair];</b>         }         si(MCTSignalingType == 0) { /* tree of stereo prediction boxes */             MultichannelCodingBoxPrediction();         }         si(MCTSignalingType == 1) { /* tree of rotation boxes */             MultichannelCodingBoxRotation();         }         si((MCTSignalingType == 2)    (MCTSignalingType == 3)) {             /* reserved */         }     } } </pre>	<p>2 1</p> <p>1</p>	<p>uimsbf uimsbf</p> <p>uimsbf</p>

5 El llenado estéreo en MCT se puede implementar, de acuerdo con algunas formas de realización, de la siguiente manera:

Como el llenado estéreo para IGF en un elemento de par de canales, descrito en la subcláusula 5.5.5.4.9 de [1], el llenado estéreo en la herramienta de codificación multicanal (MCT) llena bandas de factor de escala “vacías” (que están completamente cuantificadas a cero) en y por encima de la frecuencia inicial de llenado de ruido usando una mezcla de los espectros de salida de la trama previa.

10 Cuando el llenado estéreo es activo en un par de canales conjuntos de MCT ( $hasStereoFilling[pair] \neq 0$  en la Tabla AMD4.4), todas las bandas de factor de escala “vacías” en la región de llenado de ruido (es decir, a partir de o por encima de  $noiseFillingStartOffset$ ) del segundo canal del par se llenan hasta una energía de destino específica usando una mezcla de los correspondientes espectros de salida (después de la aplicación de MCT) de la trama previa. Esto se hace después del llenado de ruido FD (véase la subcláusula 7.2 en ISO/IEC 23003–3:2012) y antes del factor de escala y aplicación estéreo conjunta de MCT. Todos los espectros de salida después del procesamiento de MCT completado se salvan para el potencial llenado estéreo en la siguiente trama.

15 Las restricciones operativas pueden ser, por ejemplo, aquellas del algoritmo de llenado estéreo de ejecución en cascada ( $hasStereoFilling[pair] \neq 0$ ) en bandas vacías del segundo canal no soportadas para cualquier par estéreo de MCT con  $hasStereoFilling[pair] \neq 0$  si el segundo canal es el mismo. En un elemento de par de canales, el llenado estéreo activo de IGF en el segundo canal (residual) de acuerdo con la subcláusula 5.5.5.4.9 de [1] tiene precedencia respecto –y, así, deshabilita– cualquier aplicación posterior de llenado estéreo MCT en el mismo canal de la misma trama.

25 Los términos y definiciones se pueden definir, por ejemplo, de la siguiente manera:

30  $hasStereoFilling[pair]$  indica el uso de llenado estéreo en un par de canales MCT actualmente procesados  
 $ch1, ch2$  índices de canales en canal MCT actualmente procesados

	spectral_data[ ][ ]	coeficientes espectrales de canales en par de canales MCT actualmente procesados
	spectral_data_prev[ ][ ]	espectros de salida después de procesamiento de MCT completado en trama previa
	downmix_prev[ ][ ]	mezcla estimada de canales de salida de la trama previa con índices dados por par de canales MCT actualmente procesados
5	num_swb	cantidad total de bandas de factor de escala, ver ISO/IEC 23003-3, subcláusula 6.2.9.4
	ccfl	coreCoderFrameLength, longitud de la transformada, véase ISO/IEC 23003-3, subcláusula 6.1.
	noiseFillingStartOffset	línea inicial de llenado de ruidos, definida según ccfl en ISO/IEC 23003-3, Tabla 109.
	igf_WhiteningLevel	blanqueo espectral en IGF, véase ISO/IEC 23008-3, subcláusula 5.5.5.4.7
10	seed[ ]	semilla de llenado de ruidos usado por randomSign(), véase ISO/IEC 23003-3, subcláusula 7.2.

Para algunas formas de realización particulares, el proceso de decodificación se puede describir, por ejemplo, de la siguiente manera:

15 El llenado estéreo MCT se lleva a cabo usando cuatro operaciones consecutivas, que se describen a continuación:

Etapa 1: Preparación del espectro del segundo canal para el algoritmo de llenado estéreo

20 Si el indicador de llenado estéreo para el par dado del canal MCT, hasStereoFilling[pair], es igual a cero, no se usa el llenado estéreo y las siguientes etapas no se ejecutan. De otro modo, la aplicación de factor de escala no se realiza si se aplicaba previamente al espectro del segundo canal del par, spectral\_data[ch2].

Etapa 2: Generación del espectro de mezcla previo para par de canales de MCT dados

25 La mezcla previa se estima de las señales de salida de la trama previa spectral\_data\_prev[ ][ ] que se almacenó después de la aplicación del procesamiento de MCT. Si una señal del canal de salida previa no está disponible, por ejemplo, debido a una trama independiente (indepFlag>0), cambio de longitud de transformada o core\_mode == 1, la memoria intermedia del canal previo del correspondiente canal se ha de fijar en cero.

30 Para pares estéreo de predicción, es decir, MCTSignalingType == 0, la mezcla previa se calcula a partir de los canales de salida previos como downmix\_prev[ ][ ] definido en la etapa 2 de la subcláusula 5.5.5.4.9.4 de [1], donde spectrum>window[ ] se representa por spectral\_data[ ][window].

35 Para los pares estéreo de rotación, es decir, MCTSignalingType == 1, la mezcla previa se calcula a partir de los canales de salida previos invirtiendo la operación de rotación definida en la subcláusula 5.5.X.3.7.1 de [2].

```
apply_mct_rotation_inverse(*R, *L, *dmx, aIdx, nSamples)
```

```
{
  para (n=0; n<nSamples; n++) {
40   dmx = L[n] * tabIndexToCosAlpha[aIdx] + R[n] * tabIndexToSinAlpha[aIdx];
  }
}
```

45 usando L = spectral\_data\_prev[ch1][ ], R = spectral\_data\_prev[ch2][ ], dmx = downmix\_prev[ ] de la trama previa y usando aIdx, nSamples de la trama actual y par de MCT.

Etapa 3: Ejecución de algoritmo de llenado estéreo en bandas vacías del segundo canal

50 El llenado estéreo se aplica en el segundo canal del par de MCT como en la etapa 3 de la subcláusula 5.5.5.4.9.4 de [1], donde spectrum>window está representado por spectral\_data[ch2][window] y max\_sfb\_ste viene dado por num\_swb.

Etapa 4: Aplicación del factor de escala y sincronización adaptativa de semillas de llenado de ruidos

55 Como después de la etapa 3 de la subcláusula 5.5.5.4.9.4 de [1], los factores de escala se aplican al espectro resultante como en 7.3 de ISO/IEC 23003-3, procesando los factores de escala de bandas vacías como factores de escala regulares. En caso de que un factor de escala no se defina, por ejemplo, porque se ubica por encima de max\_sfb, su valor debe ser igual a cero. Si se usa IGF, igf\_WhiteningLevel es igual a 2 en cualquiera de los mosaicos del segundo canal y ambos canales no emplean transformación corta de ocho, las energías espectrales de ambos canales en el par de MCT se computan en el rango de índice noiseFillingStartOffset a índice ccfl/2 - 1 antes de ejecutar decode\_mct( ). Si la energía computada del primer canal es más de ocho veces la energía del segundo canal, la semilla [ch2] del segundo canal se fija igual a la semilla [ch1] del primer canal.

60

A pesar de que algunos aspectos fueron descritos en el contexto de un aparato, queda claro que estos aspectos

también representan una descripción del correspondiente método, donde un bloque o dispositivo corresponde a una etapa del método o una característica de una etapa de un método. Análogamente, los aspectos descritos en el contexto de una etapa del método también representan una descripción de un correspondiente bloque o artículo o característica de un correspondiente aparato. Algunas o todas las etapas del método se pueden ejecutar por (o usando) un aparato de hardware, como, por ejemplo, un microprocesador, una computadora programable o un circuito electrónico. En algunas formas de realización, una o varias de las etapas del método más importantes se pueden ejecutar por tal aparato.

Según ciertos requerimientos de implementación, las formas de realización de la invención se pueden implementar en hardware o en software o al menos parcialmente en hardware o al menos parcialmente en software. La implementación se puede llevar a cabo usando un medio de almacenamiento digital, por ejemplo, un disquete, un DVD, un Blu-Ray, un CD, ROM, PROM, EPROM, EEPROM o una memoria FLASH, que tiene señales de control electrónicamente legibles almacenadas allí que cooperan (o son capaces de cooperar) con un sistema de computadora programable de un modo tal para realizar el respectivo método. En consecuencia, el medio de almacenamiento digital puede ser legible por computadora.

Algunas formas de realización de acuerdo con la invención comprenden un portador de datos que tiene señales de control electrónicamente legibles, que son capaces de cooperar con un sistema de computadora programable, de modo tal que se realice uno de los métodos descritos en la presente.

En general, las formas de realización de la presente invención se pueden implementar como un producto de programa de computadora con un código de programa, siendo el código de programa operativo para realizar uno de los métodos cuando el producto de programa de computadora corre en una computadora. El código de programa se puede almacenar, por ejemplo, en un portador legible por máquina.

Otras formas de realización comprenden el programa de computadora para realizar uno de los métodos descritos en la presente, almacenado en un portador legible por máquina.

En otras palabras, una forma de realización del método de la invención es, en consecuencia, un programa de computadora que tiene un código de programa para realizar uno de los métodos descritos con anterioridad, cuando el programa de computadora corre en una computadora.

Otra forma de realización de los métodos de la invención es, por ende, un portador de datos (o un medio de almacenamiento digital o un medio legible por computadora) que comprende, registrado allí, el programa de computadora para realizar uno de los métodos descritos allí. El portador de datos, el medio de almacenamiento digital o el medio registrado son típicamente tangibles y/o no transitorios.

Otra forma de realización del método de la invención es, por ende, un flujo de datos o una secuencia de señales que representan el programa de computadora para llevar a cabo uno de los métodos descritos allí. El flujo de datos o la secuencia de señales se puede configurar, por ejemplo, para ser transferido por medio de una conexión de comunicación de datos, por ejemplo, a través de Internet.

Otra forma de realización comprende un medio de procesamiento, por ejemplo, una computadora o un dispositivo lógico programable, configurado o adaptado para realizar uno de los métodos descritos allí.

Otra forma de realización comprende una computadora que tiene instalado dentro el programa de computadora para llevar a cabo uno de los métodos descritos allí.

Otra forma de realización de acuerdo con la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente u ópticamente) un programa de computadora para realizar uno de los métodos descritos allí a un receptor. El receptor puede ser, por ejemplo, una computadora, un dispositivo móvil, un dispositivo de memoria o similares. El aparato o sistema puede comprender, por ejemplo, un servidor de archivos para transferir el programa de computadora al receptor.

En algunas formas de realización, se puede usar un dispositivo lógico programable (por ejemplo, una disposición de puerto programable en campo) para realizar algunas o todas las funcionalidades de los métodos descritos en la presente. En algunas formas de realización, una disposición de puerto programable en campo puede cooperar con un microprocesador a fin de realizar uno de los métodos descritos allí. En general, los métodos se realizan con preferencia por cualquier aparato de hardware.

El aparato descrito en la presente se puede implementar usando un aparato de hardware o usando una computadora o usando una combinación de un aparato de hardware y una computadora.

Los métodos descritos en la presente se pueden llevar a cabo usando un aparato de hardware o usando una

computadora o usando una combinación de un aparato de hardware y una computadora.

Las formas de realización antes descritas son meramente ilustrativas para los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en la presente serán obvias para los expertos en la técnica. Se pretende, por ello, que estén limitadas solo por el alcance de las reivindicaciones de patentes inminentes y no por los detalles específicos presentados a modo de descripción y explicación de las formas de realización en la presente.

