

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 649 194**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 21/038 (2013.01)

G10L 19/00 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.07.2014 PCT/EP2014/065021**

87 Fecha y número de publicación internacional: **29.01.2015 WO15010934**

96 Fecha de presentación y número de la solicitud europea: **14.07.2014 E 14738535 (5)**

97 Fecha y número de publicación de la concesión europea: **23.08.2017 EP 3022734**

54 Título: **Decodificador de audio, codificador de audio, procedimiento para proporcionar al menos cuatro señales de canales de audio sobre la base de una representación codificada, procedimiento para proporcionar una representación codificada sobre la base de al menos cuatro señales de canales de audio y programa informático que utiliza una extensión de ancho de banda**

30 Prioridad:

22.07.2013 EP 13177376

18.10.2013 EP 13189306

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

10.01.2018

73 Titular/es:

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

72 Inventor/es:

**DICK, SASCHA;
ERTEL, CHRISTIAN;
HELMRICH, CHRISTIAN;
HILPERT, JOHANNES;
HÖLZER, ANDREAS y
KUNTZ, ACHIM**

74 Agente/Representante:

SALVA FERRER, Joan

ES 2 649 194 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

Decodificador de audio, codificador de audio, procedimiento para proporcionar al menos cuatro señales de canales de audio sobre la base de una representación codificada, procedimiento para proporcionar una representación codificada sobre la base de al menos cuatro señales de canales de audio y programa informático que utiliza una extensión de ancho de banda

CAMPO TÉCNICO

- 10 **[0001]** Una realización según la invención crea un decodificador de audio que proporciona al menos cuatro señales de canales con ancho de banda extendido sobre la base de una representación codificada.
- [0002]** Otra realización según la invención crea un codificador de audio que proporciona una representación codificada sobre la base de al menos cuatro señales de canales de audio.
- 15 **[0003]** Otra realización según la invención crea un procedimiento que proporciona al menos cuatro señales de canales de audio sobre la base de una representación codificada.
- [0004]** Otra realización según la invención crea un procedimiento que proporciona una representación codificada sobre la base de al menos cuatro señales de canales de audio.
- 20 **[0005]** Otra realización según la invención crea un programa informático para realizar uno de los procedimientos.
- 25 **[0006]** En general, las realizaciones según la invención se refieren a una codificación conjunta de n canales.

ANTECEDENTES DE LA INVENCION

- 30 **[0007]** En los años recientes, se ha ido incrementando de modo constante una demanda por el almacenamiento y la transmisión de contenidos de audio. Más aún, los requerimientos de calidad respecto del almacenamiento y la transmisión de contenidos de audio también se ha ido incrementando constantemente. Conforme a ello, los conceptos para la codificación y la decodificación del contenido de audio se han mejorado. Por ejemplo, se ha desarrollado la así llamada "codificación de audio avanzada" (AAC), que se describe, por ejemplo, en el estándar internacional ISO/IEC 13818-7:2003. Más aún, se han creado algunas extensiones espaciales como, por ejemplo, el así llamado concepto de "Surround MPEG" que se describe, por ejemplo, en el estándar internacional ISO/IEC 23003-1:2007. Más aún, se describen mejoras adicionales para la codificación y decodificación de información espacial de señales de audio en el estándar internacional ISO/IEC 23003-2:2010, que se refiere a la codificación de objetos de audio espacial (SAOC).
- 35 **[0008]** Más aún, un concepto de codificación/decodificación de audio flexible, que proporciona la posibilidad de codificar tanto las señales generales de audio como las señales de voz con buena eficacia de codificación para manipular las señales de audio multicanal, se define en el estándar internacional ISO/IEC 23003-3:2012, que describe así el llamado concepto de "codificación de voz y audio unificados" (USAC).
- 40 **[0009]** En MPEG USAC [1], la codificación estéreo conjunta de dos canales se lleva a cabo utilizando una predicción compleja, MPS 2-1-1 o estéreo unificado con señales residuales de banda limitada o de banda completa. El Surround MPEG [2] combina jerárquicamente las cajas OTT y TTT para la codificación conjunta de audio multicanal con o sin transmisión de señales residuales. Además, el documento US 2012/0070007 A1 describe la codificación/decodificación multicanal con extensión de ancho de banda.
- 45 **[0009]** En MPEG USAC [1], la codificación estéreo conjunta de dos canales se lleva a cabo utilizando una predicción compleja, MPS 2-1-1 o estéreo unificado con señales residuales de banda limitada o de banda completa. El Surround MPEG [2] combina jerárquicamente las cajas OTT y TTT para la codificación conjunta de audio multicanal con o sin transmisión de señales residuales. Además, el documento US 2012/0070007 A1 describe la codificación/decodificación multicanal con extensión de ancho de banda.
- 50 **[0010]** Sin embargo, hay un deseo de proporcionar un concepto incluso más avanzado concepto de una codificación eficaz y decodificación de escenas de audio tridimensionales.

RESUMEN DE LA INVENCION

- 55 **[0011]** Una realización según la invención crea un decodificador de audio que proporciona al menos cuatro señales de canales con ancho de banda extendido sobre la base de una representación codificada. El decodificador de audio está configurado para proporcionar una primera señal de mezcla descendente y una segunda señal de mezcla descendente sobre la base de una representación codificada conjuntamente de la primera señal de mezcla descendente y la segunda señal de mezcla descendente utilizando una (primera) decodificación multicanal. El

5 decodificador de audio está configurado para proporcionar al menos una primera señal del canal de audio y una
 segunda señal del canal de audio sobre la base de la primera señal de mezcla descendente utilizando una (segunda)
 decodificación multicanal y para proporcionar al menos una tercera señal de canal de audio y una cuarta señal de
 canal de audio sobre la base de la segunda señal de mezcla descendente utilizando una (tercera) decodificación
 10 multicanal. El decodificador de audio está configurado para realizar una extensión de ancho de banda multicanal sobre
 la base de la primera señal del canal de audio y la tercera señal del canal de audio con el fin de obtener una primera
 señal del canal con ancho de banda extendido y una tercera señal del canal con ancho de banda extendido. Más aún,
 el decodificador de audio está configurado para realizar una extensión de ancho de banda multicanal sobre la base de
 la segunda señal del canal de audio y la cuarta señal del canal de audio con el fin de obtener una segunda señal del
 15 canal con ancho de banda extendido y una cuarta señal del canal con ancho de banda extendido.

[0012] Esta realización según la invención está basada en el hallazgo de que se pueden obtener resultados de
 extensión de ancho de banda particularmente buenos en un decodificador de audio jerárquico si se utilizan señales
 de canales de audio, que se obtienen sobre la base de diferentes señales de mezcla descendente en la segunda etapa
 15 del decodificador de audio, en una extensión de ancho de banda multicanal, en el que las diferentes señales de mezcla
 descendente se obtienen de una representación codificada conjuntamente en una primera etapa del decodificador de
 audio. Se ha encontrado que se puede obtener una calidad de audio particularmente buena si se separan señales de
 mezcla descendente, que se asocian con posiciones perceptualmente importantes en particular de una escena de
 20 audio, en la primera etapa de un decodificador de audio jerárquico, mientras que las posiciones especiales que no son
 tan importantes para una impresión auditiva, se separan en una segunda etapa del decodificador de audio jerárquico.
 Más aún, se ha encontrado que las señales de canales de audio, que se asocian con diferentes posiciones
 perceptualmente importantes de una escena de audio (por ejemplo, posiciones de la escena de audio, en las que la
 25 relación entre las señales desde dichas posiciones es perceptualmente importante) se deberá procesar conjuntamente
 en una extensión de ancho de banda multicanal, porque la extensión del ancho de banda multicanal puede tener en
 cuenta, en consecuencia, dependencias y diferencias entre las señales de estas posiciones auditivas importantes.
 Esto se consigue realizando la extensión del ancho de banda multicanal sobre la base de la primera señal del canal
 de audio (que se deriva de la primera señal de mezcla descendente en la segunda etapa del decodificador de audio
 30 jerárquico) y sobre la base de la tercera señal del canal de audio, que se deriva de la segunda señal de mezcla
 descendente en la segunda etapa del decodificador de audio jerárquico con el fin de obtener dos señales de canales
 con ancho de banda extendido (concretamente, la primera señal del canal de ancho de banda extendida y la tercera
 señal del canal de ancho de banda extendida). Conforme a ello, la extensión (conjunta) de ancho de banda multicanal
 se lleva a cabo sobre la base de señales de canales de audio que se derivan de diferentes señales de mezcla
 35 descendente en la segunda etapa del decodificador multicanal jerárquico, de manera que una relación entre la primera
 señal del canal de audio y la tercera señal del canal de audio es similar a (o determinada por) una relación entre la
 primera señal de mezcla descendente y la segunda señal de mezcla descendente. De esta manera, la extensión del
 ancho de banda multicanal puede usar esta relación (por ejemplo, entre la primera señal del canal de audio y la tercera
 40 señal del canal de audio), que se determina sustancialmente por la derivación de la primera señal de mezcla
 descendente y la segunda señal de mezcla descendente de la representación codificada conjuntamente de la primera
 señal de mezcla descendente y de la segunda señal de mezcla descendente usando la decodificación multicanal, que
 se lleva a cabo en la primera etapa del decodificador de audio. Conforme a ello, la extensión del ancho de banda
 multicanal puede aprovechar esta relación que se puede reproducir con buena precisión en la primera etapa del
 decodificador de audio jerárquico, de manera que se consigue una impresión auditiva particularmente buena.

[0013] En una realización preferente, la primera señal de mezcla descendente y la segunda señal de mezcla
 45 descendente se asocian con diferentes posiciones horizontales (o posiciones azimutales) de una escena de audio. Se
 ha encontrado que la diferenciación entre diferentes posiciones de audio horizontales (o posiciones azimutales) es
 particularmente significativa, dado que el sistema auditivo humano es particularmente sensible con respecto a
 diferentes posiciones horizontales. Conforme a ello, es ventajoso separar entre señales de mezcla descendente
 50 asociadas con diferentes posiciones horizontales de la escena de audio en la primera etapa del decodificador de audio
 jerárquico porque el procesamiento en la primera etapa del decodificador de audio jerárquico normalmente es más
 preciso que el procesamiento en las etapas posteriores. Más aún, como una consecuencia, la primera señal del canal
 de audio y la tercera señal del canal de audio, que se utilizan conjuntamente en la (primera) extensión de ancho de
 banda multicanal se asocian con diferentes posiciones horizontales de la escena de audio (porque la primera señal
 del canal de audio se obtiene de la primera señal de mezcla descendente y la tercera señal del canal de audio se
 55 deriva de la segunda señal de mezcla descendente en la segunda etapa del decodificador de audio jerárquico), que
 permite que la (primera) extensión de ancho de banda multicanal esté bien adaptada a la capacidad humana de
 distinguir entre diferentes posiciones horizontales. De modo similar, la (segunda) extensión de ancho de banda
 multicanal, que se lleva a cabo sobre la base de la segunda señal del canal de audio y la cuarta señal del canal de
 audio, trabaja con señales de canales de audio que se asocian con diferentes posiciones horizontales de la escena

de audio, de manera que la (segunda) extensión de ancho de banda multicanal también se puede adaptar bien a la importante relación psicoacústica entre las señales de canales de audio asociadas con las diferentes posiciones horizontales de la escena de audio. Conforme a ello, se puede conseguir una impresión auditiva particularmente buena.

5

[0014] En una realización preferente, la primera señal de mezcla descendente está asociada con un lado izquierdo de una escena de audio y la segunda señal de mezcla descendente está asociada con un lado derecho de la escena de audio. En consecuencia, la primera señal del canal de audio normalmente también está asociada con el lado izquierdo de la escena de audio y la tercera señal del canal de audio está asociada con el lado derecho de la escena de audio, de manera que la (primera) extensión de ancho de banda multicanal trabaja (preferentemente conjuntamente) con señales de canales de audio de diferentes lados de la escena de audio y, en consecuencia, se puede adaptar bien a la percepción izquierda/derecha humana. Lo mismo también es válido para la (segunda) extensión de ancho de banda multicanal, que trabaja con la base de la segunda señal del canal de audio y la cuarta señal del canal de audio.

10

15

[0015] En una realización preferente, la primera señal del canal de audio y la segunda señal del canal de audio se asocian con posiciones vecinas verticalmente de una escena de audio. De modo similar, la tercera señal del canal de audio y la cuarta señal del canal de audio se asocian con posiciones vecinas verticalmente de la escena de audio. Se ha encontrado que es ventajoso separar entre las señales de canales de audio asociadas con posiciones vecinas verticalmente de la escena de audio en la segunda etapa del decodificador de audio jerárquico. Más aún, se ha encontrado que las señales de canales de audio normalmente no se degradan gravemente al separar entre las señales de canales de audio asociadas con posiciones vecinas verticalmente, de manera que las señales de entrada para las extensiones de ancho de banda multicanal son muy adecuadas para una extensión de ancho de banda multicanal (por ejemplo, una extensión del ancho de banda estéreo).

20

25

[0016] En una realización preferente, la primera señal del canal de audio y la tercera señal del canal de audio se asocian con un primer plano horizontal común (o una primera elevación común) de una escena de audio pero diferentes posiciones horizontales (o posiciones azimutales) de la escena de audio y la segunda señal del canal de audio y la cuarta señal del canal de audio se asocian con un segundo plano horizontal común (o una segunda elevación común) de una escena de audio pero diferentes posiciones horizontales (o posiciones azimutales) de la escena de audio. En este caso, el primer plano horizontal común (o elevación) es diferente del segundo plano horizontal común (o elevación). Se ha encontrado que la extensión del ancho de banda multicanal se puede realizar con resultados de calidad particularmente buenos sobre la base de dos señales de canales de audio que se asocian con el mismo plano horizontal (o elevación).

30

35

[0017] En una realización preferente, la primera señal del canal de audio y la segunda señal del canal de audio se asocian con un primer plano vertical común (o posición azimutal común) de la escena de audio pero diferentes posiciones verticales (o elevaciones) de la escena de audio. De modo similar, la tercera señal del canal de audio y la cuarta señal del canal de audio se asocian con un segundo plano vertical común (o posición azimutal común) de la escena de audio pero diferentes posiciones verticales (o elevaciones) de la escena de audio. En este caso, el primer plano vertical común (o posición azimutal) es, preferentemente, diferente del segundo plano vertical común (o posición azimutal). Se ha encontrado que una división (o separación) de las señales de canales de audio asociadas con un plano vertical común (o posición azimutal) se puede realizar con buenos resultados usando la segunda etapa del decodificador de audio jerárquico, mientras que la separación (o división) entre las señales de canales de audio asociadas con diferentes planos verticales (o posiciones azimutales) se puede realizar con buenos resultados de calidad usando la primera etapa del decodificador de audio jerárquico.

40

45

[0018] En una realización preferente, la primera señal del canal de audio y la segunda señal del canal de audio se asocian con un lado izquierdo de una escena de audio y la tercera señal del canal de audio y la cuarta señal del canal de audio se asocian con un lado derecho de la escena de audio. Esta configuración permite una extensión de ancho de banda multicanal particularmente buena, que usa una relación entre una señal del canal de audio asociada con un lado izquierdo y una señal del canal de audio asociada con un lado derecho y, en consecuencia, se adapta bien a la capacidad humana para distinguir entre el sonido que procede del lado izquierdo y el sonido que procede del lado derecho.