Referencias

[1] ISO/IEC international standard 23008-3:2015, "Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 3: 3D audio," marzo de 2015

[2] ISO/IEC amendment 23008-3:2015/PDAM3, "Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 3: 3D audio, Amendment 3: MPEG-H 3D Audio Phase 2," julio de 2015

[3] International Organization for Standardization, ISO/IEC 23003-3:2012, "Information Technology – MPEG audio – Part 3: Unified speech and audio coding," Ginebra, enero de 2012

[4] ISO/IEC 23003-1:2007 – Information technology – MPEG audio technologies Part 1: MPEG Surround

[5] C. R. Helmrich, A. Niedermeier, S. Bayer, B. Edler, "Low-Complexity Semi-Parametric Joint-Stereo Audio Transform Coding," in Proc. EUSIPCO, Niza, Septiembre de 2015

[6] ETSI TS 103 190 V1.1.1 (2014-04) – Digital Audio Compression (AC-4) Standard

[7] Yang, Dai and Ai, Hongmei and Kyriakakis, Chris and Kuo, C.-C. Jay, 2001: Adaptive Karhunen-Loeve Transform for Enhanced Multichannel Audio Coding, <http://ict.usc.edu/pubs/Adaptive%20Karhunen-Loeve%20Transform%20for%20Enhanced%20Multichannel%20Audio%20Coding.pdf>

[8] European Patent Application, Publication EP 2 830 060 A1: "Noise filling in multichannel audio coding", publicada el 28 de enero de 2015

[9] Internet Engineering Task Force (IETF), RFC 6716, "Definition of the Opus Audio Codec," Int. Standard, septiembre de 2012. Disponible online en: <http://tools.ietf.org/html/rfc6716>

[10] International Organization for Standardization, ISO/IEC 14496-3:2009, "Information Technology – Coding of audio-visual objects – Part 3: Audio," Ginebra, Suiza, agosto de 2009

[11] M. Neuendorf et al., "MPEG Unified Speech and Audio Coding – The ISO/MPEG Standard for High-Efficiency Audio Coding of All Content Types," in Proc. 132<sup>nd</sup> AES Convention, Budapest, Hungría, abril de 2012. A aparecer también en la revista del AES, 2013

REIVINDICACIONES

1. Aparato (201) para decodificar una señal multicanal codificada previa de una trama previa para obtener tres o más canales de salida de audio previos y para decodificar una señal multicanal codificada actual (107) de una trama actual para obtener tres o más canales de salida de audio actuales, en donde el aparato (201) comprende una interfaz (212), un decodificador de canales (202), un procesador multicanal (204) para generar los tres o más canales de salida de audio actuales y un módulo de llenado de ruidos (220), en donde la interfaz (212) se adapta para recibir la señal multicanal codificada actual (107) y para recibir información secundaria que comprende primeros parámetros multicanal (MCH\_PAR2), en donde el decodificador de canales (202) se adapta para decodificar la señal multicanal codificada actual de la trama actual para obtener un grupo de tres o más canales decodificados (D1, D2, D3) de la trama actual, en donde el procesador multicanal (204) se adapta para seleccionar un primer par seleccionado de dos canales decodificados (D1, D2) del grupo de tres o más canales decodificados (D1, D2, D3) según los primeros parámetros multicanal (MCH\_PAR2), en donde el procesador multicanal (204) se adapta para generar un primer grupo de dos o más canales procesados (P1\*,P2\*) en función de dicho primer par seleccionado de dos canales decodificados (D1, D2) para obtener un grupo actualizado de tres o más canales decodificados (D3, P1\*, P2\*), **caracterizado porque**, antes de que el procesador multicanal (204) genere el primer grupo de dos o más canales procesados (P1\*,P2\*) en función de dicho primer par seleccionado de dos canales decodificados (D1, D2), el módulo de llenado de ruidos (220) se adapta para identificar al menos uno de los dos canales de dicho primer par seleccionado de dos canales decodificados (D1, D2), una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero y para generar un canal de mezcla usando dos o más, pero no todos los tres o más canales de salida de audio previos y para llenar las líneas espectrales de la una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero, con ruido generado usando líneas espectrales del canal de mezcla, en donde el módulo de llenado de ruidos (220) se adapta para seleccionar los dos o más canales de salida de audio previos que se usan para generar el canal de mezcla de los tres o más canales de salida de audio previos según la información secundaria.

2. Un aparato (201) de acuerdo con la reivindicación 1, en donde el módulo de llenado de ruidos (220) se adapta para generar el canal de mezcla usando exactamente dos canales de salida de audio previos de los tres o más canales de salida de audio como los dos o más de los tres o más canales de salida de audio previos; en donde el módulo de llenado de ruidos (220) se adapta para seleccionar los exactamente dos canales de salida de audio previos de los tres o más canales de salida de audio previos según la información secundaria.

3. Un aparato (201) de acuerdo con la reivindicación 2, en donde el módulo de llenado de ruidos (220) se adapta para generar el canal de mezcla usando exactamente dos canales de salida de audio previos en función de la fórmula

$$D_{ch} = (\hat{O}_1 + \hat{O}_2) \cdot d$$

o en función de la fórmula

$$D_{ch} = (\hat{O}_1 - \hat{O}_2) \cdot d$$

en donde  $D_{ch}$  es el canal de mezcla,

en donde  $\hat{O}_1$  es un primero de los exactamente dos canales de salida de audio previos,

en donde  $\hat{O}_2$  es un segundo de los exactamente dos canales de salida de audio previos, siendo diferente del primero de los exactamente dos canales de salida de audio previos, y en donde  $d$  es un escalar positivo real.

4. Un aparato (201) de acuerdo con la reivindicación 2, en donde el módulo de llenado de ruidos (220) se adapta para generar el canal de mezcla usando exactamente dos canales de salida de audio previos en función de la fórmula

$$\hat{I}_{ch} = (\cos \alpha \cdot \hat{O}_1 + \sin \alpha \cdot \hat{O}_2) \cdot d$$

o en función de la fórmula

$$\hat{I}_{ch} = (-\sin \alpha \cdot \hat{O}_1 + \cos \alpha \cdot \hat{O}_2) \cdot d$$

en donde  $\hat{I}_{ch}$  es el canal de mezcla,

en donde  $\hat{O}_1$  es un primero de los exactamente dos canales de salida de audio previos,

en donde  $\hat{O}_2$  es un segundo de los exactamente dos canales de salida de audio previos, siendo diferente del primero de los exactamente dos canales de salida de audio previos, y

5 en donde  $\alpha$  es un ángulo de rotación.

5. Un aparato (201) de acuerdo con la reivindicación 4, en donde la información secundaria es información secundaria actual que se asigna a la trama actual,

10 en donde la interfaz (212) se adapta para recibir información secundaria previa que se asigna a la trama previa, en donde la información secundaria previa comprende un ángulo previo,

en donde la interfaz (212) se adapta para recibir la información secundaria actual que comprende un ángulo actual y en donde el módulo de llenado de ruidos (220) se adapta para usar el ángulo actual de la información secundaria actual como el ángulo de rotación  $\alpha$  y se adapta para no usar el ángulo previo de la información secundaria previa como el ángulo de rotación  $\alpha$ .

15 6. Un aparato (201) de acuerdo con una de las reivindicaciones 2 a 5, en donde el módulo de llenado de ruidos (220) se adapta para seleccionar los exactamente dos canales de salida de audio previos de los tres o más canales de salida de audio previos según los primeros parámetros multicanal (MCH\_PAR2).

20 7. Un aparato (201) de acuerdo con una de las reivindicaciones 2 a 6, en donde la interfaz (212) se adapta para recibir la señal multicanal codificada actual (107) y para recibir la información secundaria que comprende los primeros parámetros multicanal (MCH\_PAR2) y los segundos parámetros multicanal (MCH\_PAR1),

25 en donde el procesador multicanal (204) se adapta para seleccionar un segundo par seleccionado de dos canales decodificados (P1\*, D3) del grupo actualizado de tres o más canales decodificados (D3, P1\*, P2\*) según los segundos parámetros multicanal (MCH\_PAR1), siendo al menos un canal (P1\*) del segundo par seleccionado de dos canales decodificados (P1\*, D3) un canal del primer grupo de dos o más canales procesados (P1\*, P2\*) y

en donde el procesador multicanal (204) se adapta para generar un segundo grupo de dos o más canales procesados (P3\*, P4\*) en función de dicho segundo par seleccionado de dos canales decodificados (P1\*, D3) para actualizar luego el grupo actualizado de tres o más canales decodificados.

30 8. Un aparato (201) de acuerdo con la reivindicación 7, en donde el procesador multicanal (204) se adapta para generar el primer grupo de dos o más canales procesados (P1\*, P2\*) generando un primer grupo de exactamente dos canales procesados (P1\*, P2\*) en función de dicho primer par seleccionado de dos canales decodificados (D1, D2);

35 en donde el procesador multicanal (204) se adapta para reemplazar dicho primer par seleccionado de dos canales decodificados (D1, D2) en el grupo de tres de más canales decodificados (D1, D2, D3) por el primer grupo de exactamente dos canales procesados (P1\*, P2\*) para obtener el grupo actualizado de tres o más canales decodificados (D3, P1\*, P2\*);

40 en donde el procesador multicanal (204) se adapta para generar el segundo grupo de dos o más canales procesados (P3\*, P4\*) generando un segundo grupo de exactamente dos canales procesados (P3\*, P4\*) en función de dicho segundo par seleccionado de dos canales decodificados (P1\*, D3), y

en donde el procesador multicanal (204) se adapta para reemplazar dicho segundo par seleccionado de dos canales decodificados (P1\*, D3) en el grupo actualizado de tres de más canales decodificados (D3, P1\*, P2\*) por el segundo grupo de exactamente dos canales procesados (P3\*, P4\*) para actualizar luego el grupo actualizado de tres o más canales decodificados.

45 9. Un aparato (201) de acuerdo con la reivindicación 8, en donde los primeros parámetros multicanal (MCH\_PAR2) indican dos canales decodificados (D1, D2) del grupo de tres o más canales decodificados;

50 en donde el procesador multicanal (204) se adapta para seleccionar el primer par seleccionado de dos canales decodificados (D1, D2) del grupo de tres o más canales decodificados (D1, D2, D3) seleccionando los dos canales decodificados (D1, D2) indicados por los primeros parámetros multicanal (MCH\_PAR2);

en donde los segundos parámetros multicanal (MCH\_PAR1) indican dos canales decodificados (P1\*, D3) del grupo actualizado de tres o más canales decodificados;

55 en donde el procesador multicanal (204) se adapta para seleccionar el segundo par seleccionado de dos canales decodificados (P1\*, D3) del grupo actualizado de tres o más canales decodificados (D3, P1\*, P2\*) seleccionando los dos canales decodificados (P1\*, D3) indicados por los segundos parámetros multicanal (MCH\_PAR1).