50

55

[0019] En una realización preferente, la primera señal del canal de audio y la tercera señal del canal de audio se asocian con una parte inferior de la escena de audio y la segunda señal del canal de audio y la cuarta señal del canal de audio se asocian con una parte superior de la escena de audio. Se ha encontrado que dicha asignación espacial de las señales de canales de audio trae consigo resultados auditivos particularmente buenos.

[0020] En una realización preferente, el decodificador de audio está configurado para realizar una división horizontal cuando se proporciona la primera señal de mezcla descendente y la segunda señal de mezcla descendente sobre la base de la representación codificada conjuntamente de la primera señal de mezcla descendente y la segunda señal de mezcla descendente usando la decodificación multicanal. Se ha encontrado que la realización de una división horizontal la primera etapa del decodificador de audio jerárquico da como resultado una impresión auditiva particularmente buena porque el procesamiento realizado en la primera etapa del decodificador de audio jerárquico normalmente se puede realizar con un mayor rendimiento que el procesamiento realizado en la segunda etapa del decodificador de audio jerárquico. Más aún, la realización de la división horizontal en la primera etapa del decodificador de audio da como resultado una buena impresión auditiva, porque el sistema auditivo humano es más sensible con respecto a una posición horizontal de un objeto de audio cuando se compara con una posición vertical del objeto de audio.

[0021] En una realización preferente, el decodificador de audio está configurado para realizar una división vertical cuando se proporciona al menos la primera señal del canal de audio y la segunda señal del canal de audio sobre la base de la primera señal de mezcla descendente usando la decodificación multicanal. De modo similar, el decodificador de audio está configurado, preferentemente, para realizar una división vertical cuando se proporciona al menos la tercera señal del canal de audio y la cuarta señal del canal de audio sobre la base de la segunda señal de mezcla descendente usando la decodificación multicanal. Se ha encontrado que la realización de la división vertical en la segunda etapa del decodificador jerárquico trae consigo una buena impresión auditiva, dado que el sistema auditivo humano no es particularmente sensible a la posición vertical de una fuente de audio (u objeto de audio).

[0022] En una realización preferente, el decodificador de audio está configurado para realizar una extensión de ancho de banda estéreo sobre la base de la primera señal del canal de audio y la tercera señal del canal de audio con el fin de obtener la primera señal del canal de ancho de banda extendida y la tercera señal del canal de ancho de banda extendida, en el que la primera señal del canal de audio y la tercera señal del canal de audio representan un primer par de canales izquierdo/derecho. De modo similar, el decodificador de audio está configurado para realizar una extensión de ancho de banda estéreo sobre la base de la segunda señal del canal de audio y la cuarta señal del canal de audio con el fin de obtener la segunda señal del canal con ancho de banda extendido y la tercera señal del canal con ancho de banda extendido, en el que la segunda señal del canal de audio y la cuarta señal del canal de audio representan un segundo par de canales izquierdo/derecho. Se ha encontrado que una extensión de ancho de banda estéreo da como resultado una impresión auditiva particularmente buena porque la extensión del ancho de banda estéreo puede tener en cuenta la relación entre el canal estéreo izquierdo y un canal estéreo derecho y realizar la extensión de ancho de banda en dependencia de la relación.

[0023] En una realización preferente, el decodificador de audio está configurado para proporcionar la primera señal de mezcla descendente y la segunda señal de mezcla descendente sobre la base de una representación codificada conjuntamente de la primera señal de mezcla descendente y la segunda señal de mezcla descendente utilizando una decodificación multicanal basada en predicciones. Se ha encontrado que el uso de una decodificación multicanal a base de predicciones en la primera etapa del decodificador de audio jerárquico trae consigo un buen equilibrio entre la tasa de bits y la calidad. Se ha encontrado que el uso de una predicción da como resultado una buena reconstrucción de diferencias entre la primera señal de mezcla descendente y la segunda señal de mezcla descendente, que es importante para una distinción izquierda/derecha de un objeto de audio.

[0024] Por ejemplo, el decodificador de audio se puede configurar para evaluar un parámetro de predicción que describe la contribución de un componente de señal que se obtiene utilizando un componente de señal de una trama anterior, a un suministro de las señales de mezcla descendente de la trama actual. Conforme a ello, la intensidad de la contribución del componente de señal, que se deriva utilizando un componente de señal de una trama anterior, se puede ajustar sobre la base de un parámetro, que se incluye en la representación codificada.

[0025] Por ejemplo, la decodificación multicanal a base de predicciones puede ser operativa en el dominio MDCT, de manera que la decodificación multicanal basada en predicciones se puede adaptar bien e interactuar fácilmente con una etapa de decodificación de audio que proporciona la señal de entrada a la decodificación multicanal que deriva la primera señal de mezcla descendente y la segunda señal de mezcla descendente. Con preferencia, pero no necesariamente, la decodificación multicanal basada en predicciones puede ser una predicción estéreo compleja USAC, que facilita la implementación del decodificador de audio.

[0026] En una realización preferente, el decodificador de audio está configurado para proporcionar la primera señal de mezcla descendente y la segunda señal de mezcla descendente sobre la base de una representación

codificada conjuntamente de la primera señal de mezcla descendente y la segunda señal de mezcla descendente utilizando una decodificación multicanal asistida con señales residuales. El uso de una decodificación multicanal asistida con señales residuales permite a particularmente precise reconstrucción de la primera señal de mezcla descendente y la segunda señal de mezcla descendente que, a su vez, mejora una percepción de la posición izquierda–derecha sobre la base de las señales de canales de audio y, en consecuencia, sobre la base de las señales del canal con ancho de banda extendido.

[0027] En una realización preferente, el decodificador de audio está configurado para proporcionar al menos la primera señal del canal de audio y la segunda señal del canal de audio sobre la base de la primera señal de mezcla descendente utilizando una decodificación multicanal basada en parámetros. Más aún, el decodificador de audio está configurado para proporcionar al menos la tercera señal del canal de audio y la cuarta señal del canal de audio sobre la base de la segunda señal de mezcla descendente utilizando una decodificación multicanal basada en parámetros. Se ha encontrado que el uso de una decodificación multicanal basada en parámetros es muy adecuada en la segunda etapa del decodificador de audio jerárquico. Se ha encontrado que una decodificación multicanal basada en parámetros conlleva un buen equilibrio entre la calidad de audio y la tasa de bits. Incluso si la calidad de reproducción de la decodificación multicanal basada en parámetros normalmente no es tan buena como la calidad de reproducción de una decodificación multicanal a base de predicciones (y posiblemente asistida con señales residuales), se ha encontrado que el uso de una decodificación multicanal basada en parámetros normalmente es suficiente, dado que el sistema auditivo humano no es particularmente sensible a la posición vertical (o elevación) de un objeto de audio que, preferentemente, se determina por dispersión (o separación) entre la primera señal del canal de audio y la segunda señal del canal de audio o entre la tercera señal del canal de audio y la cuarta señal del canal de audio.

[0028] En una realización preferente, la decodificación multicanal a base de parámetros está configurado para evaluar uno o varios parámetros que describen una correlación deseada (o covarianza) entre dos canales y/o diferencias de niveles entre dos canales a fin de proporcionar las dos o más señales de canales de audio sobre la base de una respectiva señal de mezcla descendente. Se ha encontrado que el uso de dichos parámetros que describen, por ejemplo, una correlación deseada entre dos canales y/o diferencias de niveles entre dos canales es muy adecuada para una división (o una separación) entre las señales del primer canal de audio y el segundo canal de audio (que normalmente están asociadas con diferentes posiciones verticales de una escena de audio) y para una división (o separación) entre la tercera señal del canal de audio y la cuarta señal del canal de audio (que también normalmente están asociadas con diferentes posiciones verticales).

[0029] Por ejemplo, la decodificación multicanal basada en parámetros puede ser operativa en un dominio QMF. Conforme a ello, la decodificación multicanal basada en parámetros se puede adaptar bien e interactuar fácilmente con la extensión de ancho de banda multicanal, que también puede trabajar preferentemente, pero no necesariamente, en el dominio QMF.

[0030] Por ejemplo, la decodificación multicanal basada en parámetros puede ser una decodificación Surround MPEG 2–1–2 o una decodificación estéreo unificada. El uso de dichos conceptos de codificación puede facilitar la implementación porque estos conceptos de decodificación pueden estar ya presentes en decodificadores de audio heredados.

[0031] En una realización preferente, el decodificador de audio está configurado para proporcionar al menos la primera señal del canal de audio y la segunda señal del canal de audio sobre la base de la primera señal de mezcla descendente utilizando una decodificación multicanal asistida con señales residuales. Más aún el decodificador de audio se puede configurar para proporcionar al menos la tercera señal del canal de audio y la cuarta señal del canal de audio sobre la base de la segunda señal de mezcla descendente utilizando una decodificación multicanal asistida con señales residuales. Con una decodificación multicanal asistida con señales residuales, la calidad de audio se puede mejorar incluso si bien la separación entre la primera señal del canal de audio y la segunda señal de audio y/o la separación entre la tercera señal del canal de audio y la cuarta señal del canal de audio se pueden realizar con una calidad particularmente alta.

[0032] En una realización preferente, el decodificador de audio se puede configurar para proporcionar una primera señal residual, que se utiliza para proporcionar al menos la primera señal del canal de audio y la segunda señal del canal de audio y una segunda señal residual, que se utiliza para proporcionar al menos la tercera señal del canal de audio y la cuarta señal del canal de audio, sobre la base de una representación codificada conjuntamente de la primera señal residual y la segunda señal residual utilizando una decodificación multicanal. Conforme a ello, el concepto para la decodificación jerárquica se puede extender al suministro de dos señales residuales, una de las que se utiliza para proporcionar la primera señal del canal de audio y la segunda señal del canal de audio (pero que

normalmente no se utiliza para proporcionar la tercera señal del canal de audio y la cuarta señal del canal de audio) y una de las que se utiliza para proporcionar la tercera señal del canal de audio y la cuarta señal del canal de audio (pero preferentemente no utilizada para proporcionar la primera señal del canal de audio y la segunda señal del canal de audio).

5

[0033] En una realización preferente, la primera señal residual y la segunda señal residual pueden estar asociadas con diferentes posiciones horizontales (o posiciones azimutales) de una escena de audio. Conforme a ello, el suministro de la primera señal residual y la segunda señal residual, que se lleva a cabo en la primera etapa del decodificador de audio jerárquico, puede realizar una división horizontal (o separación), en la que se ha encontrado que una división horizontal particularmente buena (o separación) se puede realizar en la primera etapa del decodificador de audio jerárquico (cuando se compara con el procesamiento realizado en la segunda etapa del decodificador de audio jerárquico). Conforme a ello, la separación horizontal, que es particularmente importante para el oyente, se lleva a cabo en la primera etapa de la decodificación de audio jerárquica, que proporciona una reproducción particularmente buena de manera que se pueda conseguir una buena impresión auditiva.

10
15

[0034] En una realización preferente, la primera señal residual está asociada con un lado izquierdo de una escena de audio y la segunda señal residual está asociada con un lado derecho de la escena de audio, que se ajusta a la sensibilidad posicional humana.

[0035] Una realización según la invención crea un codificador de audio que proporciona una representación codificada sobre la base de al menos cuatro señales de canales de audio. El codificador de audio está configurado para obtener un primer grupo de parámetros de extensión de ancho de banda comunes sobre la base de una primera señal del canal de audio y una tercera señal de canal de audio. El codificador de audio también está configurado para obtener un segundo grupo de parámetros de extensión de ancho de banda comunes sobre la base de una segunda señal del canal de audio y una cuarta señal de canal de audio. El codificador de audio está configurado para codificar conjuntamente al menos la primera señal del canal de audio y la segunda señal del canal de audio utilizando una codificación multicanal para obtener una primera señal de mezcla descendente y codificar conjuntamente al menos la tercera señal del canal de audio y la cuarta señal del canal de audio utilizando una codificación multicanal para obtener una segunda señal de mezcla descendente. Más aún, el codificador de audio está configurado para codificar conjuntamente la primera señal de mezcla descendente y la segunda señal de mezcla descendente utilizando una codificación multicanal con el fin de obtener una representación codificada de las señales de mezcla descendente.

[0036] Esta realización está basada en la idea de que el primer grupo de parámetros de extensión de ancho de banda comunes se debe obtener sobre la base de señales de canales de audio, que están representadas por diferentes señales de mezcla descendente que sólo se codifican conjuntamente en la segunda etapa del codificador de audio jerárquico. En paralelo con el decodificador de audio analizado anteriormente, la relación entre las señales de canales de audio que sólo se combinan en la segunda etapa de la codificación de audio jerárquica, se puede reproducir con una precisión particularmente alta del lado de un decodificador de audio. Conforme a ello, se ha encontrado que dos señales de audio que sólo se combinan eficazmente en la segunda etapa del codificador jerárquico son muy adecuadas para obtener un grupo de parámetros comunes de extensión de ancho de banda, dado que una extensión de ancho de banda multicanal se puede aplicar mejor a las señales de canales de audio, la relación entre ellas se reconstruye bien en el lado de un decodificador de audio. En consecuencia, se ha encontrado que es mejor, en términos de la calidad de audio que se puede conseguir, derivar un grupo de parámetros de extensión de ancho de banda comunes de dichas señales de canales de audio que sólo se combinan en la segunda etapa del codificador de audio jerárquico cuando se compara con la obtención de un grupo de parámetros de extensión de ancho de banda comunes a partir de dichas señales de canales de audio que se combinan en la primera etapa del codificador de audio jerárquico. Sin embargo, también se ha encontrado que se puede obtener una mejor calidad de audio derivando los grupos de parámetros de extensión de ancho de banda comunes de señales de canales de audio antes de codificarlas conjuntamente en la primera etapa del codificador de audio jerárquico.

50

[0037] En una realización preferente, la primera señal de mezcla descendente y la segunda señal de mezcla descendente se asocian con diferentes posiciones horizontales (o posiciones azimutales) de una escena de audio. Este concepto está basado en la idea de que una mejor impresión auditiva se puede conseguir si las señales que se asocian con diferentes posiciones horizontales se codifican sólo conjuntamente en la segunda etapa del codificador de audio jerárquico.

55

[0038] En una realización preferente, la primera señal de mezcla descendente está asociada con un lado izquierdo de una escena de audio y la segunda señal de mezcla descendente está asociada con un lado derecho de la escena de audio. De esta manera, dichas señales multicanal que se asocian con diferentes lados de la escena de

audio se utilizan para proporcionar los grupos de parámetros comunes de extensión de ancho de banda. En consecuencia, los grupos de parámetros de extensión de ancho de banda comunes están bien adaptados a la capacidad humana para distinguir entre fuentes de audio en lados diferentes.

5 **[0039]** En una realización preferente, la primera señal del canal de audio y la segunda señal del canal de audio se asocian con posiciones vecinas verticalmente de una escena de audio. Más aún, la tercera señal del canal de audio y la cuarta señal del canal de audio también están asociadas con posiciones vecinas verticalmente de la escena de audio. Se ha encontrado que una buena impresión auditiva se puede obtener si las señales de canales de audio que se asocian con posiciones vecinas verticalmente de una escena de audio se codifican conjuntamente en la primera
10 etapa del codificador jerárquico, si bien es mejor derivar los grupos de parámetros de extensión de ancho de banda comunes de señales de canales de audio que no están asociadas con posiciones vecinas verticalmente (pero que se asocian con diferentes posiciones horizontales o diferentes posiciones azimutales).

[0040] En una realización preferente, la primera señal del canal de audio y la tercera señal del canal de audio
15 se asocian con un primer plano horizontal común (o una primera elevación común) de una escena de audio pero diferentes posiciones horizontales (o posiciones azimutales) de la escena de audio y la segunda señal del canal de audio y la cuarta señal del canal de audio se asocian con un segundo plano horizontal común (o una segunda elevación común) de la escena de audio pero diferentes posiciones horizontales (o posiciones azimutales) de la escena de audio, en la que el primer plano horizontal es diferente del segundo plano horizontal. Se ha encontrado que se pueden
20 conseguir unos resultados de codificación de audio particularmente buenos (y, en consecuencia, resultados de decodificación de audio) utilizando dicha asociación espacial de las señales de canales de audio.

[0041] En una realización preferente, la primera señal del canal de audio y la segunda señal del canal de audio se asocian con un primer plano vertical (o una primera posición azimutal) de la escena de audio pero diferentes
25 posiciones verticales (o diferentes elevaciones) de la escena de audio. Más aún, la tercera señal del canal de audio y la cuarta señal del canal de audio están asociadas, preferentemente, con un segundo plano vertical (o una segunda posición azimutal) de la escena de audio pero diferentes posiciones verticales (o diferentes elevaciones) de la escena de audio, en la que el primer plano vertical común es diferente del segundo plano vertical común. Se ha encontrado que dicha asociación espacial de las señales de canales de audio da como resultado una buena calidad de codificación
30 de audio.

[0042] En una realización preferente, la primera señal del canal de audio y la segunda señal del canal de audio se asocian con un lado izquierdo de la escena de audio y la tercera señal del canal de audio y la cuarta señal del canal de audio se asocian con un lado derecho de la escena de audio. En consecuencia, se puede conseguir una buena
35 impresión auditiva mientras que la decodificación es normalmente eficaz en la tasa de bits.

[0043] En una realización preferente, la primera señal del canal de audio y la tercera señal del canal de audio se asocian con una parte inferior de la escena de audio y la segunda señal del canal de audio y la cuarta señal del canal de audio se asocian con una parte superior de la escena de audio. Esta disposición también ayuda a obtener
40 una codificación de audio eficaz con una buena impresión auditiva.

[0044] En una realización preferente, el codificador de audio está configurado para realizar una combinación horizontal cuando se proporciona la representación codificada de las señales de mezcla descendente sobre la base de la primera señal de mezcla descendente y la segunda señal de mezcla descendente utilizando una codificación
45 multicanal. En paralelo con las explicaciones anteriores hechas con respecto al decodificador de audio, se ha encontrado que se puede obtener una impresión auditiva particularmente buena si la combinación horizontal se lleva a cabo en la segunda etapa del codificador de audio (cuando se compara con la primera etapa del codificador de audio), dado que la posición horizontal de un objeto de audio tiene una relevancia particularmente alta para un oyente y dado que la segunda etapa del codificador de audio jerárquico normalmente corresponde a la primera etapa del
50 decodificador de audio jerárquico descrita anteriormente.

[0045] En una realización preferente, el codificador de audio está configurado para realizar una combinación vertical cuando se proporciona la primera señal de mezcla descendente sobre la base de la primera señal del canal de audio y la segunda señal del canal de audio utilizando una decodificación multicanal. Más aún, el decodificador de
55 audio está configurado preferentemente para realizar una combinación vertical cuando se proporciona la segunda señal de mezcla descendente sobre la base de la tercera señal del canal de audio y la cuarta señal del canal de audio. Conforme a ello, se lleva a cabo una combinación vertical en la primera etapa del codificador de audio. Esto es ventajoso dado que la posición vertical de un objeto de audio normalmente no es tan importante para el oyente como la posición horizontal del objeto de audio, de manera que las degradaciones de la reproducción, que son causadas

por la codificación jerárquica (y, en consecuencia, decodificación jerárquica), se pueden mantener razonablemente bajas.

5 **[0046]** En una realización preferente, el codificador de audio está configurado para proporcionar la representación codificada conjuntamente de la primera señal de mezcla descendente y la segunda señal de mezcla descendente sobre la base de la primera señal de mezcla descendente y la segunda señal de mezcla descendente utilizando una codificación multicanal basada en predicciones. Se ha encontrado que dicha codificación multicanal basada en predicciones es muy adecuada para la codificación conjunta que se lleva a cabo en la segunda etapa del codificador jerárquico. Se hace referencia a las explicaciones anteriores respecto del decodificador de audio, que
10 también se aplican en esta invención de una manera paralela.

[0047] En una realización preferente, se proporciona un parámetro de predicción que describe una contribución del componente de señal, que se ha derivado utilizando un componente de señal de una trama anterior, al suministro de la señal de mezcla descendente de la trama actual que utiliza la codificación multicanal basada en predicciones.
15 Conforme a ello, una buena reconstrucción de señales se puede conseguir en este lado del codificador de audio, que aplica este parámetro de predicción que describe una contribución del componente de señal, que se obtiene utilizando un componente de señal de una trama anterior, al suministro de la señal de mezcla descendente de la trama actual.

[0048] En una realización preferente, la codificación multicanal basada en predicciones funciona en el dominio MDCT. Conforme a ello, la codificación multicanal basada en predicciones se adapta bien a la codificación final de una señal de salida de la codificación multicanal basada en predicciones (por ejemplo, de una señal común de mezcla descendente), en la que esta codificación final se lleva a cabo normalmente en el dominio MDCT para mantener los artefactos de bloqueo razonablemente pequeños.
20

25 **[0049]** En una realización preferente, la codificación multicanal basada en predicciones es una codificación de predicción estéreo compleja USAC. El uso de la codificación de predicción estéreo compleja USAC facilita la implementación dado que el hardware y/o código de programa existentes se pueden reutilizar con facilidad para implementar el codificador de audio jerárquico.

30 **[0050]** En una realización preferente, el codificador de audio está configurado para proporcionar una representación codificada conjuntamente de la primera señal de mezcla descendente y la segunda señal de mezcla descendente sobre la base de la primera señal de mezcla descendente y la segunda señal de mezcla descendente utilizando una codificación multicanal asistida con señales residuales. Conforme a ello, se puede conseguir una calidad de reproducción particularmente buena del lado de un decodificador de audio.
35

[0051] En una realización preferente, el codificador de audio está configurado para proporcionar la primera señal de mezcla descendente sobre la base de la primera señal del canal de audio y la segunda señal del canal de audio utilizando una codificación multicanal basada en parámetros. Más aún, el codificador de audio está configurado para conducir la segunda señal de mezcla descendente sobre la base de la tercera señal del canal de audio y la cuarta
40 señal del canal de audio utilizando una codificación multicanal basada en parámetros. Se ha encontrado que el uso de una codificación multicanal basada en parámetros proporciona un buen compromiso entre la calidad de reproducción y la tasa de bits cuando se aplica en la primera etapa del codificador de audio jerárquico.

[0052] En una realización preferente, la codificación multicanal basada en parámetros está configurado para
45 proporcionar uno o varios parámetros que describen una correlación deseada entre dos canales y/o diferencias de niveles entre dos canales. Conforme a ello, es posible una codificación eficaz con una tasa de bits moderada sin degradar significativamente la calidad de audio.

50 **[0053]** En una realización preferente, la codificación multicanal a base de parámetros funciona en el dominio QMF, que se adapta bien a un preprocesamiento que se puede realizar en las señales de canales de audio.

[0054] En una realización preferente, la codificación multicanal a base de parámetros es una codificación Surround MPEG 2-1-2 o una codificación estéreo unificada. El uso de dichos conceptos de codificación puede reducir
55 significativamente el esfuerzo de la implementación.

[0055] En una realización preferente, el codificador de audio está configurado para proporcionar la primera señal de mezcla descendente sobre la base de la primera señal del canal de audio y la segunda señal del canal de audio utilizando una codificación multicanal asistida con señales residuales. Más aún, el codificador de audio se puede

configurar para proporcionar la segunda señal de mezcla descendente sobre la base de la tercera señal del canal de audio y la cuarta señal del canal de audio utilizando una codificación multicanal asistida con señales residuales. Conforme a ello, es posible para obtener una calidad de audio incluso mejor.

5 **[0056]** En una realización preferente, el codificador de audio está configurado para proporcionar una representación codificada conjuntamente de una primera señal residual, que se obtiene cuando se codifica conjuntamente al menos la primera señal del canal de audio y la segunda señal del canal de audio y de una segunda señal residual, que se obtiene cuando se codifica conjuntamente al menos la tercera señal del canal de audio y la cuarta señal del canal de audio, utilizando una codificación multicanal. Se ha encontrado que el concepto de
10 codificación jerárquica se puede incluso aplicar a las señales residuales, que se proporcionan en la primera etapa de la codificación de audio jerárquica. Con una codificación conjunta de las señales residuales, se pueden aprovechar las dependencias (o correlaciones) entre las señales de canales de audio porque estas dependencias (o correlaciones) también se reflejan normalmente en las señales residuales.

15 **[0057]** En una realización preferente, la primera señal residual y la segunda señal residual se asocian con diferentes posiciones horizontales (o posiciones azimutales) de una escena de audio. Conforme a ello, las dependencias entre las señales residuales se pueden codificar con buena precisión en la segunda etapa de la codificación jerárquica. Esto permite una reproducción de las dependencias (o correlaciones) entre las diferentes posiciones horizontales (o posiciones azimutales) con una buena impresión auditiva del lado de un decodificador de
20 audio.

[0058] En una realización preferente, la primera señal residual está asociada con un lado izquierdo de una escena de audio y la segunda señal residual está asociada con un lado derecho de la escena de audio. Conforme a ello, la codificación conjunta de la primera señal residual y de la segunda señal residual que se asocian con diferentes
25 posiciones horizontales (o posiciones azimutales) de la escena de audio, se lleva a cabo en la segunda etapa del codificador de audio que permite una reproducción de alta calidad del lado del decodificador de audio.

[0059] Una realización preferente según la invención crea un procedimiento que proporciona al menos cuatro señales de canales de audio sobre la base de una representación codificada. El procedimiento comprende
30 proporcionar una primera señal de mezcla descendente y una segunda señal de mezcla descendente sobre la base de una representación codificada conjuntamente de la primera señal de mezcla descendente y la segunda señal de mezcla descendente utilizando una (primera) decodificación multicanal. El procedimiento también comprende proporcionar al menos una primera señal del canal de audio y una segunda señal del canal de audio sobre la base de la primera señal de mezcla descendente utilizando una (segunda) decodificación multicanal y proporcionar al menos
35 una tercera señal de canal de audio y una cuarta señal de canal de audio sobre la base de la segunda señal de mezcla descendente utilizando una (tercera) decodificación multicanal. El procedimiento también comprende realizar una (primera) extensión de ancho de banda multicanal sobre la base de la primera señal del canal de audio y la tercera señal del canal de audio con el fin de obtener una primera señal del canal con ancho de banda extendido y una tercera señal del canal con ancho de banda extendido. El procedimiento también comprende realizar una (segunda) extensión
40 de ancho de banda multicanal sobre la base de la segunda señal del canal de audio y la cuarta señal del canal de audio con el fin de obtener la segunda señal del canal con ancho de banda extendido. Este procedimiento está basado en las mismas consideraciones que el decodificador de audio descrito anteriormente.

[0060] Una realización preferente según la invención crea un procedimiento que proporciona una representación codificada sobre la base de al menos cuatro señales de canales de audio. El procedimiento comprende
45 obtener un primer grupo de parámetros de extensión de ancho de banda comunes sobre la base de una primera señal del canal de audio y una tercera señal de canal de audio. El procedimiento también comprende obtener un segundo grupo de parámetros de extensión de ancho de banda comunes sobre la base de una segunda señal del canal de audio y una cuarta señal de canal de audio. El procedimiento también comprende la codificación conjunta de al menos
50 la primera señal del canal de audio y la segunda señal del canal de audio utilizando una codificación multicanal con el fin de obtener una primera señal de mezcla descendente y la codificación conjunta de al menos la tercera señal del canal de audio y la cuarta señal del canal de audio utilizando una codificación multicanal para obtener una segunda señal de mezcla descendente. El procedimiento también comprende la codificación conjunta de la primera señal de mezcla descendente y la segunda señal de mezcla descendente que utilizan una codificación multicanal con el fin de
55 obtener una representación codificada de las señales de mezcla descendente. Este procedimiento está basado en las mismas consideraciones que el codificador de audio descrito anteriormente.

[0061] Otras realizaciones según la invención crea programas informáticos que llevan a cabo los procedimientos mencionados en esta invención.

BREVE DESCRIPCIÓN DE LAS FIGURAS

[0062] Las realizaciones según la presente invención se describirán posteriormente haciendo referencia a las
5 figures adjuntas, en las que:

la Fig. 1 muestra un diagrama esquemático en bloques de un codificador de audio según una realización de la presente invención;

10 la Fig. 2 muestra un diagrama esquemático en bloques de un decodificador de audio según una realización de la presente invención;

la Fig. 3 muestra un diagrama esquemático en bloques de un decodificador de audio según otra realización de la presente invención;

15 la Fig. 4 muestra un diagrama esquemático en bloques de un codificador de audio según una realización de la presente invención;

20 la Fig. 5 muestra un diagrama esquemático en bloques de un decodificador de audio según una realización de la presente invención;

la Fig. 6 muestra un diagrama esquemático en bloques de un decodificador de audio según otra realización de la presente invención;

25 la Fig. 7 muestra un diagrama de flujo de un procedimiento que proporciona una representación codificada sobre la base de al menos cuatro señales de canales de audio según una realización de la presente invención;

la Fig. 8 muestra un diagrama de flujo de un procedimiento que proporciona al menos cuatro señales de canales de audio sobre la base de una representación codificada según una realización de la invención;

30 la Fig. 9 muestra un diagrama de flujo de un procedimiento que proporciona una representación codificada sobre la base de al menos cuatro señales de canales de audio según una realización de la invención; y

35 la Fig. 10 muestra un diagrama de flujo de un procedimiento que proporciona al menos cuatro señales de canales de audio sobre la base de una representación codificada según una realización de la invención;

la Fig. 11 muestra un diagrama esquemático en bloques de un codificador de audio según una realización de la invención;

40 la Fig. 12 muestra un diagrama esquemático en bloques de un codificador de audio, según otra realización de la invención;

la Fig. 13 muestra un diagrama esquemático en bloques de un decodificador de audio según una realización de la invención;

45 la Fig. 14a muestra una representación de sintaxis de una corriente de bits que se puede utilizar con el codificador de audio según la Fig. 13.

la Fig. 14b muestra una representación en tabla de diferentes valores del parámetro qcelIndex;

50 la Fig. 15 muestra un diagrama esquemático en bloques de un codificador de audio 3D en el que se pueden utilizar los conceptos según la presente invención;

la Fig. 16 muestra un diagrama esquemático en bloques de un decodificador de audio 3D, en el que se pueden utilizar
55 los conceptos según la presente invención; y

la Fig. 17 muestra un diagrama esquemático en bloques de un convertidor de formato.

la Fig. 18 muestra una representación gráfica de una estructura topológica de un elemento de canal Quad (QCE)

según una realización de la presente invención;

la Fig. 19 muestra un diagrama esquemático en bloques de un decodificador de audio según una realización de la presente invención;

5

la Fig. 20 muestra un diagrama esquemático en bloques detallado de un decodificador QCE según una realización de la presente invención; y

la Fig. 21 muestra un diagrama esquemático en bloques detallado de un codificador de canal Quad según una realización de la presente invención.