60 10. Un aparato (201) de acuerdo con la reivindicación 9, donde el aparato (201) se adapta para asignar un identificador de un grupo de identificadores a cada canal de salida de audio previo de los tres o más canales de salida de audio previos, de modo que cada canal de salida de audio previo de los tres o más canales de salida de audio previos se asigne exactamente a un identificador del grupo de identificadores y de modo que cada identificador del grupo de identificadores se asigne a exactamente un canal de salida de audio previo de los tres o más canales de salida de

audio previos,

donde el aparato (201) se adapta para asignar un identificador de dicho grupo de identificadores a cada canal del grupo de los tres o más canales decodificados (D1, D2, D3), de modo que cada canal del grupo de los tres o más canales decodificados se asigne exactamente a un identificador del grupo de identificadores y de modo que cada

5 identificador del grupo de identificadores se asigne a exactamente un canal del grupo de los tres o más canales decodificados (D1, D2, D3),

en donde los primeros parámetros multicanal (MCH\_PAR2) indican un primer par de dos identificadores del grupo de los tres o más identificadores,

10 en donde el procesador multicanal (204) se adapta para seleccionar el primer par seleccionado de dos canales decodificados (D1, D2) del grupo de tres o más canales decodificados (D1, D2, D3) seleccionando los dos canales decodificados (D1, D2) asignados a los dos identificadores del primer par de dos identificadores;

en donde el aparato (201) se adapta para asignar un primero de los dos identificadores del primer par de dos identificadores a un primer canal procesado del primer grupo de exactamente dos canales procesados (P1\*,P2\*) y en

15 donde el aparato (201) se adapta para asignar un segundo de los dos identificadores del primer par de dos identificadores a un segundo canal procesado del primer grupo de exactamente dos canales procesados (P1\*,P2\*).

11. Un aparato (201) de acuerdo con la reivindicación 10, en donde los segundos parámetros multicanal (MCH\_PAR1) indican un segundo par de dos identificadores del grupo de los tres o más identificadores,

20 en donde el procesador multicanal (204) se adapta para seleccionar el segundo par seleccionado de dos canales decodificados (P1\*, D3) del grupo actualizado de tres o más canales decodificados (D3, P1\*, P2\*) seleccionando los dos canales decodificados (D3, P1\*) asignados a los dos identificadores del segundo par de dos identificadores;

en donde el aparato (201) se adapta para asignar un primero de los dos identificadores del segundo par de dos identificadores a un primer canal procesado del segundo grupo de exactamente dos canales procesados (P3\*, P4\*) y

25 en donde el aparato (201) se adapta para asignar un segundo de los dos identificadores del segundo par de dos identificadores a un segundo canal procesado del segundo grupo de exactamente dos canales procesados (P3\*, P4\*).

12. Un aparato (201) de acuerdo con la reivindicación 10 u 11, en donde los primeros parámetros multicanal (MCH\_PAR2) indican dicho primer par de dos identificadores del grupo de los tres o más identificadores, y

30 en donde el módulo de llenado de ruidos (220) se adapta para seleccionar los exactamente dos canales de salida de audio previos de los tres o más canales de salida de audio previos seleccionando los dos canales de salida de audio previos asignados a los dos identificadores de dicho primer par de dos identificadores.

13. Un aparato (201) de acuerdo con una de las reivindicaciones precedentes, en donde, antes de que el procesador multicanal (204) genere el primer grupo de dos o más canales procesados (P1\*,P2\*) en función de dicho primer par

35 seleccionado de dos canales decodificados (D1, D2), el módulo de llenado de ruidos (220) se adapta para identificar al menos uno de los dos canales de dicho primer par seleccionado de dos canales decodificados (D1, D2), siendo una o más bandas de factor de escala las una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero y para generar el canal de mezcla usando dichos dos o más, pero no todos los tres o más canales

40 de salida de audio previos y para llenar las líneas espectrales de las una o más bandas de factor de escala, dentro de las que todas las líneas espectrales se cuantifican a cero, con el ruido generado usando las líneas espectrales del canal de mezcla según un factor de escala de cada una de las una o más bandas de factor de escala dentro de las que todas las líneas espectrales se cuantifican a cero.

14. Un aparato (201) de acuerdo con la reivindicación 13, en donde la interfaz de recepción (212) se configura para recibir el factor de escala de cada una de dichas una o más bandas de factor de escala, y

45 en donde el factor de escala de cada una de dichas una o más bandas de factor de escala indica una energía de las líneas espectrales de dicha banda de factor de escala antes de la cuantificación, y

50 en donde el módulo de llenado de ruidos (220) se adapta para generar el ruido para cada una de las una o más bandas de factor de escala, dentro de las que todas las líneas espectrales se cuantifican a cero, de modo que una energía de las líneas espectrales después de añadir el ruido en una de las bandas de frecuencia corresponde a la energía indicada por el factor de escala para dicha banda de factor de escala.

15. Sistema que comprende

55 un aparato (100) para codificar una señal multicanal (101) que tiene al menos tres canales (CH1:CH3), y un aparato (201) para decodificar de acuerdo con una de las reivindicaciones 1 a 14, en donde el aparato (201) para decodificar se configura para recibir una señal multicanal codificada (107), generada por el aparato (100) para decodificar, desde el aparato (100) para codificar, en donde el aparato (100) para codificar la señal multicanal (101) comprende:

60 un procesador de iteración (102) que se adapta para calcular, en una primera etapa de iteración, valores de correlación intercanal entre cada par de los al menos tres canales (CH:CH3), para seleccionar, en la primera etapa de iteración, un par que tiene un valor máximo o que tiene un valor por encima del umbral y para procesar el par seleccionado usando una operación de procesamiento multicanal (110, 112) para derivar parámetros multicanal iniciales (MCH\_PAR1) para el par seleccionado y para derivar primeros canales procesados (P1,

P2),

en donde el procesador de iteración (102) se adapta para realizar el cálculo, la selección y el procesamiento en una segunda etapa de iteración usando al menos uno de los canales procesados (P1) para derivar otros parámetros multicanal (MCH\_PAR2) y segundos canales procesados (P3, P4);

5 un codificador de canales que se adapta para codificar canales (P2:P4) que resultan de un procesamiento de iteración realizado por el procesador de iteración (104) para obtener canales codificados (E1:E3); y  
 una interfaz de salida (106) que se adapta para generar la señal multicanal codificada (107) que tiene los canales codificados (E1:E3), los parámetros multicanal iniciales y los otros parámetros multicanal (MCH\_PAR1, MCH\_PAR2) y que tiene una información que indica si un aparato para decodificar ha de rellenar  
 10 o no las líneas espectrales de una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero, con ruido generado en función de canales de salida de audio previamente decodificados que fueron previamente decodificados por el aparato para decodificar.

16. Un sistema de acuerdo con la reivindicación 15, en donde cada uno de los parámetros multicanal iniciales y los otros parámetros multicanal (MCH\_PAR1, MCH\_PAR2) indican exactamente dos canales, donde cada uno de los exactamente dos canales es uno de los canales codificados (E1:E3) o es uno de los primeros o los segundos canales procesados (P1, P2, P3, P4) o es uno de los al menos tres canales (CH:CH3), y  
 15 en donde la interfaz de salida (106) del aparato (100) para codificar la señal multicanal (101) se adapta para generar la señal multicanal codificada (107), de modo que la información que indica si un aparato para decodificar ha de rellenar o no líneas espectrales de una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero, comprenda información que indica para cada uno de los parámetros iniciales y multicanal (MCH\_PAR1, MCH\_PAR2), ya sea o no para al menos un canal de los exactamente dos canales que se indican por dichos uno de los parámetros iniciales y los otros parámetros multicanal (MCH\_PAR1, MCH\_PAR2), si el aparato para decodificar ha de llenar líneas espectrales de una o varias bandas de frecuencia, dentro de las que todas las líneas  
 20 espectrales se cuantifican a cero, de dicho al menos un canal, con los datos espectrales generados en función de los canales de salida de audio previamente decodificados que fueron previamente decodificados por el aparato para decodificar.

17. Método para decodificar una señal multicanal codificada previa de una trama previa para obtener tres o más canales de salida de audio previos y para decodificar una señal multicanal codificada 10 actual (107) de una trama actual para obtener tres o más canales de salida de audio actuales, en donde el método comprende:

30 recibir la señal multicanal codificada actual (107) y recibir información secundaria que comprende primeros parámetros multicanal (MCH\_PAR2);  
 35 decodificar la señal multicanal codificada actual de la trama actual para obtener un grupo de tres o más canales decodificados (D1, D2, D3) de la trama actual;  
 seleccionar un primer par seleccionado de dos canales decodificados (D1, D2) del grupo de tres o más canales decodificados (D1, D2, D3) según los primeros parámetros multicanal (MCH\_PAR2);  
 40 generar un primer grupo de dos o más canales procesados (P1\*, P2\*) en función de dicho primer par seleccionado de dos canales decodificados (D1, D2) para obtener un grupo actualizado de tres o más canales decodificados (D3, P1\*, P2\*);  
**caracterizado porque**, antes de generar el primer grupo de dos o más canales procesados (P1\*, P2\*) en función de dicho primer par seleccionado de dos canales decodificados (D1, D2), se llevan a cabo las siguientes etapas:  
 45 identificar para al menos uno de los dos canales de dicho primer par seleccionado de dos canales decodificados (D1, D2), una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero y generar un canal de mezcla usando dos o más, pero no todos los tres o más canales de salida de audio previos y llenar las líneas espectrales de la una o varias bandas de frecuencia, dentro de las que todas las líneas espectrales se cuantifican a cero, con ruido generado usando líneas espectrales del canal de mezcla, en donde la selección de los dos o más canales de salida de audio previos que se usan para generar el canal de mezcla de los  
 50 tres o más canales de salida de audio previos se lleva a cabo según la información secundaria.

18. Un programa de computadora para implementar el método de la reivindicación 17 cuando se ejecuta en una computadora o procesador de señales.