10

DESCRIPCIÓN DETALLADA DE LAS REALIZACIONES

1. Codificador de audio según la Fig. 1

15

[0063] La Fig. 1 muestra un diagrama esquemático en bloques de un codificador de audio, que se indica en su totalidad con 100. El codificador de audio 100 está configurado para proporcionar una representación codificada sobre la base de al menos cuatro señales de canales de audio. El codificador de audio 100 está configurado para recibir una primera señal del canal de audio 110, una segunda señal del canal de audio 112, una tercera señal de canal de audio 114 y una cuarta señal de canal de audio 116. Más aún, el codificador de audio 100 está configurado para proporcionar una representación codificada de una primera señal de mezcla descendente 120 y de una segunda señal de mezcla descendente 122, así como una representación codificada conjuntamente 130 de señales residuales. El codificador de audio 100 comprende un codificador multicanal asistido con señales residuales 140, que está configurado para codificar conjuntamente la primera señal del canal de audio 110 y la segunda señal del canal de audio 112 utilizando una codificación multicanal asistida con señales residuales con el fin de obtener la primera señal de mezcla descendente 120 y una primera señal residual 142. El codificador de señales de audio 100 también comprende un codificador multicanal asistido con señales residuales 150, que está configurado para codificar conjuntamente al menos la tercera señal del canal de audio 114 y la cuarta señal del canal de audio 116 que utiliza una codificación multicanal asistida con señales residuales con el fin de obtener la segunda señal de mezcla descendente 122 y una segunda señal residual 152. El decodificador de audio 100 también comprende un codificador multicanal 160, que está configurado para codificar conjuntamente la primera señal residual 142 y la segunda señal residual 152 utilizando una codificación multicanal con el fin de obtener la representación codificada conjuntamente 130 de las señales residuales 142, 152.

20

25

30

35

[0064] Respecto de la funcionalidad del codificador de audio 100, se debería observar que el codificador de audio 100 realiza una codificación jerárquica, en el que la primera señal del canal de audio 110 y la segunda señal del canal de audio 112 se codifican conjuntamente utilizando la codificación multicanal asistida con señales residuales 140, en la que se proporcionan tanto la primera señal de mezcla descendente 120 como la primera señal residual 142. La primera señal residual 142 puede describir, por ejemplo, diferencias entre la primera señal del canal de audio 110 y la segunda señal del canal de audio 112, y/o puede describir algunas o cualquier característica de señal que no se puede representar por la primera señal de mezcla descendente 120 y parámetros opcionales, que se pueden proporcionar mediante el codificador multicanal asistido con señales residuales 140. En otras palabras, la primera señal residual 142 puede ser una señal residual que permite un refinamiento de un resultado de decodificación que se puede obtener sobre la base de la primera señal de mezcla descendente 120 y cualquier parámetro posible que puede ser proporcionado mediante el codificador multicanal asistido con señales residuales 140. Por ejemplo, la primera señal residual 142 puede permitir al menos una reconstrucción de forma de onda parcial de la primera señal del canal de audio 110 y de la segunda señal del canal de audio 112 del lado de un decodificador de audio cuando se compara con una mera reconstrucción de características de señal de alto nivel (como, por ejemplo, características de correlación, características de covarianza, características de diferencia de nivel, y similares). De modo similar, el codificador multicanal asistido con señales residuales 150 proporciona tanto la segunda señal de mezcla descendente 122 como la segunda señal residual 152 sobre la base de la tercera señal del canal de audio 114 y la cuarta señal del canal de audio 116, de manera que la segunda señal residual permita un refinamiento de una reconstrucción de señales de la tercera señal del canal de audio 114 y de la cuarta señal del canal de audio 116 del lado de un decodificador de audio. La segunda señal residual 152 puede servir, en consecuencia, para la misma funcionalidad que la primera señal residual 142. Sin embargo, si las señales de canales de audio 110, 112, 114, 116 comprenden la misma correlación, la primera señal residual 142 y la segunda señal residual 152 normalmente también se correlacionan hasta cierto grado. Conforme a ello, la codificación conjunta de la primera señal residual 142 y de la segunda señal residual 152 que utiliza el codificador multicanal 160 normalmente comprende una alta eficacia dado que una codificación multicanal de señales correlacionadas normalmente reduce la tasa de bits explotando las

40

45

50

55

dependencias. En consecuencia, la primera señal residual 142 y la segunda señal residual 152 se pueden codificar con buena precisión mientras se mantiene la tasa de bits de la representación codificada conjuntamente 130 de las señales residuales razonablemente pequeña.

5 **[0065]** Para resumir, la realización según la Fig. 1 proporciona una codificación multicanal jerárquica, en la que una buena calidad de reproducción se puede conseguir con los codificadores multicanal asistidos con señales residuales 140, 150 y en la que la demanda de una tasa de bits se puede mantener moderada por codificación conjunta de una primera señal residual 142 y una segunda señal residual 152.

10 **[0066]** Otra mejora opcional del codificador de audio 100 es posible. Algunas de estas mejoras se describirán haciendo referencia a las Figs. 4, 11 y 12. Sin embargo, se debería observar que el codificador de audio 100 también se puede adaptar en paralelo con los decodificadores de audio descritos en esta invención, en el que la funcionalidad del codificador de audio normalmente es inverso a la funcionalidad del decodificador de audio.

15 2. Decodificador de audio según la Fig. 2

[0067] La Fig. 2 muestra un diagrama esquemático en bloques de un decodificador de audio, que se indica en su totalidad con 200.

20 **[0068]** El decodificador de audio 200 está configurado para recibir una representación codificada que comprende una representación codificada conjuntamente 210 de una primera señal residual y una segunda señal residual. El decodificador de audio 200 también recibe una representación de una primera señal de mezcla descendente 212 y de una segunda señal de mezcla descendente 214. El decodificador de audio 200 está configurado para proporcionar una primera señal del canal de audio 220, una segunda señal del canal de audio 222, una tercera
25 señal de canal de audio 224 y una cuarta señal de canal de audio 226.

[0069] El decodificador de audio 200 comprende un decodificador multicanal 230, que está configurado para proporcionar una primera señal residual 232 y una segunda señal residual 234 sobre la base de la representación codificada conjuntamente 210 de la primera señal residual 232 y de la segunda señal residual 234. El decodificador
30 de audio 200 también comprende un (primer) decodificador multicanal asistido con señales residuales 240 que está configurado para proporcionar la primera señal del canal de audio 220 y la segunda señal del canal de audio 222 sobre la base de la primera señal de mezcla descendente 212 y la primera señal residual 232 utilizando una decodificación multicanal. El decodificador de audio 200 también comprende un (segundo) decodificador multicanal asistido con
35 señales residuales 250, que está configurado para proporcionar la tercera señal del canal de audio 224 y la cuarta señal del canal de audio 226 sobre la base de la segunda señal de mezcla descendente 214 y la segunda señal residual 234.

[0070] Respecto de la funcionalidad del decodificador de audio 200, se debería observar que el decodificador de señales de audio 200 proporciona la primera señal del canal de audio 220 y la segunda señal del canal de audio
40 222 sobre la base de una (primera) decodificación multicanal asistida con señales residuales comunes 240, en el que la calidad de decodificación de la decodificación multicanal es incrementada por la primera señal residual 232 (cuando se compara con una decodificación asistida por señales no residuales). En otras palabras, la primera señal de mezcla descendente 212 proporciona una información "gruesa" acerca de la primera señal del canal de audio 220 y la segunda
45 señal del canal de audio 222, en la que, por ejemplo, diferencias entre la primera señal del canal de audio 220 y la segunda señal del canal de audio 222 se pueden describir por parámetros (opcionales), que pueden ser recibidos por el decodificador multicanal asistido con señales residuales 240 y por la primera señal residual 232. En consecuencia, la primera señal residual 232 puede permitir, por ejemplo, una reconstrucción de forma de onda parcial de la primera
señal del canal de audio 220 y de la segunda señal del canal de audio 222.

50 **[0071]** De modo similar, el (segundo) decodificador multicanal asistido con señales residuales 250 proporciona la tercera señal del canal de audio 224 en la cuarta señal del canal de audio 226 sobre la base de la segunda señal de mezcla descendente 214, en el que la segunda señal de mezcla descendente 214 puede describir, por ejemplo, "en forma gruesa" la tercera señal del canal de audio 224 y la cuarta señal del canal de audio 226. Más aún, se pueden
describir diferencias entre la tercera señal del canal de audio 224 y la cuarta señal del canal de audio 226, por ejemplo,
55 por parámetros (opcionales), que pueden ser recibidos por el (segundo) decodificador multicanal asistido con señales residuales 250 y por la segunda señal residual 234. Conforme a ello, la evaluación de la segunda señal residual 234 puede permitir, por ejemplo, una reconstrucción de forma de onda parcial de la tercera señal del canal de audio 224 y la cuarta señal del canal de audio 226. Conforme a ello, la segunda señal residual 234 puede permitir una mejora de la calidad de reconstrucción de la tercera señal del canal de audio 224 y la cuarta señal del canal de audio 226.

[0072] Sin embargo, la primera señal residual 232 y la segunda señal residual 234 se obtienen de una representación codificada conjuntamente 210 de la primera señal residual y de la segunda señal residual. Dicha decodificación multicanal, que se lleva a cabo por el decodificador multicanal 230, permite una alta eficacia de decodificación dado que la primera señal del canal de audio 220, la segunda señal del canal de audio 222, la tercera señal del canal de audio 224 y la cuarta señal del canal de audio 226 son normalmente similares o "correlacionadas". Conforme a ello, la primera señal residual 232 y la segunda señal residual 234 también son normalmente similares o "correlacionadas", que se pueden aprovechar derivando la primera señal residual 232 y la segunda señal residual 234 a partir de una representación codificada conjuntamente 210 que utiliza una decodificación multicanal.

[0073] En consecuencia, es posible obtener una alta calidad de decodificación con una tasa de bits moderada mediante la decodificación de las señales residuales 232, 234 sobre la base de una representación codificada conjuntamente 210 de las mismas y utilizando cada una de las señales residuales para la decodificación de dos o más señales de canales de audio.

[0074] Para concluir, el decodificador de audio 200 permite una elevada eficacia de codificación al proporcionar señales de canales de audio de alta calidad 220, 222, 224, 226.

[0075] Se deberá observar que las características y funcionalidades adicionales, que se pueden implementar opcionalmente en el decodificador de audio 200, se describirán posteriormente haciendo referencia a las Figs. 3, 5, 6 y 13. Sin embargo, se debería observar que el codificador de audio 200 puede comprender las ventajas antes mencionadas sin ninguna modificación adicional.

3 Decodificador de audio según la Fig. 3

[0076] La Fig. 3 muestra un diagrama esquemático en bloques de un decodificador de audio según otra realización de la presente invención. El decodificador de audio de Fig. 3 se ha indicado en su totalidad con 300. El decodificador de audio 300 es similar al decodificador de audio 200 según la Fig. 2, de manera que también se aplican las explicaciones anteriores. Sin embargo, el decodificador de audio 300 se complementa con características y funcionalidades adicionales cuando se compara con el decodificador de audio 200 como se explicará a continuación.

[0077] El decodificador de audio 300 está configurado para recibir una representación codificada conjuntamente 310 de una primera señal residual y de una segunda señal residual. Más aún, el decodificador de audio 300 está configurado para recibir una representación codificada conjuntamente 360 de una primera señal de mezcla descendente y de una segunda señal de mezcla descendente. Más aún, el decodificador de audio 300 está configurado para proporcionar una primera señal del canal de audio 320, una segunda señal del canal de audio 322, una tercera señal de canal de audio 324 y una cuarta señal de canal de audio 326. El decodificador de audio 300 comprende un decodificador multicanal 330 que está configurado para recibir la representación codificada conjuntamente 310 de la primera señal residual y de la segunda señal residual y para proporcionar, sobre la base de las mismas, una primera señal residual 332 y una segunda señal residual 334. El decodificador de audio 300 también comprende una (primera) decodificación multicanal asistida con señales residuales 340, que recibe la primera señal residual 332 y una primera señal de mezcla descendente 312 y proporciona la primera señal del canal de audio 320 y la segunda señal del canal de audio 322. El decodificador de audio 300 también comprende una (segunda) decodificación multicanal asistida con señales residuales 350, que está configurado para recibir la segunda señal residual 334 y una segunda señal de mezcla descendente 314 y para proporcionar la tercera señal del canal de audio 324 y la cuarta señal del canal de audio 326.

[0078] El decodificador de audio 300 también comprende otro decodificador multicanal 370, que está configurado para recibir la representación codificada conjuntamente 360 de la primera señal de mezcla descendente y de la segunda señal de mezcla descendente y para proporcionar, sobre la base de las mismas, la primera señal de mezcla descendente 312 y la segunda señal de mezcla descendente 314.

[0079] A continuación, se describirán otros detalles específicos del decodificador de audio 300. Sin embargo, se debería observar que un decodificador de audio real no necesita implementar una combinación de todas estas características y funcionalidades adicionales. Más bien, las características y funcionalidades descritas a continuación se pueden añadir individualmente al decodificador de audio 200 (o cualquier otro decodificador de audio), para mejorar de forma gradual el decodificador de audio 200 (o cualquier otro decodificador de audio).

[0080] En una realización preferente, el decodificador de audio 300 recibe una representación codificada

conjuntamente 310 de la primera señal residual y la segunda señal residual, en el que esta representación codificada conjuntamente 310 puede comprender una señal de mezcla descendente de la primera señal residual 332 y de la segunda señal residual 334 y una señal residual común de la primera señal residual 332 y la segunda señal residual 334. Además, la representación codificada conjuntamente 310 puede comprender, por ejemplo, uno o varios parámetros de predicción. Conforme a ello, el decodificador multicanal 330 puede ser un decodificador multicanal asistido con señales residuales basada en predicciones. Por ejemplo, el decodificador multicanal 330 puede ser una predicción estéreo compleja USAC, tal como se describe, por ejemplo, en la sección “predicción estéreo compleja” del estándar internacional ISO/IEC 23003–3:2012. Por ejemplo, el decodificador multicanal 330 se puede configurar para evaluar un parámetro de predicción que describe una contribución de un componente de señal, que se obtiene utilizando un componente de señal de una trama anterior, a un suministro de la primera señal residual 332 y la segunda señal residual 334 de una trama actual. Más aún, el decodificador multicanal 330 se puede configurar para aplicar la señal residual común (que se incluye en la representación codificada conjuntamente 310) con un primer signo para obtener la primera señal residual 332 y para aplicar la señal residual común (que se incluye en la representación codificada conjuntamente 310) con un segundo signo, que es opuesto al primer signo con el fin de obtener la segunda señal residual 334. De esta manera, la señal residual común puede describir, al menos en parte, las diferencias entre la primera señal residual 332 y la segunda señal residual 334. Sin embargo, el decodificador multicanal 330 puede evaluar la señal de mezcla descendente, la señal residual común y los uno o varios parámetros de predicción, que están todos incluidos en la representación codificada conjuntamente 310 con el fin de obtener la primera señal residual 332 y la segunda señal residual 334 tal como se describe en el estándar internacional ISO/IEC 23003–3:2012 antes mencionado. Más aún, se debería observar que la primera señal residual 332 puede estar asociada con una primera posición horizontal (o posición azimutal), por ejemplo, una posición horizontal izquierda y que la segunda señal residual 334 puede estar asociada con una segunda posición horizontal (o posición azimutal), por ejemplo, una posición horizontal derecha, de una escena de audio.

[0081] La representación codificada conjuntamente 360 de la primera señal de mezcla descendente y de la segunda señal de mezcla descendente comprende, preferentemente, una señal de mezcla descendente de la primera señal de mezcla descendente y de la segunda señal de mezcla descendente, una señal residual común de la primera señal de mezcla descendente y de la segunda señal de mezcla descendente y uno o varios parámetros de predicción. En otras palabras, hay una señal de mezcla descendente “común”, en la que la primera señal de mezcla descendente 312 y la segunda señal de mezcla descendente 314 se mezclan de forma descendente y hay una señal residual “común” que pueden describir, al menos en parte, diferencias entre la primera señal de mezcla descendente 312 y la segunda señal de mezcla descendente 314. El decodificador multicanal 370 es, preferentemente, un decodificador multicanal asistido con señales residuales basada en predicciones, por ejemplo, un decodificador USAC de predicción estéreo compleja. En otras palabras, el decodificador multicanal 370, que proporciona la primera señal de mezcla descendente 312 y la segunda señal de mezcla descendente 314 puede ser sustancialmente idéntico al decodificador multicanal 330, que proporciona la primera señal residual 332 y la segunda señal residual 334, de manera que también se pueden aplicar las explicaciones y referencias anteriores. Más aún, se debería observar que la primera señal de mezcla descendente 312 está asociada preferentemente con una primera posición horizontal o posición azimutal (por ejemplo, posición horizontal izquierda o posición azimutal) de la escena de audio y que la segunda señal de mezcla descendente 314 está asociada preferentemente con una segunda posición horizontal o posición azimutal (por ejemplo, posición horizontal derecha o posición azimutal) de la escena de audio. Conforme a ello, la primera señal de mezcla descendente 312 y la primera señal residual 332 pueden estar asociadas con la misma primera posición horizontal o posición azimutal (por ejemplo, posición horizontal izquierda) y la segunda señal de mezcla descendente 314 y la segunda señal residual 334 puede estar asociada con la misma segunda posición horizontal o posición azimutal (por ejemplo, posición horizontal derecha). Conforme a ello, tanto el decodificador multicanal 370 y el decodificador multicanal 330 pueden realizar una división horizontal (o separación horizontal o distribución horizontal).

[0082] El decodificador multicanal asistido con señales residuales 340 puede estar, preferentemente, basado en parámetros y puede recibir, en consecuencia, uno o varios parámetros 342 que describen una correlación deseada entre dos canales (por ejemplo, entre la primera señal del canal de audio 320 y la segunda señal del canal de audio 322) y/o diferencias de niveles entre dichos dos canales. Por ejemplo, la decodificación multicanal asistida con señales residuales 340 se puede basar en una codificación MPEG–Surround (tal como se describe, por ejemplo, en ISO/IEC 23003–1:2007) con una extensión de señal residual o un decodificador de “decodificación estéreo unificada” (tal como se describe, por ejemplo, en ISO/IEC 23003–3, capítulo 7.11 (Decoder) & Anexo B.21 (Description of the Encoder & Definition of the Term “Unified Stereo”). Conforme a ello, el decodificador multicanal asistido con señales residuales 340 puede proporcionar la primera señal del canal de audio 320 y la segunda señal del canal de audio 322, en el que la primera señal del canal de audio 320 y la segunda señal del canal de audio 322 se asocian con posiciones vecinas verticalmente de la escena de audio. Por ejemplo, la primera señal del canal de audio puede estar asociada con una posición izquierda inferior de la escena de audio y la segunda señal del canal de audio puede estar asociada con una

posición izquierda superior de la escena de audio (de manera que la primera señal del canal de audio 320 y la segunda señal del canal de audio 322, por ejemplo, estén asociadas con idénticas posiciones horizontales o posiciones azimutales de la escena de audio o con posiciones azimutales separadas por no más de 30 grados). En otras palabras, el decodificador multicanal asistido con señales residuales 340 puede realizar una división vertical (o distribución o separación).

[0083] La funcionalidad del decodificador multicanal asistido con señales residuales 350 puede ser idéntica a la funcionalidad del decodificador multicanal asistido con señales residuales 340, en la que la tercera señal del canal de audio puede estar asociada, por ejemplo, con una posición derecha inferior de la escena de audio y en la que la cuarta señal del canal de audio puede estar asociada, por ejemplo, con una posición derecha superior de la escena de audio. En otras palabras, la tercera señal del canal de audio y la cuarta señal del canal de audio puede estar asociada con posiciones vecinas verticalmente de la escena de audio y puede estar asociada con la misma posición horizontal o posición azimutal de la escena de audio, en la que el decodificador multicanal asistido con señales residuales 350 realiza una división vertical (o separación o distribución).

[0084] Para resumir, el decodificador de audio 300 según la Fig. 3 realiza una decodificación de audio jerárquica, en el que una división izquierda–derecha se lleva a cabo en las primeras etapas (decodificador multicanal 330, decodificador multicanal 370) y en el que una división superior–inferior se lleva a cabo en la segunda etapa (decodificadores multicanal asistidos con señales residuales 340, 350). Más aún, las señales residuales 332, 334 también se codifican utilizando una representación codificada conjuntamente 310, así como las señales de mezcla descendente 312, 314 (representación codificada conjuntamente 360). De esta manera, las correlaciones entre los diferentes canales se aprovechan tanto para la codificación (and decodificación) de las señales de mezcla descendente 312, 314 como para la codificación (y decodificación) de las señales residuales 332, 334. Conforme a ello, se consigue una elevada eficacia de codificación y las correlaciones entre las señales se aprovechan bien.

4. Codificador de audio según la Fig. 4

[0085] La Fig. 4 muestra un diagrama esquemático en bloques de un codificador de audio según otra realización de la presente invención. El codificador de audio según la Fig. 4 se indica en su totalidad con 400. El codificador de audio 400 está configurado para recibir cuatro señales de canales de audio, concretamente, una primera señal del canal de audio 410, una segunda señal del canal de audio 412, una tercera señal de canal de audio 414 y una cuarta señal de canal de audio 416. Más aún, el codificador de audio 400 está configurado para proporcionar una representación codificada sobre la base de las señales de canales de audio 410, 412, 414 y 416, en el que dicha representación codificada comprende una representación codificada conjuntamente 420 de dos señales de mezcla descendente, así como una representación codificada de un primer grupo 422 de parámetros de extensión de ancho de banda comunes y de un segundo grupo 424 de parámetros de extensión de ancho de banda comunes. El codificador de audio 400 comprende un primer extractor de parámetros de extensión de ancho de banda 430, que está configurado para obtener el primer grupo 422 de parámetros de extracción de ancho de banda comunes sobre la base de la primera señal del canal de audio 410 y la tercera señal del canal de audio 414. El codificador de audio 400 también comprende un segundo extractor de parámetros de extensión de ancho de banda 440, que está configurado para obtener el segundo grupo 424 de parámetros de extensión de ancho de banda comunes sobre la base de la segunda señal del canal de audio 412 y la cuarta señal del canal de audio 416.

[0086] Más aún, el codificador de audio 400 comprende un (primer) codificador multicanal 450, que está configurado para codificar conjuntamente al menos la primera señal del canal de audio 410 y la segunda señal del canal de audio 412 que utiliza una codificación multicanal con el fin de obtener una primera señal de mezcla descendente 452. Además, el codificador de audio 400 también comprende un (segundo) codificador multicanal 460, que está configurado para codificar conjuntamente al menos la tercera señal del canal de audio 414 y la cuarta señal del canal de audio 416 que utiliza una codificación multicanal con el fin de obtener una segunda señal de mezcla descendente 462. Además, el codificador de audio 400 también comprende un (tercer) codificador multicanal 470, que está configurado para codificar conjuntamente la primera señal de mezcla descendente 452 y la segunda señal de mezcla descendente 462 que utiliza una codificación multicanal con el fin de obtener la representación codificada conjuntamente 420 de las señales de mezcla descendente.

[0087] Respecto de la funcionalidad del codificador de audio 400, se debería observar que el codificador de audio 400 realiza una codificación multicanal jerárquica, en el que la primera señal del canal de audio 410 y la segunda señal del canal de audio 412 se combinan en una primera etapa y en el que la tercera señal del canal de audio 414 y la cuarta señal del canal de audio 416 también se combinan en la primera etapa para obtener de este modo la primera señal de mezcla descendente 452 y la segunda señal de mezcla descendente 462. La primera señal de mezcla

descendente 452 y la segunda señal de mezcla descendente 462 se codifican luego conjuntamente en una segunda etapa. Sin embargo, se debería observar que el primer extractor de parámetros de extensión de ancho de banda 430 proporciona el primer grupo 422 de parámetros de extracción de ancho de banda comunes sobre la base de señales de canales de audio 410, 414 que están manejados por diferentes codificadores multicanal 450, 460 en la primera
 5 etapa de la codificación multicanal jerárquica. De modo similar, el segundo extractor de parámetros de extensión de ancho de banda 440 proporciona un segundo grupo 424 de parámetros de extracción de ancho de banda comunes sobre la base de diferentes señales de canales de audio 412, 416, que están manejados por diferentes codificadores multicanal 450, 460 en la primera etapa de procesamiento. Este orden específico de procesamiento trae consigo la
 10 ventaja de que los grupos 422, 424 de parámetros de extensión de ancho de banda se basan en canales que sólo se combinan en la segunda etapa de la codificación jerárquica (es decir, en el codificador multicanal 470). Esto es ventajoso dado que es deseable combinar dichos canales de audio en la primera etapa de la codificación jerárquica, cuya relación no es muy significativa con respecto a la percepción de una posición de la fuente de sonido. Más bien, es recomendable que la relación entre la primera señal de mezcla descendente y la segunda señal de mezcla descendente determine principalmente la percepción de una ubicación de la fuente de sonido, porque la relación entre
 15 la primera señal de mezcla descendente 452 y la segunda señal de mezcla descendente 462 se puede mantener mejor que la relación entre las señales individuales de canal de audio 410, 412, 414, 416. En otras palabras, se ha encontrado que es deseable que el primer grupo 422 de parámetros de extensión de ancho de banda comunes esté basado en dos audio canales (señales de canales de audio) que contribuyen con diferentes de las señales de mezcla descendente 452, 462 y que el segundo grupo 424 de parámetros de extensión de ancho de banda comunes se
 20 proporciona sobre la base de señales de canales de audio 412, 416, que también contribuyen con diferentes de las señales de mezcla descendente 452, 462, que se alcanza por el procesamiento de las señales de canales de audio antes descrito en la codificación multicanal jerárquica. En consecuencia, el primer grupo 422 de parámetros de extensión de ancho de banda comunes está basado en una relación de canales similar cuando se compara con la relación de canales entre la primera señal de mezcla descendente 452 y la segunda señal de mezcla descendente
 25 462, en las que la última, normalmente, domina la impresión espacial generada del lado de un decodificador de audio. Conforme a ello, el suministro del primer grupo 422 de parámetros de extensión de ancho de banda y también el suministro del segundo grupo 424 de parámetros de extensión de ancho de banda se adapta bien a una impresión acústica espacial que se genera del lado de un decodificador de audio.

30 5. Decodificador de audio según la Fig. 5

[0088] La Fig. 5 muestra un diagrama esquemático en bloques de un decodificador de audio, según otra realización de la presente invención. El decodificador de audio según la Fig. 5 se indica en su totalidad con 500.

35 **[0089]** El decodificador de audio 500 está configurado para recibir una representación codificada conjuntamente 510 de una primera señal de mezcla descendente y una segunda señal de mezcla descendente. Más aún, el decodificador de audio 500 está configurado para proporcionar una primera señal del canal con ancho de banda extendido 520, una segunda señal del canal con ancho de banda extendido 522, una tercera señal del canal con ancho de banda extendido 524 y una cuarta señal del canal con ancho de banda extendido 526.

40 **[0090]** El decodificador de audio 500 comprende un (primer) decodificador multicanal 530 que está configurado para proporcionar una primera señal de mezcla descendente 532 y una segunda señal de mezcla descendente 534 sobre la base de la representación codificada conjuntamente 510 de la primera señal de mezcla descendente y la segunda señal de mezcla descendente utilizando una decodificación multicanal. El decodificador de audio 500 también
 45 comprende un (segundo) decodificador multicanal 540, que está configurado para proporcionar al menos una primera señal del canal de audio 542 y una segunda señal del canal de audio 544 sobre la base de la primera señal de mezcla descendente 532 utilizando una decodificación multicanal. El decodificador de audio 500 también comprende un (tercer) decodificador multicanal 550 que está configurado para proporcionar al menos una tercera señal de canal de audio 556 y una cuarta señal de canal de audio 558 sobre la base de la segunda señal de mezcla descendente 544
 50 utilizando una decodificación multicanal. Más aún, el decodificador de audio 500 comprende una (primera) extensión de ancho de banda multicanal 560, que está configurado para realizar una extensión de ancho de banda multicanal sobre la base de la primera señal del canal de audio 542 y la tercera señal del canal de audio 556 con el fin de obtener una primera señal del canal con ancho de banda extendido 520 y la tercera señal del canal de ancho de banda extendida 524. Más aún, el decodificador de audio comprende una (segunda) extensión de ancho de banda multicanal
 55 570, que está configurado para realizar una extensión de ancho de banda multicanal sobre la base de la segunda señal del canal de audio 544 y la cuarta señal del canal de audio 558 con el fin de obtener la segunda señal del canal con ancho de banda extendido 522 y la cuarta señal del canal con ancho de banda extendido 526.

[0091] Respecto de la funcionalidad del decodificador de audio 500, se debería observar que el decodificador

de audio 500 realiza una decodificación multicanal jerárquica, en el que una división entre una primera señal de mezcla descendente 532 y una segunda señal de mezcla descendente 534 se lleva a cabo en una primera etapa de la decodificación jerárquica y en el que la primera señal del canal de audio 542 y la segunda señal del canal de audio 544 se obtienen de la primera señal de mezcla descendente 532 en una segunda etapa de la decodificación jerárquica y en el que la tercera señal del canal de audio 556 y la cuarta señal del canal de audio 558 se obtienen de la segunda señal de mezcla descendente 550 en la segunda etapa de la decodificación jerárquica. Sin embargo, tanto la primera extensión del ancho de banda multicanal 560 como la segunda extensión del ancho de banda multicanal 570 reciben cada una señal del canal de audio que se deriva de la primera señal de mezcla descendente 532 y una señal del canal de audio que se deriva de la segunda señal de mezcla descendente 534. Si bien una mejor separación del canal se consigue normalmente por la (primera) decodificación multicanal 530, que se lleva a cabo como una primera etapa de la decodificación multicanal jerárquica, cuando se compara con la segunda etapa de la decodificación jerárquica, se puede ver que cada extensión de ancho de banda multicanal 560, 570 recibe señales de entrada que se separan bien (porque se originan de la primera señal de mezcla descendente 532 y la segunda señal de mezcla descendente 534, que se separan bien por canales). De esta manera, la extensión del ancho de banda multicanal 560, 570 pueden tener en cuenta características estéreo, que son importantes para una impresión auditiva y que están bien representadas por la relación entre la primera señal de mezcla descendente 532 y la segunda señal de mezcla descendente 534 y, por ello, pueden proporcionar una buena impresión auditiva.

[0092] En otras palabras, la estructura "cruzada" del decodificador de audio, en el que cada una de las etapas de extensión del ancho de banda multicanal 560, 570 recibe señales de entrada de decodificadores multicanal de (segunda etapa) 540, 550 permite una buena extensión de ancho de banda multicanal, que considera una relación estéreo entre los canales.

[0093] Sin embargo, debe observarse que el decodificador de audio 500 se puede complementar con cualquiera de las características y funcionalidades descritas en este documento con respecto a los decodificadores de audio de acuerdo con las Figs. 2, 3, 6 y 13, en el cual es posible introducir características individuales en el decodificador de audio 500 que mejoran gradualmente el resultado del decodificador de audio.