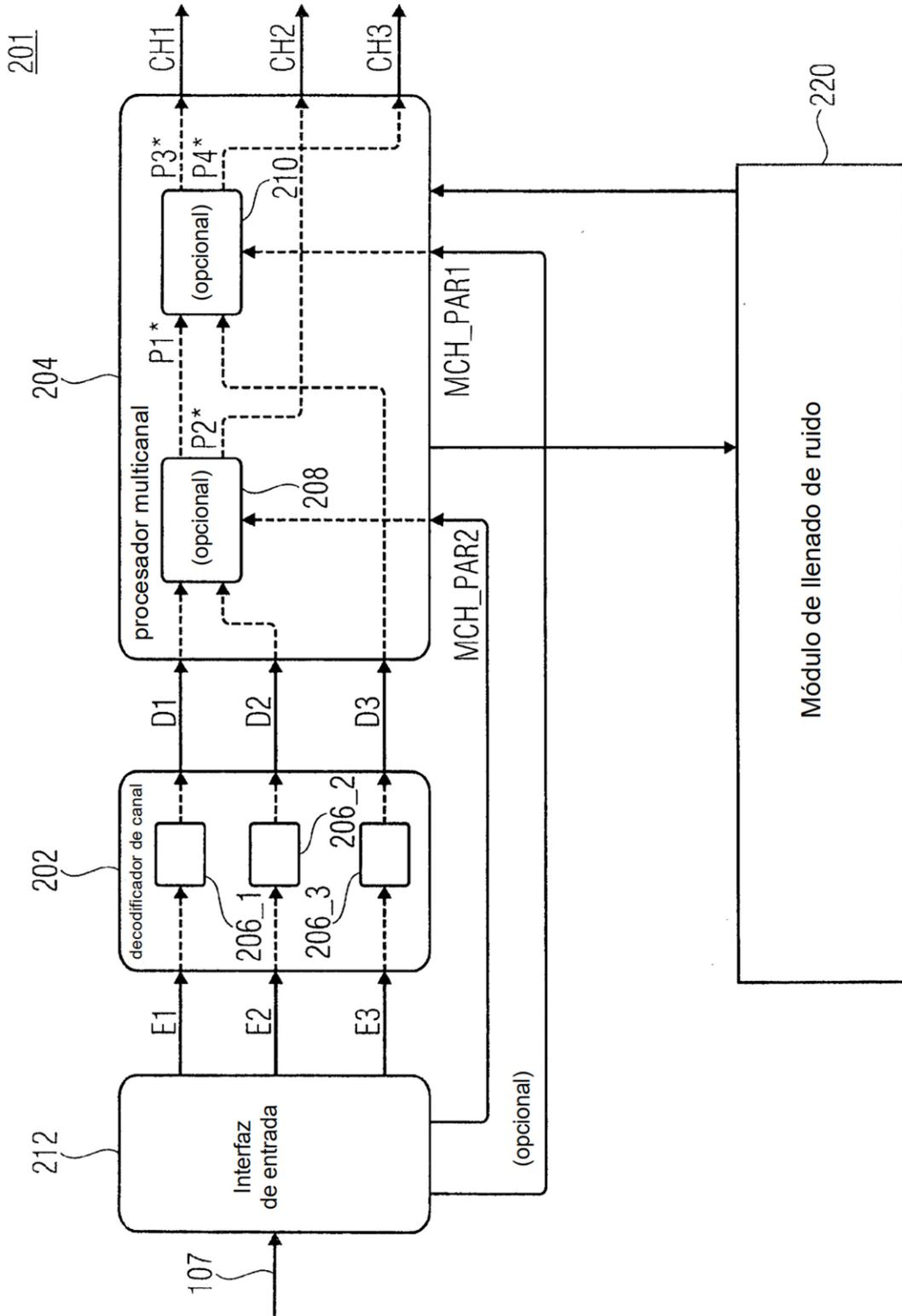


Fig. 1a

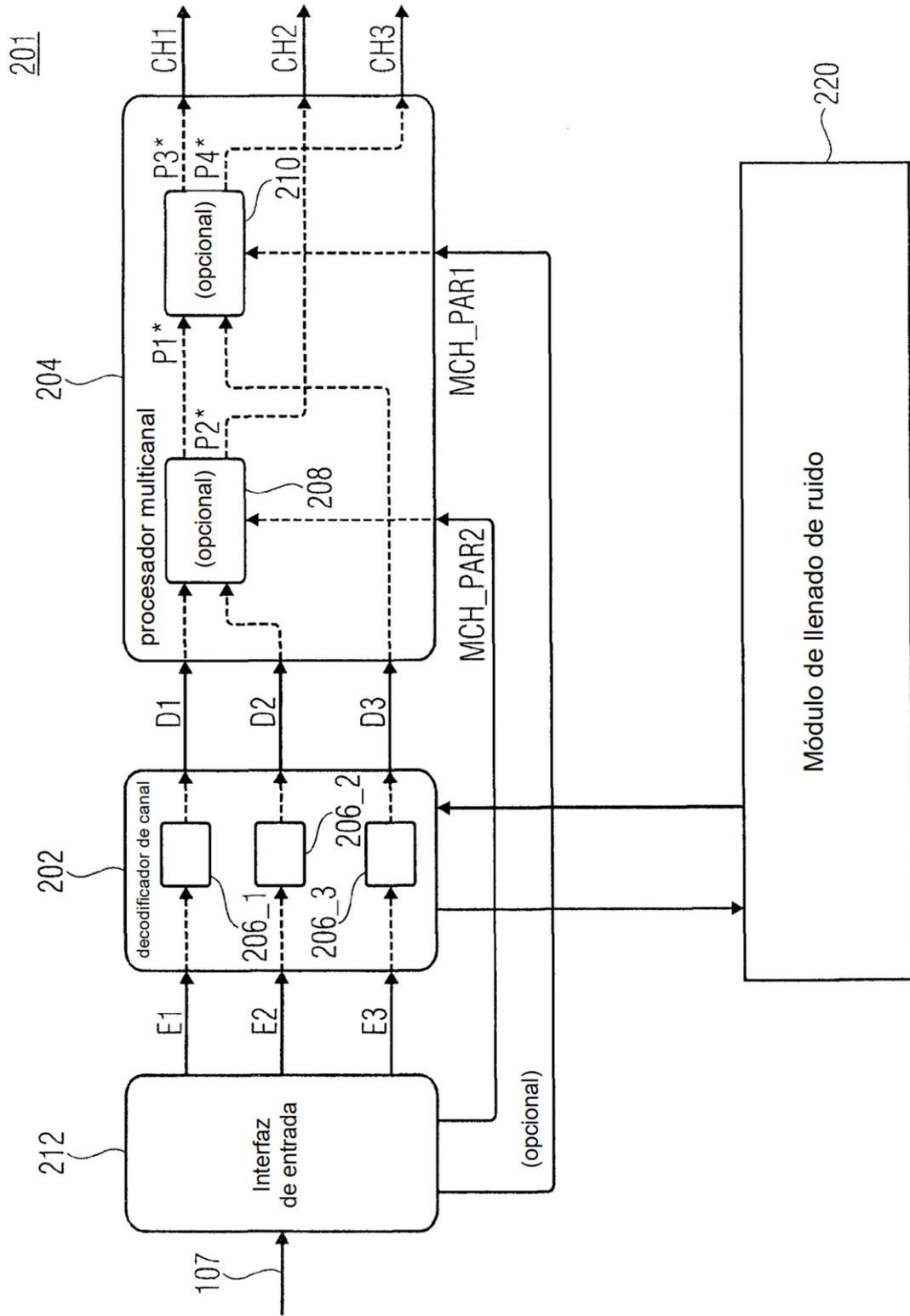


Fig. 1b

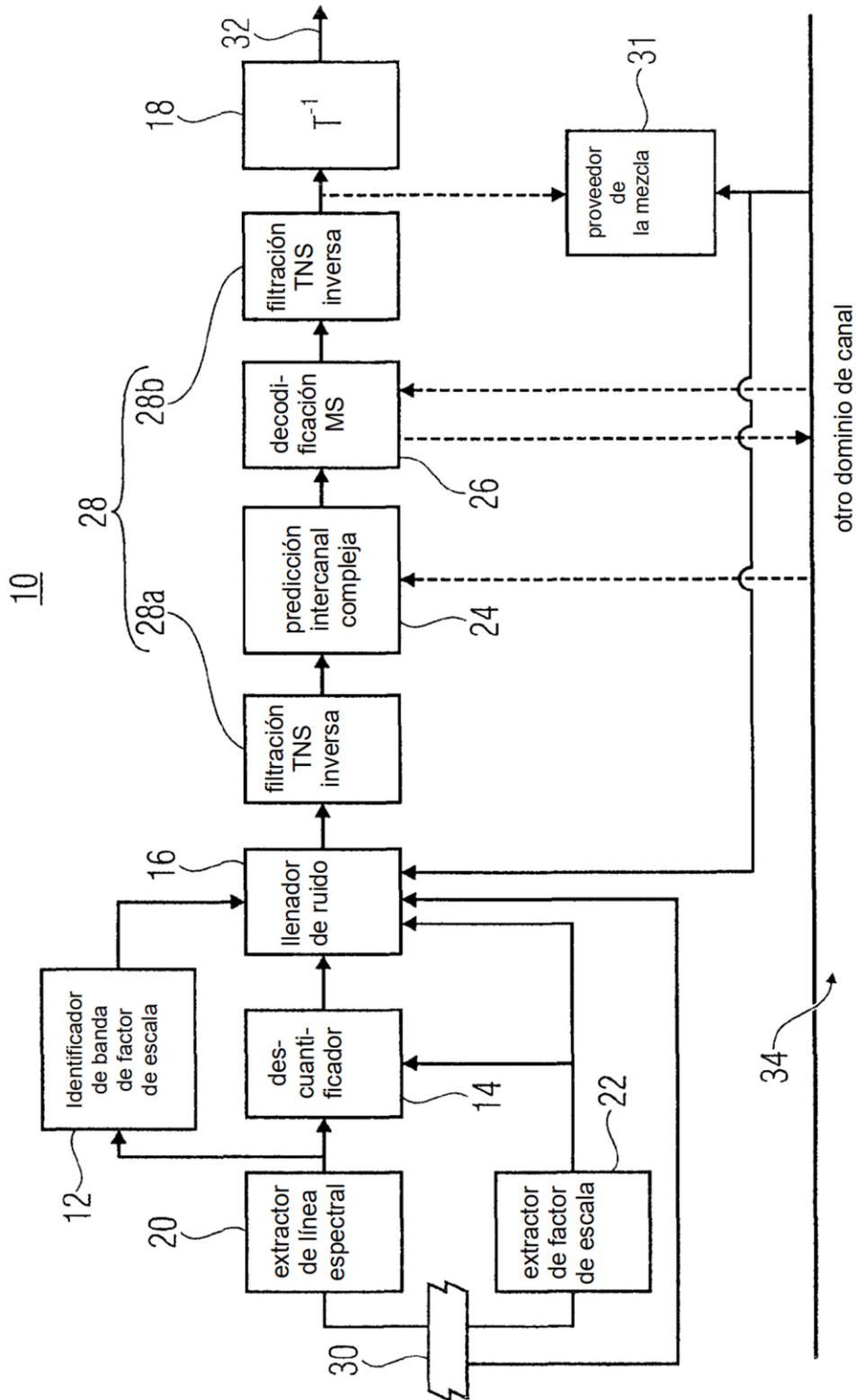


Fig. 2

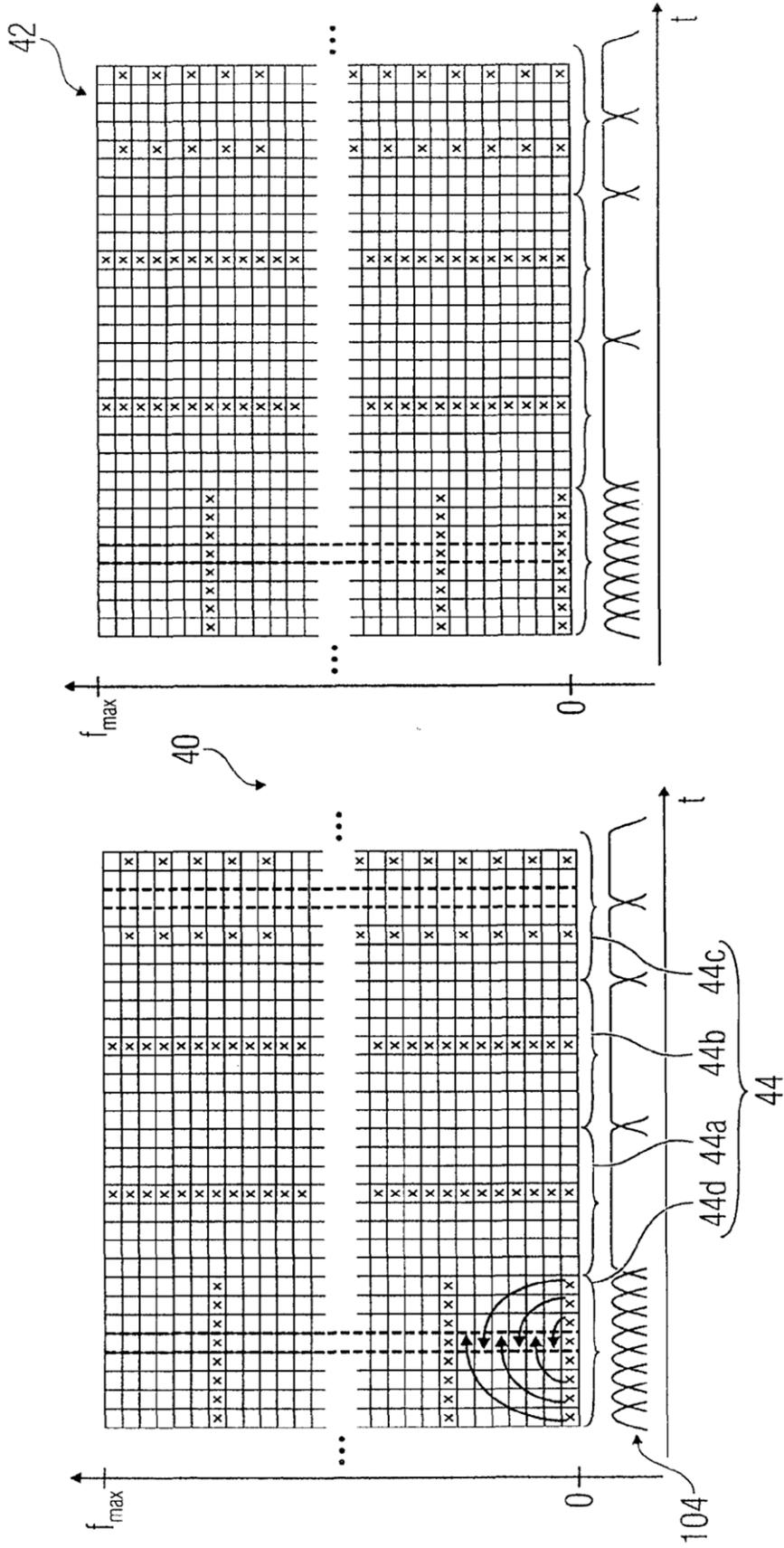