6. Decodificador de audio según la Fig. 6

[0094] La Fig. 6 muestra un diagrama esquemático en bloques de un decodificador de audio según otra realización de la presente invención. El decodificador de audio según la Fig. 6 se indica en su totalidad con 600. El decodificador de audio 600 según la Fig. 6 es similar al decodificador de audio 500 según la Fig. 5 de manera que también se aplican las explicaciones anteriores. Sin embargo, el decodificador de audio 600 se ha suplementado por algunas características y funcionalidades, que también se pueden introducir, individualmente o en combinación, en el decodificador de audio 500 para mejora.

[0095] El decodificador de audio 600 está configurado para recibir una representación codificada conjuntamente 610 de una primera señal de mezcla descendente y de una segunda señal de mezcla descendente y para proporcionar una primera señal con ancho de banda extendido 620, una segunda señal con ancho de banda extendido 622, una tercera señal con ancho de banda extendido 624 y una cuarta señal con ancho de banda extendido 626. El decodificador de audio 600 comprende un decodificador multicanal 630, que está configurado para recibir la representación codificada conjuntamente 610 de la primera señal de mezcla descendente y de la segunda señal de mezcla descendente y para proporcionar, sobre la base de las mismas, la primera señal de mezcla descendente 632 y la segunda señal de mezcla descendente 634. El decodificador de audio 600 también comprende un decodificador multicanal 640, que está configurado para recibir la primera señal de mezcla descendente 632 y para proporcionar, sobre la base de la misma, una primera señal del canal de audio 542 y una segunda señal del canal de audio 544. El decodificador de audio 600 también comprende un decodificador multicanal 650, que está configurado para recibir la segunda señal de mezcla descendente 634 y para proporcionar una tercera señal de canal de audio 656 y una cuarta señal de canal de audio 658. El decodificador de audio 600 también comprende una (primera) extensión de ancho de banda multicanal 660, que está configurado para recibir la primera señal del canal de audio 642 y la tercera señal del canal de audio 656 y para proporcionar, sobre la base de las mismas, la primera señal del canal con ancho de banda extendido 620 y la tercera señal del canal con ancho de banda extendido 624. Además, una (segunda) extensión de ancho de banda multicanal 670 recibe la segunda señal del canal de audio 644 y la cuarta señal del canal de audio 658 y proporciona, sobre la base de las mismas, la segunda señal del canal con ancho de banda extendido 622 y la tercera señal del canal con ancho de banda extendido 626.

[0096] El decodificador de audio 600 también comprende otro decodificador multicanal 680, que está configurado para recibir una representación codificada conjuntamente 682 de una primera señal residual y de una

segunda señal residual y que proporciona, sobre la base de las mismas, una primera señal residual 684 para ser usada por el decodificador multicanal 640 y una segunda señal residual 686 para ser usada por el decodificador multicanal 650.

5 **[0097]** El decodificador multicanal 630 es, preferentemente, un decodificador multicanal asistido con señales residuales basada en predicciones. Por ejemplo, el decodificador multicanal 630 puede ser sustancialmente idéntico al decodificador multicanal 370 descrito anteriormente. Por ejemplo, el decodificador multicanal 630 puede ser un decodificador de predicción estéreo complejo USAC, como se mencionó anteriormente y tal como se describe en el estándar USAC mencionado anteriormente. Conforme a ello, la representación codificada conjuntamente 610 de la
10 primera señal de mezcla descendente y de la segunda señal de mezcla descendente pueden comprender, por ejemplo, una señal (común) de mezcla descendente de la primera señal de mezcla descendente y de la segunda señal de mezcla descendente, una señal residual (común) de la primera señal de mezcla descendente y de la segunda señal de mezcla descendente y uno o varios parámetros de predicción, que son evaluadas por el decodificador multicanal 630.

15 **[0098]** Más aún, se debería observar que la primera señal de mezcla descendente 632 puede estar asociada, por ejemplo, con una primera posición horizontal o posición azimutal (por ejemplo, a posición horizontal izquierda) de una escena de audio y que la segunda señal de mezcla descendente 634 puede estar asociada, por ejemplo, con una segunda posición horizontal o posición azimutal (por ejemplo, una posición horizontal derecha) de la escena de audio.
20

[0099] Más aún, el decodificador multicanal 680 puede ser, por ejemplo, un decodificador multicanal asociado con señales residuales basada en predicciones. El decodificador multicanal 680 puede ser sustancialmente idéntico al decodificador multicanal 330 descrito anteriormente. Por ejemplo, el decodificador multicanal 680 puede ser un decodificador USAC de predicción estéreo compleja, como se mencionó anteriormente. En consecuencia, la
25 representación codificada conjuntamente 682 de la primera señal residual y de la segunda señal residual puede comprender una señal de mezcla descendente (común) de la primera señal residual y de la segunda señal residual, una señal residual (común) de la primera señal residual y de la segunda señal residual y uno o varios parámetros de predicción, que son evaluadas por el decodificador multicanal 680. Más aún, se debería observar que la primera señal residual 684 puede estar asociada con una primera posición horizontal o posición azimutal (por ejemplo, una posición
30 horizontal izquierda) de la escena de audio y que la segunda señal residual 686 puede estar asociada con una segunda posición horizontal o posición azimutal (por ejemplo, una posición horizontal derecha) de la escena de audio.

[0100] El decodificador multicanal 640 puede ser, por ejemplo, una decodificación multicanal basada en parámetros como, por ejemplo, una decodificación multicanal Surround MPEG, tal como se ha descrito anteriormente y en el estándar de referencia. Sin embargo, en presencia del decodificador multicanal (opcional) 680 y la primera
35 señal residual (opcional) 684, el decodificador multicanal 640 puede ser un decodificador multicanal asistido con señales residuales basado en parámetros como, por ejemplo, un decodificador estéreo unificado. De esta manera, el decodificador multicanal 640 puede ser sustancialmente idéntico al decodificador multicanal 340 descrito anteriormente y el decodificador multicanal 640 puede recibir, por ejemplo, los parámetros 342 descritos anteriormente.
40

[0101] De modo similar, el decodificador multicanal 650 puede ser sustancialmente idéntico al decodificador multicanal 640. Conforme a ello, el decodificador multicanal 650 puede estar basado, por ejemplo, en parámetros y puede estar asistido opcionalmente por señales residuales (en presencia del decodificador multicanal opcional 680).

45 **[0102]** Más aún, se debería observar que la primera señal del canal de audio 642 y la segunda señal del canal de audio 644 están asociadas, preferentemente, con posiciones espaciales verticalmente adyacentes de la escena de audio. Por ejemplo, la primera señal del canal de audio 642 está asociada con una posición izquierda inferior de la escena de audio y la segunda señal del canal de audio 644 está asociada con una posición izquierda superior de la escena de audio. Conforme a ello, el decodificador multicanal 640 realiza una división vertical (o separación o
50 distribución) del contenido de audio descrito por la primera señal de mezcla descendente 632 (y, opcionalmente, por la primera señal residual 684). De modo similar, la tercera señal del canal de audio 656 y la cuarta señal del canal de audio 658 se asocian con posiciones verticalmente adyacentes de la escena de audio y están asociadas, preferentemente, con la misma posición horizontal o posición azimutal de la escena de audio. Por ejemplo, la tercera señal del canal de audio 656 está asociada preferentemente con una posición derecha inferior de la escena de audio
55 y la cuarta señal del canal de audio 658 está asociada preferentemente con una posición derecha superior de la escena de audio. De esta manera, el decodificador multicanal 650 realiza una división vertical (o separación o distribución) del contenido de audio descrito por la segunda señal de mezcla descendente 634 (y, opcionalmente, la segunda señal residual 686).

[0103] Sin embargo, la primera extensión del ancho de banda multicanal 660 recibe la primera señal del canal de audio 642 y el tercer canal de audio 656, que se asocian con la posición izquierda inferior y una posición derecha inferior de la escena de audio. Conforme a ello, la primera extensión del ancho de banda multicanal 660 realiza una extensión de ancho de banda multicanal sobre la base de dos señales de canales de audio que se asocian con el mismo plano horizontal (por ejemplo, plano horizontal inferior) o elevación de la escena de audio y diferentes lados (izquierdo/derecho) de la escena de audio. Conforme a ello, la extensión del ancho de banda multicanal puede tener en cuenta características estéreo (por ejemplo, la percepción estéreo humana) cuando se realiza la extensión del ancho de banda. De modo similar, la segunda extensión del ancho de banda multicanal 670 también puede tener en cuenta características estéreo, dado que la segunda extensión del ancho de banda multicanal trabaja con señales de canales de audio del mismo plano horizontal (por ejemplo, plano horizontal superior) o elevación pero en diferentes posiciones horizontales (diferentes lados) (izquierdo/derecho) de la escena de audio.

[0104] Para concluir además, el decodificador de audio jerárquico 600 comprende una estructura en la que una división izquierda/derecha (o separación o distribución) se lleva a cabo en una primera etapa (decodificación multicanal 630, 680), en la que una división vertical (separación o distribución) se lleva a cabo en una segunda etapa (decodificación multicanal 640, 650) y en la que la extensión del ancho de banda multicanal trabaja con un par de señales izquierda/derecha (extensión de ancho de banda multicanal 660, 670). Esta "cruza" de las vías de decodificación permite que se pueda realizar la separación izquierda/derecha, que es particularmente importante para la impresión auditiva (por ejemplo, más importante que la división superior/inferior) en la primera etapa de procesamiento del decodificador de audio jerárquico y que la extensión del ancho de banda multicanal también se puede realizar en un par de señales de canales de audio izquierda/derecha, que puede resultar otra vez en una impresión auditiva particularmente buena. La división superior/inferior se lleva a cabo como una etapa intermedia entre la separación izquierda-derecha y la extensión del ancho de banda multicanal, que permite derivar cuatro señales de canales de audio (o señales de canales con ancho de banda extendido) sin degradar de forma significativa la impresión auditiva.

7. Procedimiento según la Fig. 7

[0105] La Fig. 7 muestra un diagrama de flujo de un procedimiento 700 para proporcionar una representación codificada sobre la base de al menos cuatro señales de canales de audio.

[0106] El procedimiento 700 comprende la codificación conjunta 710 de al menos una primera señal del canal de audio y una segunda señal del canal de audio utilizando una codificación multicanal asistida con señales residuales con el fin de obtener una primera señal de mezcla descendente y una primera señal residual. El procedimiento también comprende la codificación conjunta 720 de al menos una tercera señal de canal de audio y una cuarta señal de canal de audio utilizando una codificación multicanal asistida con señales residuales con el fin de obtener una segunda señal de mezcla descendente y una segunda señal residual. El procedimiento también comprende la codificación conjunta 730 de la primera señal residual y la segunda señal residual utilizando una codificación multicanal con el fin de obtener una representación codificada de las señales residuales. Sin embargo, se debería observar que el procedimiento 700 puede ser suplementada por cualquiera de las características y funcionalidades descritas en esta invención con respecto a los codificadores de audio y decodificadores de audio.

8. Procedimiento según la Fig. 8

[0107] La Fig. 8 muestra un diagrama de flujo de un procedimiento 800 para proporcionar al menos cuatro señales de canales de audio sobre la base de una representación codificada.

[0108] El procedimiento 800 comprende el suministro 810 de una primera señal residual y una segunda señal residual sobre la base de una representación codificada conjuntamente de la primera señal residual y la segunda señal residual utilizando una decodificación multicanal. El procedimiento 800 también comprende el suministro 820 de una primera señal del canal de audio y una segunda señal del canal de audio sobre la base de una primera señal de mezcla descendente y la primera señal residual utilizando una decodificación multicanal asistida con señales residuales. El procedimiento también comprende el suministro 830 de una tercera señal de canal de audio y una cuarta señal de canal de audio sobre la base de una segunda señal de mezcla descendente y la segunda señal residual utilizando una decodificación multicanal asistida con señales residuales.

[0109] Más aún, se debería observar que el procedimiento 800 puede ser suplementado por cualquiera de las características y funcionalidades descritas en esta invención con respecto a los decodificadores de audio y codificadores de audio.

9. Procedimiento según la Fig. 9

[0110] La Fig. 9 muestra un diagrama de flujo de un procedimiento 900 para proporcionar una representación
5 codificada sobre la base de al menos cuatro señales del canal de audio.

[0111] El procedimiento 900 comprende la obtención 910 de un primer grupo de parámetros de extensión de
ancho de banda comunes sobre la base de una primera señal del canal de audio y una tercera señal de canal de
audio. El procedimiento 900 también comprende la obtención 920 de un segundo grupo de parámetros de extensión
10 de ancho de banda comunes sobre la base de una segunda señal del canal de audio y una cuarta señal de canal de
audio. El procedimiento también comprende la codificación conjunta de al menos la primera señal del canal de audio
y la segunda señal del canal de audio utilizando una codificación multicanal con el fin de obtener una primera señal
de mezcla descendente y la codificación conjunta 940 de al menos la tercera señal del canal de audio y la cuarta señal
del canal de audio utilizando una codificación multicanal para obtener una segunda señal de mezcla descendente. El
15 procedimiento también comprende la codificación conjunta 950 de la primera señal de mezcla descendente y la
segunda señal de mezcla descendente utilizando una codificación multicanal con el fin de obtener una representación
codificada de las señales de mezcla descendente.

[0112] Se deberá observar que algunas de las etapas del procedimiento 900, que no comprenden
20 interdependencias específicas, se pueden realizar por orden arbitrario o en paralelo. Más aún, se debería observar
que el procedimiento 900 se puede suplementar por medio de cualquiera de las características y funcionalidades
descritas en esta invención con respecto a los codificadores de audio y decodificadores de audio.

10. Procedimiento según la Fig. 10

25

[0113] La Fig. 10 muestra un diagrama de flujo de un procedimiento 1000 que proporciona al menos cuatro
señales de canales de audio sobre la base de una representación codificada.

[0114] El procedimiento 1000 comprende el suministro 1010 de una primera señal de mezcla descendente y
30 una segunda señal de mezcla descendente sobre la base de una representación codificada conjuntamente de la
primera señal de mezcla descendente y la segunda señal de mezcla descendente que utilizan una decodificación
multicanal, el suministro 1020 de al menos una primera señal del canal de audio y una segunda señal del canal de
audio sobre la base de la primera señal de mezcla descendente que utilizan una decodificación multicanal, el
suministro 1030 de al menos una tercera señal de canal de audio y una cuarta señal de canal de audio sobre la base
35 de la segunda señal de mezcla descendente que utilizan una decodificación multicanal, la realización 1040 de una
extensión de ancho de banda multicanal sobre la base de la primera señal del canal de audio y la tercera señal del
canal de audio con el fin de obtener una primera señal del canal con ancho de banda extendido y una tercera señal
del canal con ancho de banda extendido y la realización 1050 de una extensión del ancho de banda multicanal sobre
la base de la segunda señal del canal de audio y la cuarta señal del canal de audio con el fin de obtener una segunda
40 señal del canal con ancho de banda extendido y una cuarta señal del canal con ancho de banda extendido.

[0115] Se deberá observar que algunas de las etapas del procedimiento 1000 se pueden realizar en paralelo
o en un orden diferente. Más aún, se debería observar que el procedimiento 1000 puede ser suplementado por
cualquiera de las características y funcionalidades descritas en esta invención con respecto al codificador de audio y
45 al decodificador de audio.

11. Las realizaciones según las Figs. 11, 12 y 13

[0116] A continuación, se describirán algunas realizaciones adicionales según la presente invención y las
50 consideraciones subyacentes.