Fig. 3

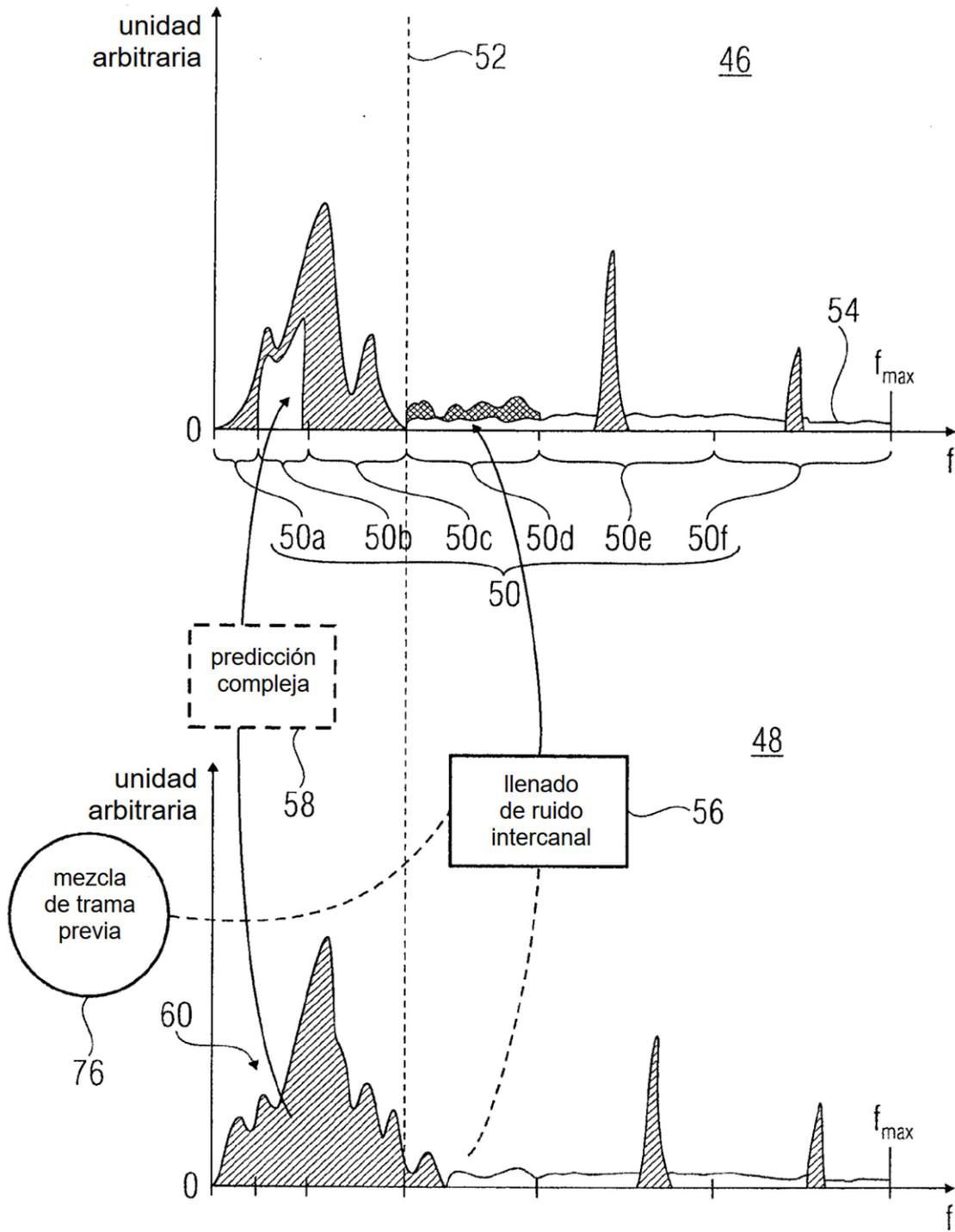


Fig. 4

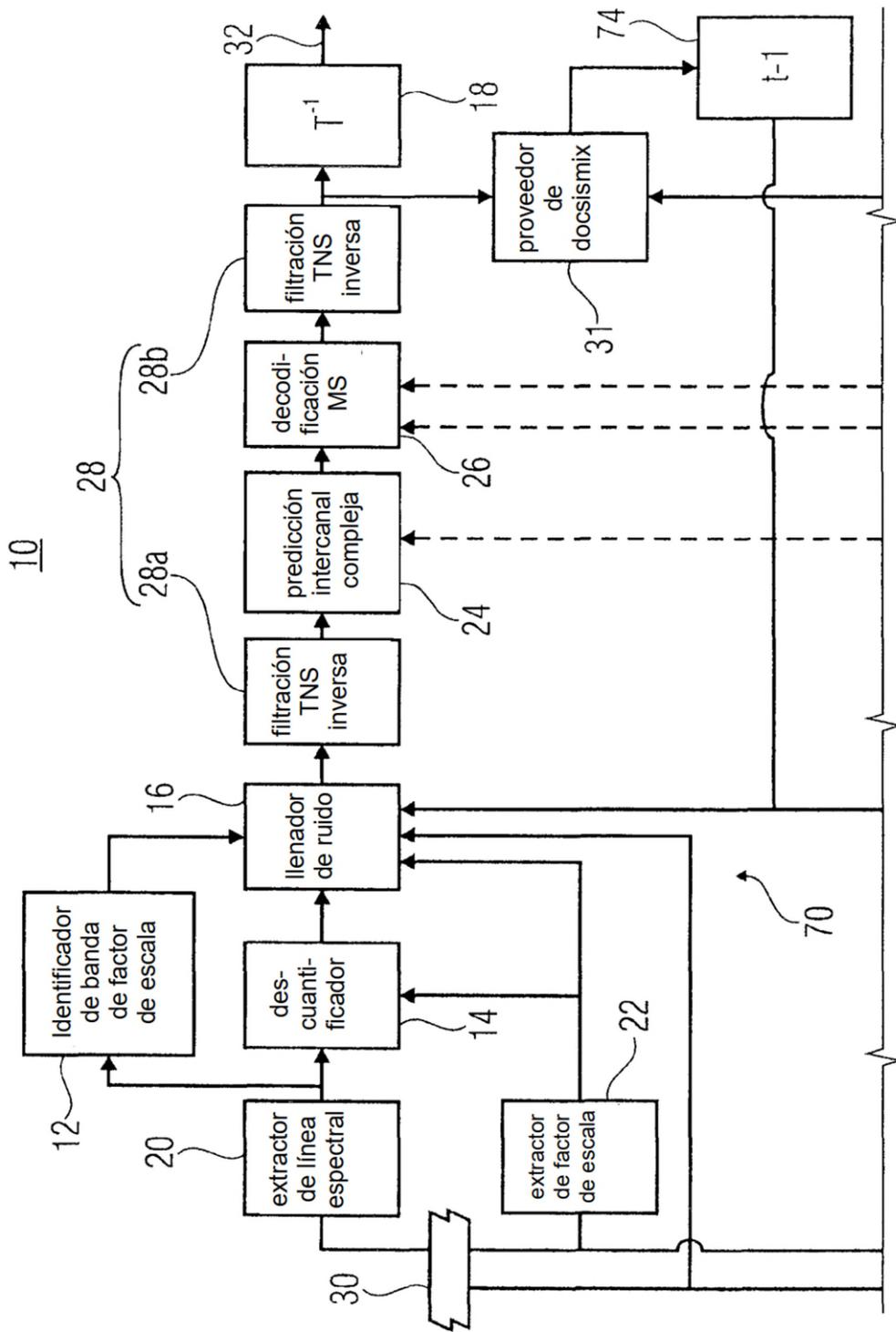


Fig. 5	Fig. 5a
	Fig. 5b

Fig. 5a

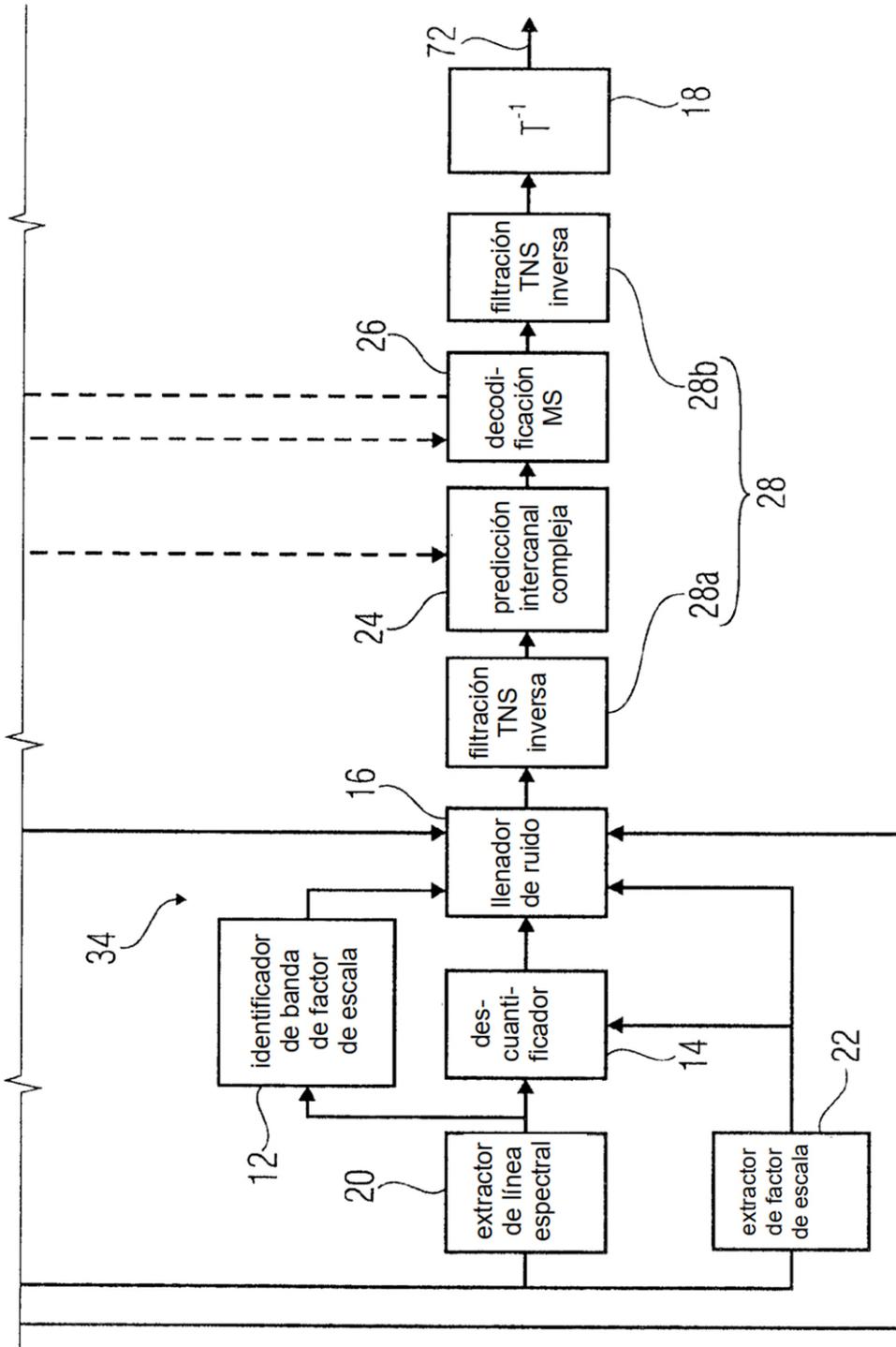