[0117] La Fig. 11 muestra un diagrama esquemático en bloques de un codificador de audio 1100 según una
realización de la invención. El codificador de audio 1100 está configurado para recibir una señal del canal inferior
izquierdo 1110, una señal del canal superior izquierdo 1112, una señal del canal inferior derecho 1114 y una señal del
55 canal superior derecho 1116.

[0118] El codificador de audio 1100 comprende un primer codificador multicanal de audio (o codificación) 1120,
que es un codificador de audio Surround MPEG 2-1-2 (o codificación) o un codificador de audio estéreo unificado (o
codificación) y que recibe la señal del canal inferior izquierdo 1110 y la señal del canal superior izquierdo 1112. El

primer codificador multicanal de audio 1120 proporciona una señal de mezcla descendente 1122 y, opcionalmente, una señal residual izquierda 1124. Más aún, el codificador de audio 1100 comprende un segundo codificador multicanal (o codificación) 1130, que es un codificador MPEG–surround 2–1–2 (o codificación) o un codificador estéreo unificado (o codificación) que recibe la señal del canal inferior derecho 1114 y la señal del canal superior derecho 1116. El segundo codificador multicanal de audio 1130 proporciona una señal de mezcla descendente derecho 1132 y, opcionalmente, una señal residual derecha 1134. El codificador de audio 1100 también comprende un codificador estéreo (o codificación) 1140, que recibe la señal de mezcla descendente izquierda 1122 y la señal de mezcla descendente derecha 1132.

10 **[0119]** Más aún, la primera codificación estéreo 1140, que es una codificación estéreo de predicción compleja, recibe una información de modelo psicoacústico 1142 de un modelo psicoacústico. Por ejemplo, la información del psicomodelo 1142 puede describir la relevancia psicoacústica de diferentes bandas de frecuencia o subbandas de frecuencia, efectos de enmascaramiento psicoacústico y similares. La codificación estéreo 1140 proporciona un elemento del par de canales (CPE) de “mezcla descendente”, que se indica con 1144 y que describe la señal de
15 mezcla descendente izquierda 1122 y la señal de mezcla descendente derecha 1132 en una forma conjuntamente codificada. Más aún, el codificador de audio 1100 opcionalmente comprende un segundo codificador estéreo (o codificación) 1150, que está configurado para recibir la señal residual izquierda opcional 1124 y la señal residual derecha opcional 1134, así como la información del modelo psicoacústico 1142. La segunda codificación estéreo 1150, que es una codificación estéreo de predicción compleja, está configurado para proporcionar un elemento del par de
20 canales (CPE) “residual”, que representa la señal residual izquierda 1124 y la señal residual derecha 1134 en una forma conjuntamente codificada.

[0120] El codificador 1100 (así como los otros codificadores de audio descritos en esta invención) está basado en la idea de que las dependencias de señales horizontales y verticales se aprovechan combinando jerárquicamente
25 herramientas estéreo USAC disponibles (es decir, conceptos de codificación que están disponibles en la codificación USAC). Los pares de canales vecinos verticalmente se combinan utilizando Surround MPEG 2–1–2 o estéreo unificado (designado con 1120 y 1130) con una señal residual de banda limitada o de banda completa (designado con 1124 y 1134). La salida de cada par de canales verticales es una señal de mezcla descendente 1122, 1132 y, para el estéreo unificado, una señal residual 1124, 1134. A fin de satisfacer los requerimientos perceptuales para
30 desenmascaramiento binaural, ambas señales de mezcla descendente 1122, 1132 se combinan de forma horizontal y se codifican de forma conjunta por uso de predicción compleja (codificador 1140) en el dominio MDCT, que incluye la posibilidad de codificación izquierda–derecha y del lado medio. El mismo procedimiento se puede aplicar a las señales residuales combinadas horizontalmente 1124, 1134. Este concepto se ilustra en la Fig. 11

35 **[0121]** La estructura jerárquica explicada en relación con la Fig. 11 se puede conseguir habilitando ambas herramientas estéreo (por ejemplo, ambas herramientas estéreo USAC) y recurriendo a canales entre medio. De esta manera, no es necesaria una etapa adicional de pre/posprocesamiento y la sintaxis de la corriente de bits para la transmisión de las cargas útiles de la herramienta queda inalterada (por ejemplo, sustancialmente inalterada cuando se compara con el estándar USAC). Esta idea da como resultado la estructura del codificador que se muestra en la
40 Fig. 12

[0122] La Fig. 12 muestra un diagrama esquemático en bloques de un codificador de audio 1200, según una realización de la invención. El codificador de audio 1200 está configurado para recibir una primera señal del canal 1210, una segunda señal del canal 1212, una tercera señal del canal 1214 y una cuarta señal del canal 1216. El
45 codificador de audio 1200 está configurado para proporcionar una corriente de bits 1220 de un primer elemento del par de canales y una corriente de bits 1222 de un segundo elemento del par de canales. El codificador de audio 1200 comprende un primer codificador multicanal 1230, que es un codificador MPEG–Surround 2–1–2 o un codificador estéreo unificado y que recibe la primera señal del canal 1210 y la segunda señal del canal 1212. Más aún, el primer codificador multicanal 1230 proporciona una primera señal de mezcla descendente 1232, una carga útil Surround
50 MPEG 1236 y, opcionalmente, una primera señal residual 1234. El codificador de audio 1200 también comprende un segundo codificador multicanal 1240 que es un codificador Surround MPEG 2–1–2 o un codificador estéreo unificado y que recibe la tercera señal del canal 1214 y la cuarta señal del canal 1216. El segundo codificador multicanal 1240 proporciona una primera señal de mezcla descendente 1242, una carga útil MPEG 1246 y, opcionalmente, una
segunda señal residual 1244.

55 **[0123]** El codificador de audio 1200 también comprende una primera codificación estéreo 1250, que es una codificación estéreo de predicción compleja. La primera codificación estéreo 1250 recibe la primera señal de mezcla descendente 1232 y la segunda señal de mezcla descendente 1242. La primera codificación estéreo 1250 proporciona una representación codificada conjuntamente 1252 de la primera señal de mezcla descendente 1232 y la segunda

señal de mezcla descendente 1242, en la que la representación codificada conjuntamente 1252 puede comprender una representación de una señal de mezcla descendente (común) (de la primera señal de mezcla descendente 1232 y de la segunda señal de mezcla descendente 1242) y de una señal residual común (de la primera señal de mezcla descendente 1232 y de la segunda señal de mezcla descendente 1242). Más aún, la (primera) codificación estéreo de predicción compleja 1250 proporciona una carga útil de predicción compleja 1254, que normalmente comprende uno o varios coeficientes de predicción compleja. Más aún, el codificador de audio 1200 también comprende una segunda codificación estéreo 1260, que es una codificación estéreo de predicción compleja. La segunda codificación estéreo 1260 recibe la primera señal residual 1234 y la segunda señal residual 1244 (o valores de entrada cero, si no hay ninguna señal residual proporcionada por los codificadores multicanal 1230, 1240). La segunda codificación estéreo 1260 proporciona una representación codificada conjuntamente 1262 de la primera señal residual 1234 y de la segunda señal residual 1244, que puede comprender, por ejemplo, una señal de mezcla descendente (común) (de la primera señal residual 1234 y de la segunda señal residual 1244) y una señal residual común (de la primera señal residual 1234 y de la segunda señal residual 1244). Más aún, la codificación estéreo de predicción compleja 1260 proporciona una carga útil de predicción compleja 1264 que normalmente comprende uno o varios coeficientes de predicción.

[0124] Más aún, el codificador de audio 1200 comprende un modelo psicoacústico 1270, que proporciona una información que controla la primera codificación estéreo de predicción compleja 1250 y la segunda codificación estéreo de predicción compleja 1260. Por ejemplo, la información proporcionada por el modelo psicoacústico 1270 puede describir qué bandas de frecuencia o intervalos de frecuencia tienen mucha relevancia psicoacústica y deberían ser codificadas con alta precisión. Sin embargo, se debería observar que el uso de la información proporcionada por el modelo psicoacústico 1270 es opcional.

[0125] Más aún, el codificador de audio 1200 comprende un primer codificador y multiplexador 1280 que recibe la representación codificada conjuntamente 1252 de la primera codificación estéreo de predicción compleja 1250, la carga útil de predicción compleja 1254 de la primera codificación estéreo de predicción compleja 1250 y la carga útil Surround MPEG 1236 del primer codificador multicanal de audio 1230. Más aún, la primera codificación y multiplexación 1280 puede recibir información del modelo psicoacústico 1270, que describe, por ejemplo, qué precisión de codificación se deberá aplicar a qué bandas de frecuencia o subbandas de frecuencia, teniendo en cuenta los efectos de enmascaramiento psicoacústico y similares. Conforme a ello, la primera codificación y multiplexación 1280 proporciona la primera corriente de bits del elemento de par del canal 1220.

[0126] Más aún, el codificador de audio 1200 comprende una segunda codificación y multiplexación 1290, el cual está configurado para recibir la representación codificada conjuntamente 1262 proporcionada por la segunda codificación estéreo de predicción compleja 1260, la carga útil de predicción compleja 1264 proporcionada por la segunda codificación estéreo de predicción compleja 1260 y la carga útil Surround MPEG 1246 proporcionada por el segundo codificador multicanal de audio 1240. Más aún, la segunda codificación y multiplexación 1290 puede recibir una información del modelo psicoacústico 1270. Conforme a ello, la segunda codificación y multiplexación 1290 proporciona la segunda corriente de bits del elemento de par de canales 1222.

[0127] Respecto de la funcionalidad del codificador de audio 1200, se hace referencia a las explicaciones anteriores y también a las explicaciones con respecto a los codificadores de audio según las Figs. 2, 3, 5 y 6.

[0128] Más aún, se debería observar que este concepto se puede extender al uso de múltiples cajas Surround MPEG para la codificación conjunta de canales relacionados de forma horizontal, vertical o de otra forma geométrica y combinación de las señales de mezcla descendente y las señales residuales con los pares estéreo de predicción compleja, considerando sus propiedades geométricas y perceptuales. Esto lleva a una estructura del decodificador generalizada.

[0129] A continuación, se describirá la implementación de un elemento del canal quad. En un sistema de codificación de auto tridimensional, se utiliza la combinación jerárquica de cuatro canales para formar un elemento del canal quad (QCE). Un QCE consiste en dos elementos de pares del canal USAC (CPE) (o proporciona dos elementos del par de canales USAC o recibe elementos del par de canales USAC). Los pares de canales verticales se combinan utilizando MPS 2–1–2 o estéreo unificado. Los canales de mezcla descendente se codifican conjuntamente en el primer elemento del par de canales CPE. Si se aplica la codificación residual, las señales residuales se codifican conjuntamente en el segundo elemento del par de canales CPE, incluso la señal en el segundo CPE se establece en cero. Ambos elementos del par de canales CPE usan predicción compleja para una codificación estéreo conjunta, incluso la posibilidad de una codificación izquierda–derecha y del lado medio. Para preservar las propiedades estéreo perceptuales de la parte de alta frecuencia de la señal, se aplica SBR estéreo (replicación de ancho de banda

espectral) entre el par de canales izquierdo/derecho superior y el par de canales izquierdo/derecho inferior, por una etapa adicional de recurso antes de la aplicación de SBR.

[0130] Una posible estructura de decodificador se describirá tomando como referencia la Fig. 13 que muestra un diagrama esquemático en bloques de un decodificador de audio según una realización. El decodificador de audio 1300 está configurado para recibir una primera corriente de bits 1310 que representa un primer elemento del par de canales y una segunda corriente de bits 1312 que representan un segundo elemento del par de canales. Sin embargo, la primera corriente de bits 1310 y la segunda corriente de bits 1312 se puede incluir en una corriente de bits global común.

[0131] El decodificador de audio 1300 está configurado para proporcionar una primera señal de canal con ancho de banda extendido 1320, que puede, por ejemplo, representar una posición izquierda inferior de una escena de audio, una segunda señal de canal con ancho de banda extendido 1322, que puede, por ejemplo, representar una posición izquierda superior de la escena de audio, una tercera señal de canal con ancho de banda extendido 1324, que puede, por ejemplo, estar asociada con una posición derecha inferior de la escena de audio y una cuarta señal de canal con ancho de banda extendido 1326, que puede, por ejemplo, estar asociado con una posición derecha superior de la escena de audio.

[0132] El decodificador de audio 1300 comprende una primera corriente de bits decodificadora 1330, el cual está configurado para recibir la corriente de bits 1310 para el primer elemento del par de canales y para proporcionar, sobre la base del mismo, una representación codificada en forma conjunta de dos señales de la mezcla descendente, una carga útil de predicción compleja 1334, una carga útil de Surround MPEG 1336 y una carga útil de replicación de ancho de banda espectral 1338. El decodificador de audio 1300 también comprende una primera decodificación estéreo de predicción compleja 1340, el cual está configurado para recibir la representación codificada en forma conjunta 1332 y la carga útil de predicción compleja 1334 y para proporcionar, sobre la base de las mismas, una primera señal de la mezcla descendente 1342 y una segunda señal de la mezcla descendente 1344. De modo similar, el decodificador de audio 1300 comprende un segundo decodificador de la corriente de bits 1350 que está configurado para recibir la corriente de bits 1312 del segundo elemento del canal y para proporcionar, sobre la base del mismo, una representación codificada en forma conjunta 1352 de dos señales residuales, una carga útil de predicción compleja 1354, una carga útil de Surround MPEG 1356 y una carga de bits de replicación del ancho de banda espectral 1358. El decodificador de audio también comprende una segunda decodificación estéreo de predicción compleja 1360, que proporciona una primera señal residual 1362 y una segunda señal residual 1364 sobre la base de la representación codificada en forma conjunta 1352 y la carga útil de predicción compleja 1354.

[0133] Además, el decodificador de audio 1300 comprende una primera decodificación multicanal tipo Surround MPEG 1370, que es una decodificación 2–1–2 Surround MPEG o una decodificación estéreo unificada. El primer decodificador multicanal tipo Surround MPEG 1370 recibe la primera señal de la mezcla descendente 1342, la primera señal residual 1362 (opcional) y la carga útil de Surround MPEG 1336 y proporciona, sobre la base de las mismas, una primera señal del canal de audio 1372 y una segunda señal del canal de audio 1374. El decodificador de audio 1300 también comprende un segundo decodificador multicanal tipo Surround MPEG 1380, que es un Decodificador multicanal 2–1–2 Surround MPEG o un decodificador multicanal estéreo unificado. El segundo decodificador multicanal tipo Surround MPEG 1380 recibe la segunda señal de la mezcla descendente 1344 y la segunda señal residual 1364 (opcional), así como la carga útil de Surround MPEG 1356, y proporciona, sobre la base de la misma, una tercera señal del canal de audio 1382 y la cuarta señal del canal de audio 1384. El decodificador de audio 1300 también comprende una primera replicación del ancho de banda espectral estéreo 1390, el cual está configurado para recibir la primera señal del canal de audio 1372 y la tercera señal del canal de audio 1382, así como la carga útil de replicación de ancho de banda espectral 1338, y para proporcionar, sobre la base de la misma, la primera señal de canal con ancho de banda extendido 1320 y la tercera señal de canal con ancho de banda extendido 1324. Además, el decodificador de audio comprende una segunda replicación del ancho de banda espectral estéreo 1394, el cual está configurado para recibir la segunda señal del canal de audio 1374 y la cuarta señal del canal de audio 1384, así como la carga útil de replicación de ancho de banda espectral 1358 y para proporcionar, sobre la base de la misma, la segunda señal de canal con ancho de banda extendido 1322 y la cuarta señal de canal con ancho de banda extendido 1326.