Fig. 5	Fig. 5a
	Fig. 5b

Fig. 5b

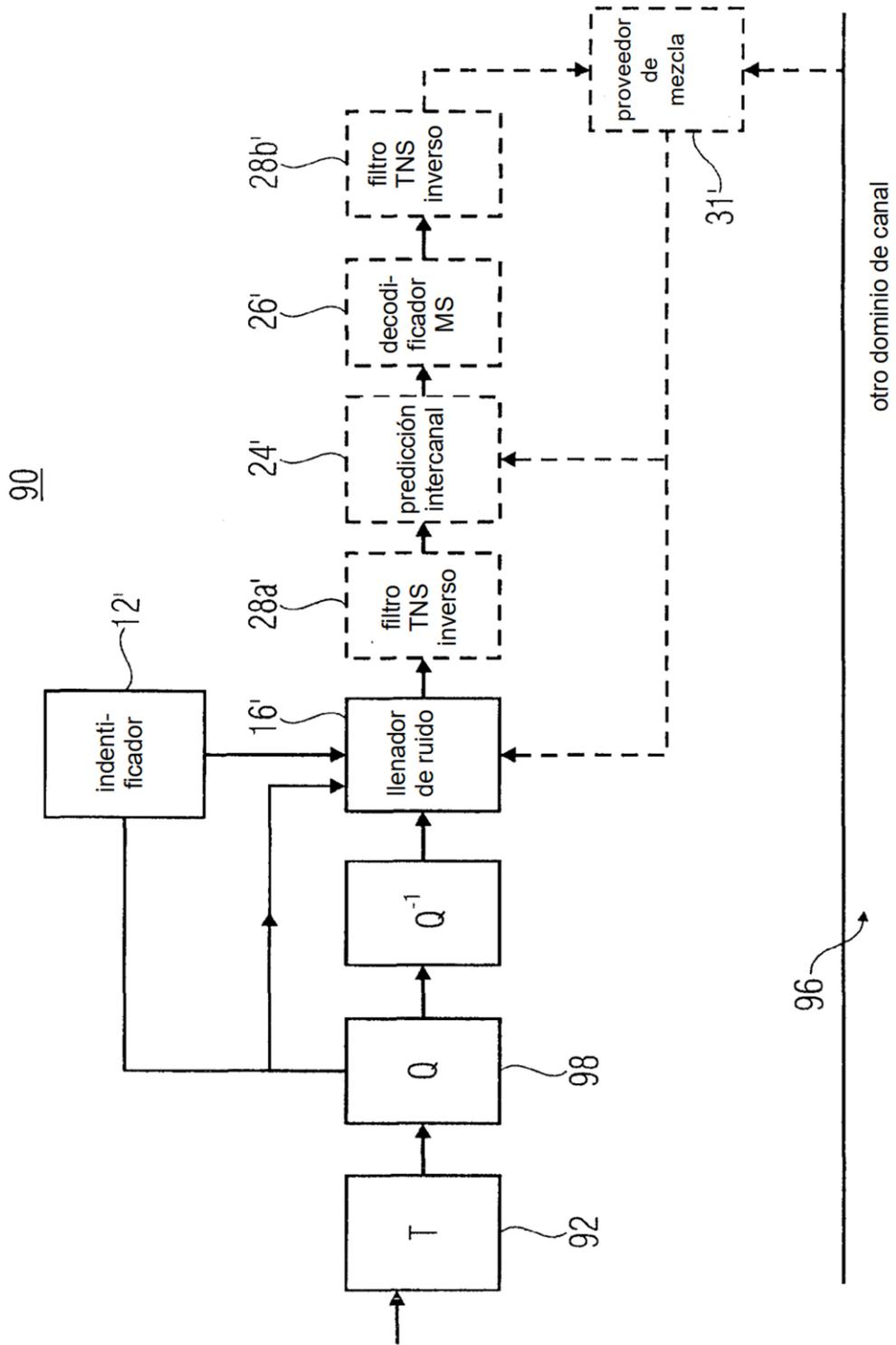


Fig. 6

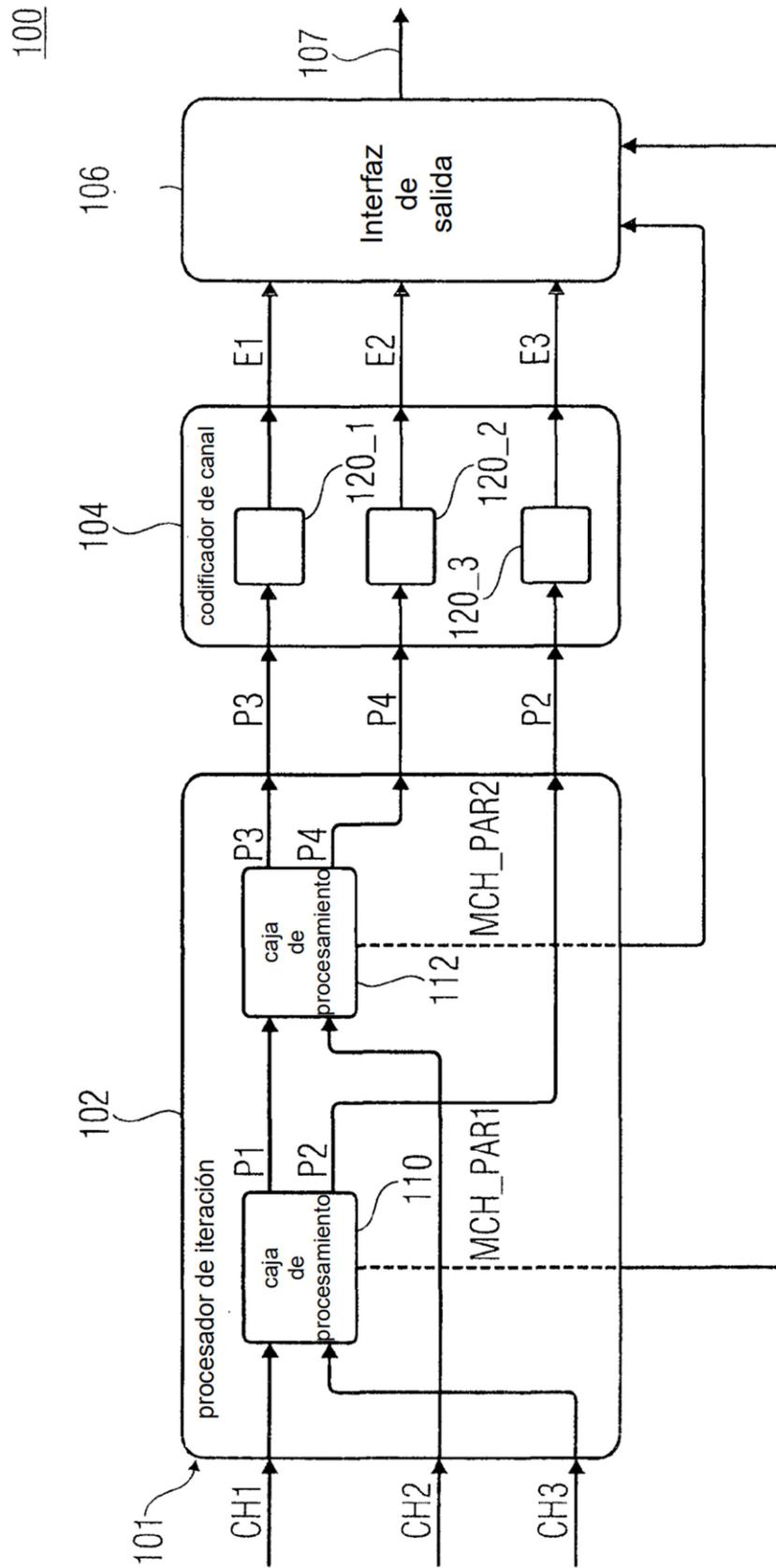


Fig. 7

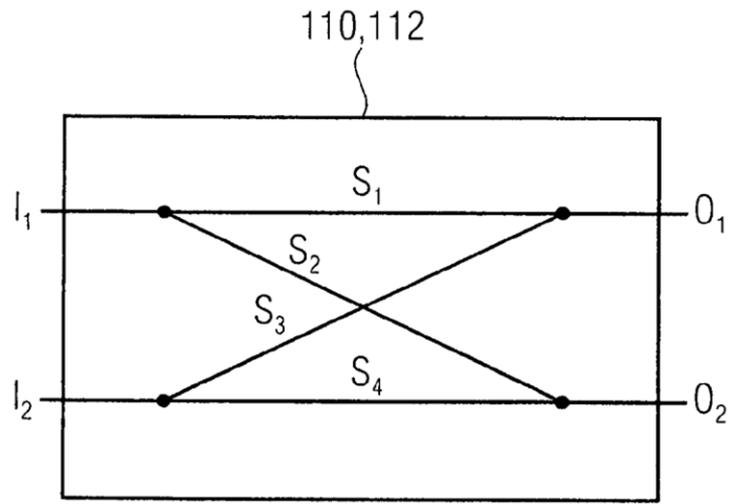


Fig. 8

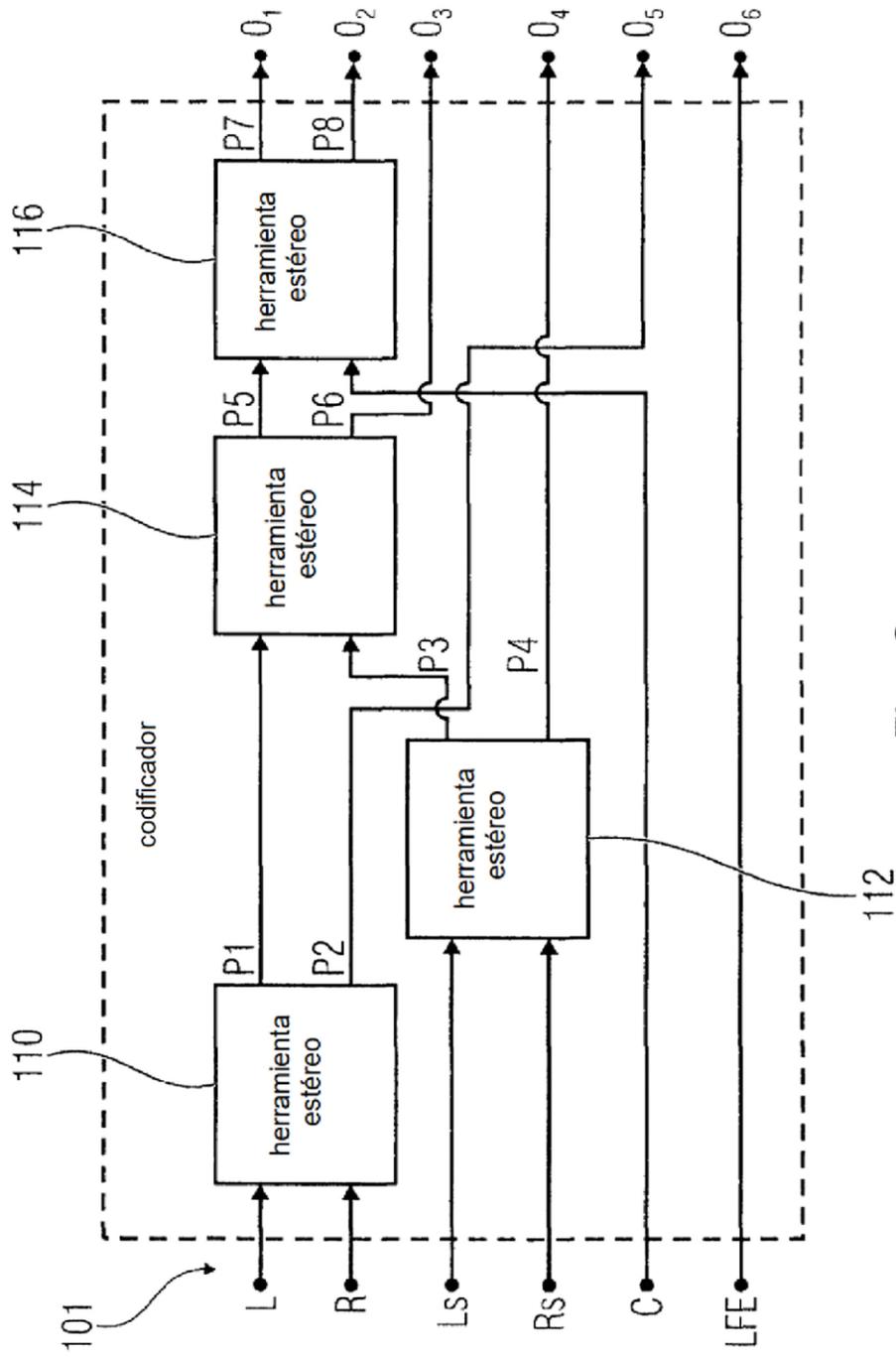


Fig. 9

200

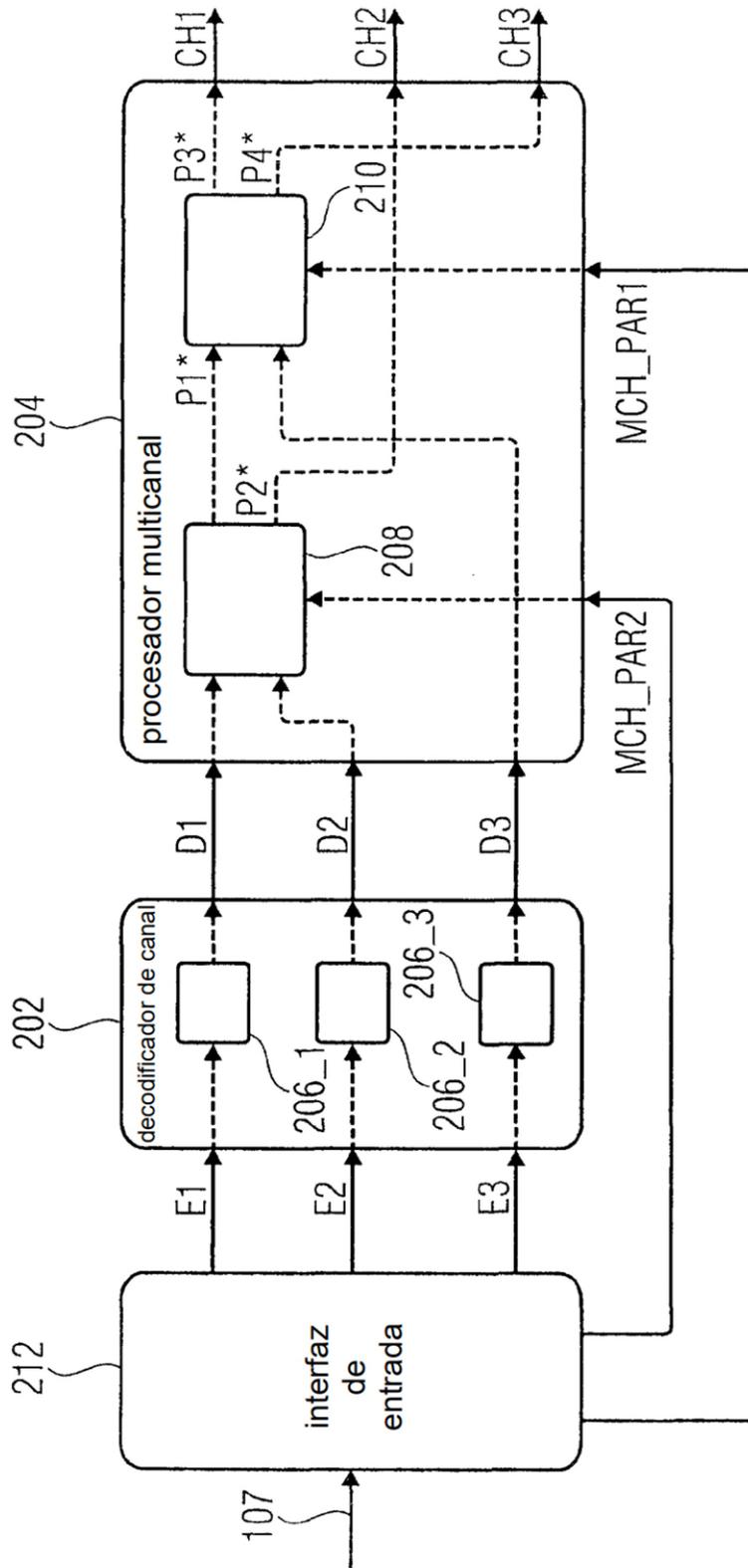


Fig. 10

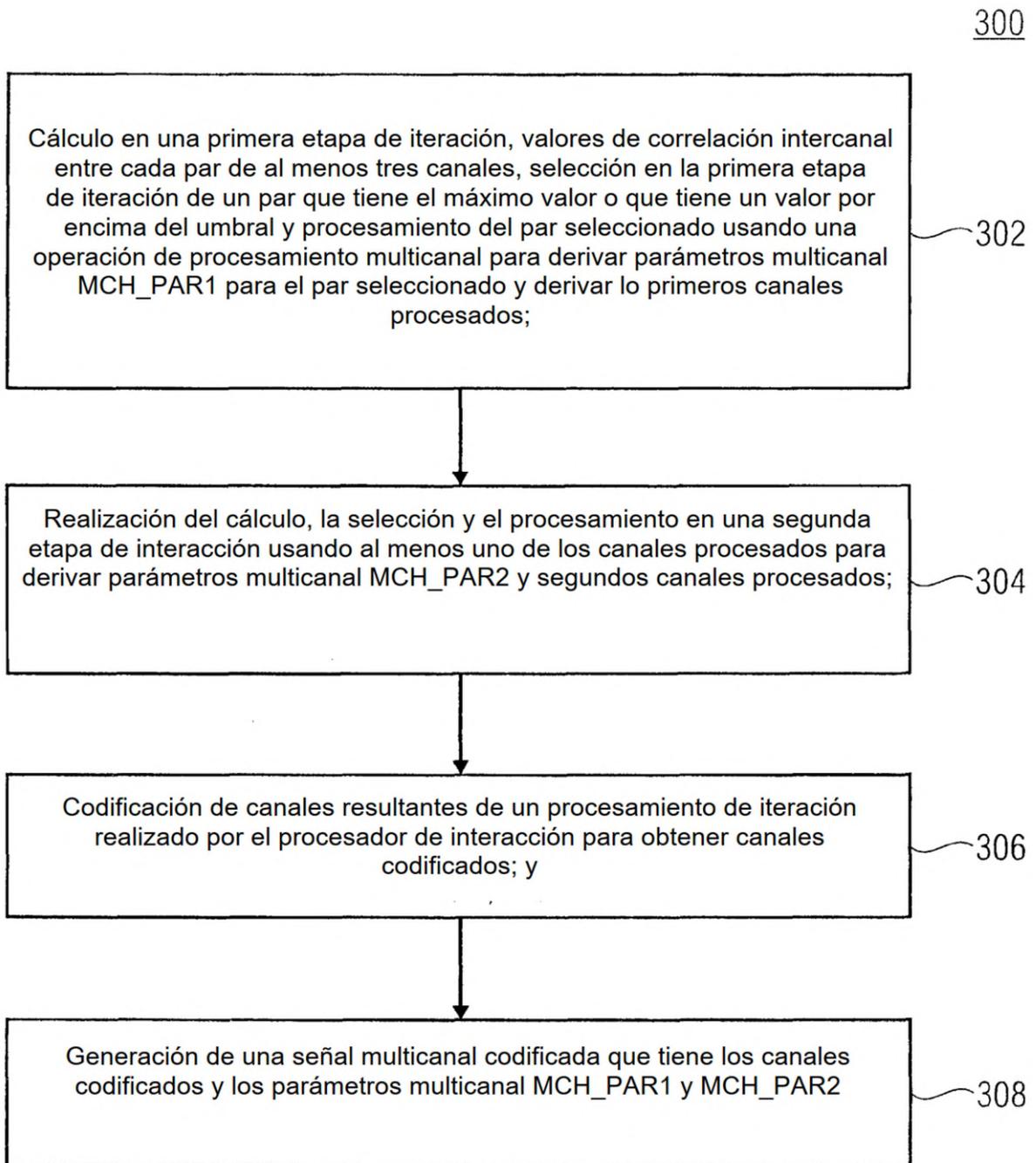


Fig. 11

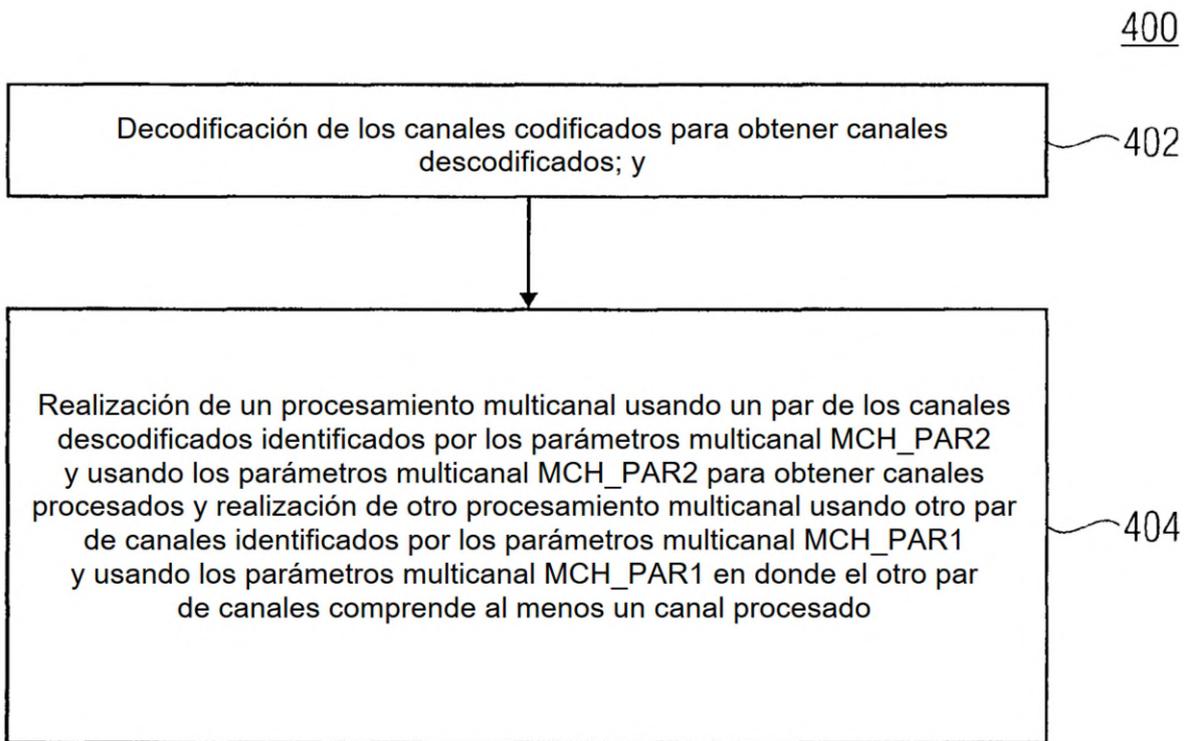


Fig. 12

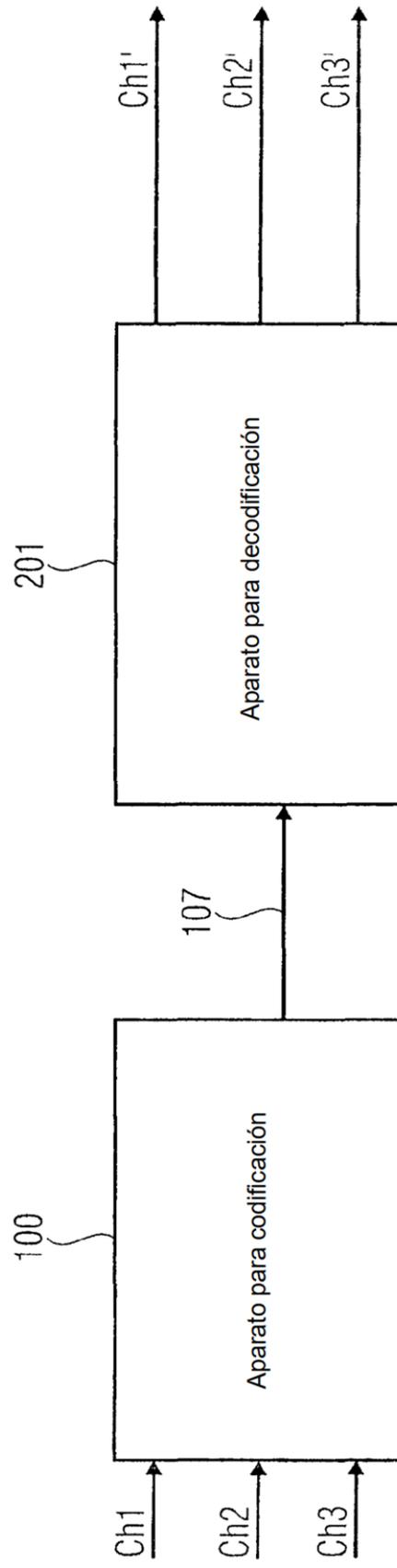


Fig. 13

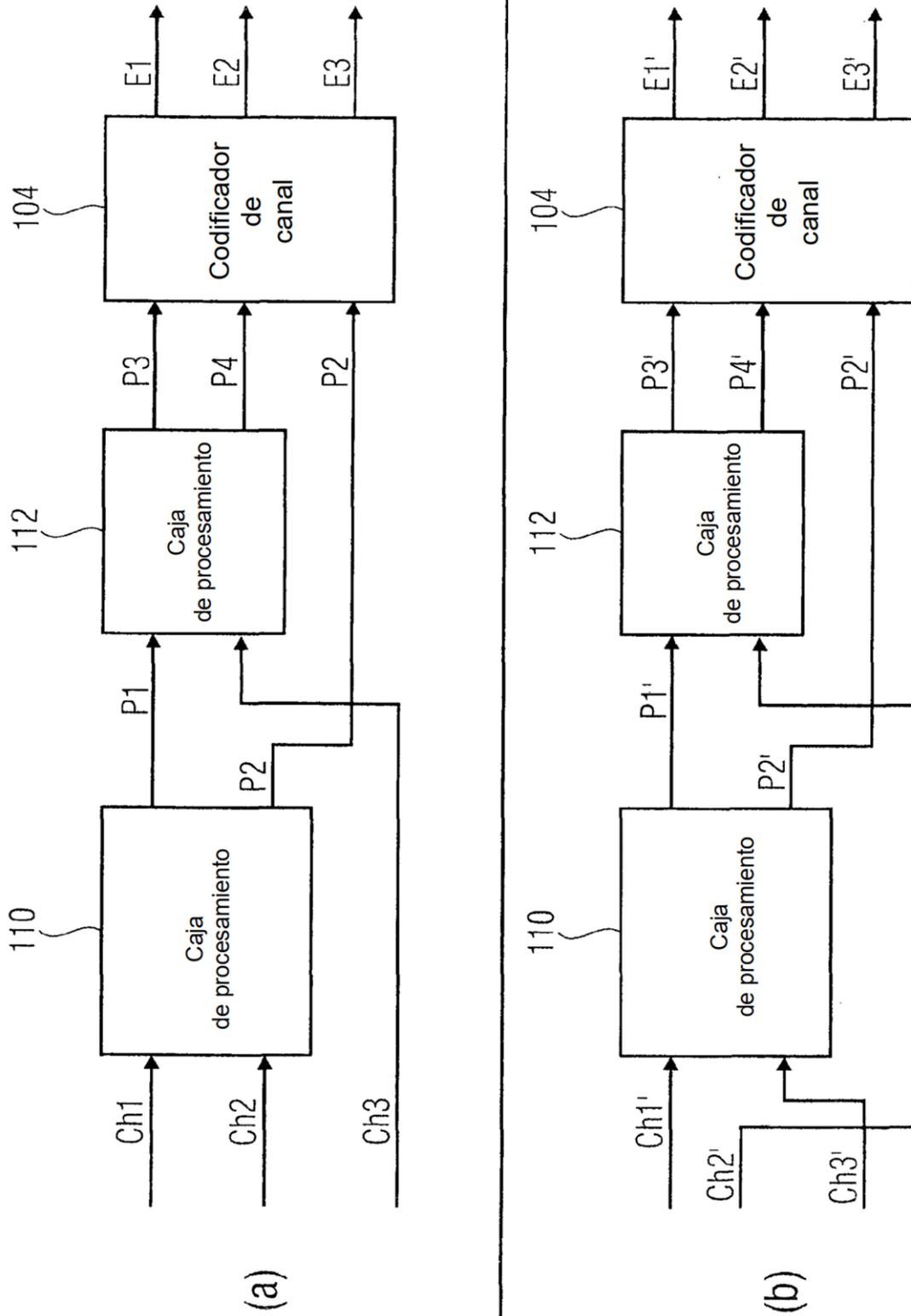


Fig. 14

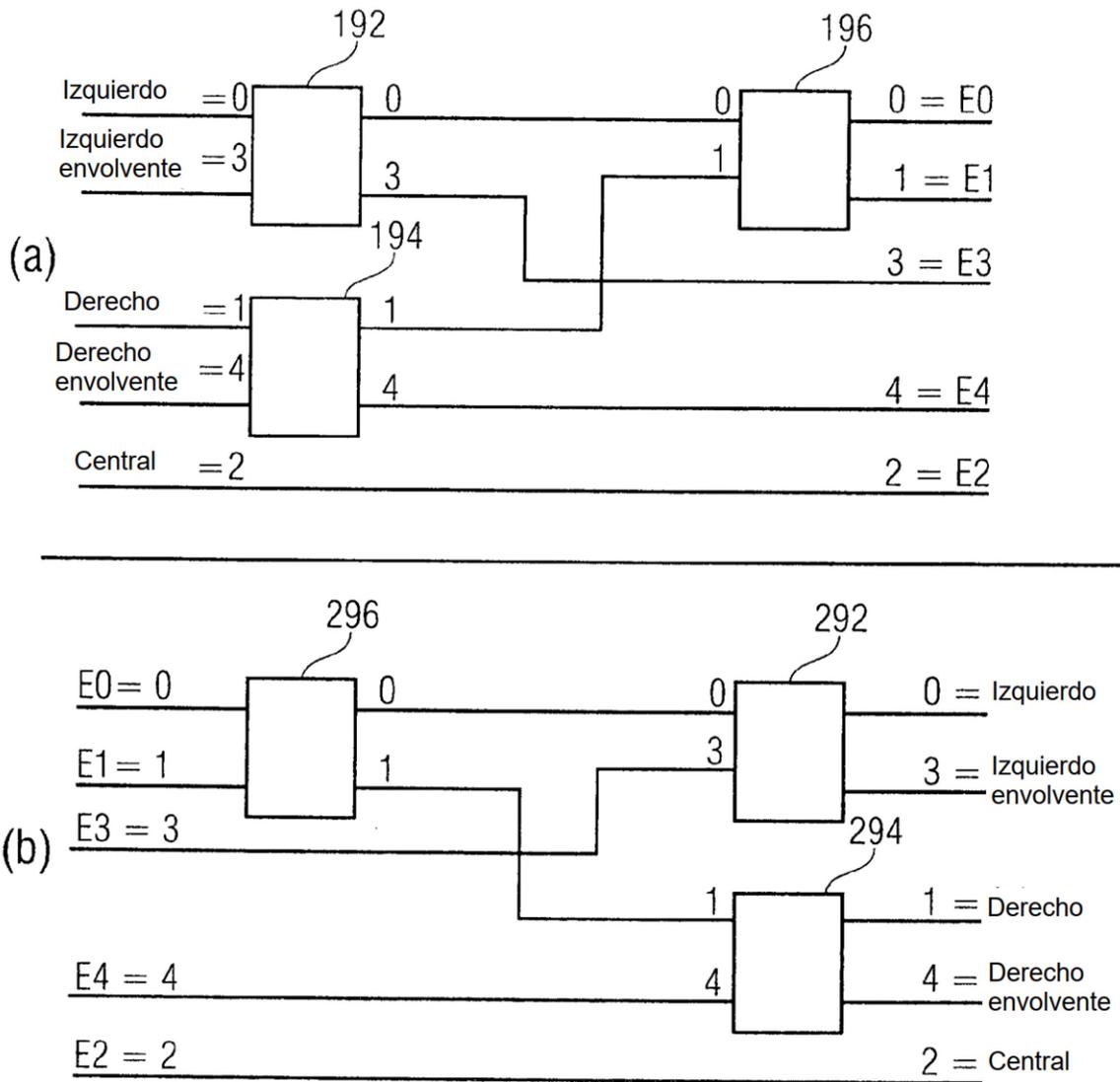


Fig. 15