[0134] Con respecto a la funcionalidad del decodificador de audio 1300, se hace referencia a la anterior discusión, y también a la discusión del decodificador de audio según las Fig. 2, 3, 5 y 6.

[0135] A continuación, se describirá un ejemplo de una corriente de bits que se puede utilizar para la codificación/decodificación de audio descrita en esta invención con referencia a las Figs. 14a y 14b. Cabe mencionar

que la corriente de bits, por ejemplo, puede ser una extensión de la corriente de bits usada en el codificación de audio y voz unificada (USAC), que se describe en el estándar mencionado anteriormente (ISO/IEC 23003-3:2012). Por ejemplo, las cargas útiles de Surround MPEG 1236, 1246, 1336, 1356 y las cargas útiles de predicción compleja 1254, 1264, 1334, 1354 se pueden transmitir como para los elementos del par de canales heredados (es decir, para los elementos del par de canales según el estándar USAC). Para la señalización del uso de un elemento del canal cuádruple QCE, la configuración del par de canal USAC se puede extender en dos bits, como se muestra en la Fig. 14A. En otras palabras, dos bits denominados "qcelIndex" se pueden añadir al elemento de la corriente de bits USAC "UsacChannelPairElementConfig()". El significado del parámetro representado por los bits "qcelIndex" se puede definir, por ejemplo, como se muestra en la tabla de la Fig. 14B.

[0136] Por ejemplo, dos elementos del par de canales que forman un QCE se pueden transmitir como elementos consecutivos, primero el CPE que contiene los canales de la mezcla descendente la carga útil MPS a la primera caja MPS, segundo el CPE que contiene la señal residual (o señal de audio cero para la codificación MPS 2-1-2) y la carga útil MPS para la segunda caja MPS.

[0137] En otras palabras, existe solo una pequeña sobrecarga de señalización cuando se compara con la corriente de bits USAC convencional para transmitir un elemento del canal QCE cuádruple.

[0138] Sin embargo, también se pueden utilizar naturalmente diferentes formatos de corriente de bits.

12. Ambiente de codificación/decodificación

[0139] A continuación, se describirá un ambiente de codificación/decodificación de audio en el que se pueden aplicar los conceptos según la presente invención.

[0140] Sistema de códec de audio en 3D, en el que se pueden utilizar los conceptos según la presente invención, está basado en un códec de MPEG-D USAC para la decodificación de señales de canal y objeto. Para aumentar la eficacia de la codificación de una gran cantidad de objetos, se ha adaptado la tecnología MPEG SAOC. Tres tipos de renderizadores realizan las tareas de renderización de los objetos a los canales, renderización de los canales a los auriculares o renderización de los canales a una configuración de altavoces diferentes. Cuando las señales de objetos se transmiten de manera explícita o codifican paramétricamente utilizando SAOC, la información de metadatos de objeto correspondiente se comprime y multiplexa en la corriente de bits de audio 3D.

[0141] La Fig. 15 muestra un diagrama esquemático en bloques de dicho codificador de audio, y la Fig. 16 muestra un diagrama esquemático en bloques de tal decodificador de audio. 15 y 16 muestran los diferentes bloques algorítmicos del sistema de audio 3D.

[0142] Tomando como referencia a la Fig. 15, que muestra un diagrama esquemático en bloques de un codificador de audio 3D 1500, se explicarán algunos detalles. El codificador 1500 comprende un prrenderizador/mezclador 1510 opcional, que recibe una o más señales del canal 1512 y una o más señales del objeto 1514 y proporciona, sobre la base de las mismas, una o más señales del canal 1516 así como una o más señales del objeto 1518, 1520. El codificador de audio también comprende un codificador USAC 1530 y, opcionalmente, un codificador SAOC 1540. El codificador SAOC 1540 está configurado para proporcionar uno o más canales de transporte SAOC 1542 y una información secundaria SAOC 1544 sobre la base de uno o más objetos 1520 provistos al codificador SAOC. Además, el codificador USAC 1530 está configurado para recibir las señales del canal 1516 que comprende canales y objetos prrenderizados desde el prrenderizador/mezclador, para recibir una o más señales del objeto 1518 desde el prrenderizador/mezclador y para recibir uno o más canales de transporte SAOC 1542 e información secundaria SAOC 1544, y proporciona, sobre la base de la misma, una representación codificada 1532. Además, el codificador de audio 1500 también comprende un codificador de metadatos del objeto 1550 y que está configurado para recibir metadatos del objeto 1552 (que se puede evaluar mediante el prrenderizador/mezclador 1510) y para codificar los metadatos del objeto para obtener metadatos del objeto codificado 1554. Los metadatos codificados también es recibido por el codificador USAC 1530 y usado para proporcionar la representación codificada 1532.

[0143] Algunos detalles respecto de los componentes individuales del codificador de audio 1500 se describirán a continuación.

[0144] Tomando como referencia a la Fig. 16, se describirá un decodificador de audio 1600. El decodificador de audio 1600 está configurado para recibir una representación codificada 1610 y para proporcionar, sobre la base de

la misma, señales del altavoz multicanal 1612, señales de auriculares 1614 y/o señales del altavoz 1616 en un formato alternativo (por ejemplo, en un formato 5.1).

[0145] El decodificador de audio 1600 comprende un decodificador USAC 1620, y proporciona una o más 5 señales del canal 1622, una o más señales del objeto prrenderizadas 1624, una o más señales del objeto 1626, uno o más canales de transporte SAOC 1628, una información secundaria SAOC 1630 y una información de metadatos del objeto comprimida 1632 sobre la base de la representación codificada 1610. El decodificador de audio 1600 también comprende un renderizador del objeto 1640 que está configurado para proporcionar una o más señales del 10 objeto renderizadas 1642 sobre la base de la señal del objeto 1626 y una información de metadatos del objeto 1644, en el que la información de metadatos del objeto 1644 se proporciona mediante un decodificador de metadatos del objeto 1650 sobre la base de la información de metadatos del objeto comprimida 1632. El decodificador de audio 1600 también comprende, opcionalmente, un decodificador SAOC 1660, que está configurado para recibir el canal de transporte de SAOC 1628 y la información secundaria SAOC 1630, y para proporcionar, sobre la base de la misma, una o más señales del objeto renderizadas 1662. El decodificador de audio 1600 también comprende un mezclador 15 1670, que está configurado para recibir las señales del canal 1622, las señales del objeto prrenderizadas 1624, las señales del objeto renderizadas 1642, y las señales del objeto renderizadas 1662, y para proporcionar, sobre la base de las mismas, una pluralidad de señales del canal mixtas 1672 que pueden constituir, por ejemplo, las señales del altavoz multicanal 1612. El decodificador de audio 1600, por ejemplo, también puede comprender un renderizador binaural 1680, que está configurado para recibir las señales del canal mixtas 1672 y para proporcionar, sobre la base 20 de las mismas, las señales para los auriculares 1614. Además, el decodificador de audio 1600 puede comprender una conversión de formato 1690, que está configurado para recibir las señales del canal mixtas 1672 y una información del diseño de reproducción 1692 y para proporcionar, sobre la base de la misma, una señal del altavoz 1616 para una configuración de altavoz alternativa.

25 **[0146]** A continuación, se describirán algunos detalles de los componentes del codificador de audio 1500 y del decodificador de audio 1600.

Prrenderizador/mezclador

30 **[0147]** El prrenderizador/mezclador 1510 se puede utilizar opcionalmente para convertir una escena de entrada de canal más objeto en una escena de canal antes de la codificación. Desde el punto de vista funcional, por ejemplo, puede ser idéntico al renderizador del objeto/mezclador que se describe a continuación. La prrenderización de los objetos, por ejemplo, puede asegurar una entropía de señal determinista en la entrada del codificador que es básicamente independiente del número de señales del objeto simultáneamente activas. En la prrenderización de los 35 objetos, no se requiere la transmisión de los metadatos del objeto. Las señales del objeto discretas se renderizan a la disposición del canal que el codificador está configurado para usar. Los pesos de los objetos para cada canal se obtienen de los metadatos del objeto asociados (OAM) 1552.

Códec de núcleo USAC

40 **[0148]** El códec de núcleo 1530, 1620 para las señales de altavoz de los canales, señales de objetos discretos, señales de la mezcla descendente del objeto y señales prrenderizados está basado en la tecnología MPEG-D USAC. Se ocupa de la codificación de la multitud de señales mediante la creación de canal y la información de mapeo de objetos basado en la información geométrica y semántica del canal de la entrada y la asignación de objetos. Esta 45 información de mapeo describe cómo los canales de entrada y los objetos se mapean en los elementos del canal USAC (EPP, SCE, LFE) y la información correspondiente se transmite al decodificador. Todas las cargas útiles adicionales como los datos SAOC o metadatos de objetos han pasado a través de los elementos de extensión y se han tenido en cuenta en el control de la frecuencia de los codificadores.

50 **[0149]** La codificación de objetos es posible en diferentes formas, según los requerimientos de frecuencia/distorsión y los requerimientos de interactividad para el renderizador. Las siguientes variantes de codificación del objeto son posibles:

1. Objetos prrenderizados: las señales del objeto se prrenderizan y se mezclan con las señales del canal 22.2 55 antes de la codificación. La subsiguiente cadena de codificación ve las señales del canal. 22.2

2. Formas de onda de objetos discretas: los objetos se suministran como formas de onda monofónicas al codificador. El codificador usa elementos de canal de SCE individuales para transferir los objetos además de las señales del canal. Los objetos decodificados se renderizan y mezclan en el lado del receptor. La información de metadatos del objeto

comprimida se transmite al receptor/renderizador en paralelo.

3. Formas de onda del objeto paramétrico: las propiedades del objeto y la relación entre sí se describen mediante los parámetros SAOC. La mezcla descendente de las señales del objeto se codifica con USAC. La información paramétrica se transmite en paralelo. El número de canales de mezcla descendente se elige en función del número de objetos y la velocidad de datos global. La información de metadatos del objeto comprimida se transmite al renderizador AOC.

SAOC

10

[0150] El codificador SAOC 1540 y el decodificador SAOC 1660 de señales de objeto están basadas en la tecnología SAOC MPEG. El sistema es capaz de volver a crear, modificar y renderizar un número de objetos de audio sobre la base de un número menor de canales de transmisión y datos paramétricos adicionales (diferencias de nivel de objeto OLD, correlaciones interobjetos IOC, ganancias de la mezcla descendente DMG). Los datos paramétricos adicional exhiben una velocidad de datos significativamente menor que la requerida para la transmisión de todos los objetos individualmente, lo que produce una codificación muy eficaz. El codificador SAOC toma como entrada las señales de objetos/canal en forma de ondas monofónicas y emite la información paramétrica (que se empaqueta en la corriente de bits de audio 3D–1532, 1610) y los canales de transporte SAOC (que se codifican mediante elementos de canal único y se transmiten).

20

[0151] El decodificador SAOC 1600 reconstruye las señales del objeto/canal procedentes de los canales de transporte SAOC decodificados 1628 y la información paramétrica 1630, y genera la escena de audio de salida basada en el diseño de reproducción, la información de metadatos del objeto descomprimida y, opcionalmente, en la información de la interacción del usuario.

25

Códec de metadatos del objeto

[0152] Para cada objeto, los metadatos asociados que especifican la posición geométrica y el volumen del objeto en el espacio 3D se codifican eficazmente mediante la cuantificación de las propiedades de los objetos en el tiempo y el espacio. Los metadatos de objeto comprimido cOAM 1554, 1632 se transmiten al receptor como información secundaria.

30

Renderizador del objeto/Mezclador

[0153] El renderizador del objeto utiliza los metadatos de objeto comprimidos para generar las formas de onda del objeto según el formato de reproducción dado. Cada objeto se renderiza en ciertos canales de salida según sus metadatos. La salida de este bloque resulta de la suma de los resultados parciales. Si se decodifica el contenido basado en el canal, así como los objetos discretos/paramétricos, las formas de onda basadas en el canal y las formas de onda de los objetos renderizados se mezclan antes de la salida de las formas de onda resultantes (o antes de alimentar a un módulo de posprocesador como el renderizador binaural o el módulo renderizador del altavoz).

40

Renderizador binaural

[0154] El módulo del renderizador binaural 1680 produce una mezcla descendente binaural del material de audio multicanal, de manera que cada canal de entrada está representado por una fuente de sonido virtual. El procesamiento se realiza por trama en un dominio QMF. La binauralización está basada en las respuestas de impulso del espacio binaural medido.

45

Renderizador de altavoz/Conversión de formato

50

[0155] El renderizador de altavoz 1690 convierte entre la configuración del canal transmitida y el formato de reproducción deseada. En consecuencia, se llama “convertidor de formato” en la descripción siguiente. El convertidor de formato realiza conversiones para reducir el número de canales de salida, es decir, se crean mezclas descendentes. El sistema genera automáticamente las matrices de la mezcla descendente optimizadas para una determinada combinación de formatos de entrada y salida y aplica estas matrices en un proceso de mezcla descendente. El convertidor de formato permite configuraciones de altavoz estándares, así como para configuraciones aleatorias con posiciones de altavoces no estándares.

55

[0156] La Fig. 17 muestra un diagrama esquemático en bloques del convertidor de formato. Como se puede

observar, el convertidor de formato 1700 recibe las señales de salida del mezclador 1710, por ejemplo, las señales del canal mixtas 1672 y proporciona las señales del altavoz 1712, por ejemplo, las señales del hablante 1616. El convertidor de formato comprende un proceso de mezcla descendente 1720 en el dominio QMF y un configurador de mezcla descendente 1730, en el que el configurador de la mezcla descendente proporciona la información de configuración para el proceso de mezcla descendente 1720 sobre la base de una información de la disposición de salida del mezclador 1732 y una información de la disposición de reproducción 1734.

[0157] Además, cabe mencionar que los conceptos descritos anteriormente, por ejemplo el codificador de audio 100, el decodificador de audio 200 o 300, el codificador de audio 400, el decodificador de audio 500 o 600, los procedimientos 700, 800, 900, o 1000, el codificador de audio 1100 o 1200 y el decodificador de audio 1300 se pueden utilizar dentro del codificador de audio 1500 y/o dentro del decodificador de audio 1600. Por ejemplo, los codificadores/decodificadores de audio mencionados antes se pueden utilizar para codificar o decodificar las señales del canal que se asocian con posiciones espaciales diferentes.

15 13. Realizaciones alternativas

[0158] A continuación se describirán algunas realizaciones adicionales.

[0159] Tomando como referencia las Figs. 18 a 21, se explicarán realizaciones adicionales según la invención.

[0160] Cabe mencionar que un llamado "elemento de canal cuádruple" (QCE) se puede tener en cuenta como una herramienta de un decodificador de audio, que se puede utilizar, por ejemplo, para decodificar el contenido de audio tridimensional.

[0161] En otras palabras, el elemento de canal cuádruple (QCE) es un procedimiento para la codificación conjunta de cuatro canales para una codificación más eficaz de los canales distribuidos en forma horizontal y vertical. Un QCE consiste en dos CPE consecutivos y se forma mediante la combinación jerárquica de la herramienta de estéreo conjunto con la posibilidad de la herramienta de predicción de estéreo compleja en dirección horizontal y la herramienta de estéreo basada en Surround MPEG en dirección vertical. Esto se obtiene al permitir que ambas herramientas de estéreo intercambien los canales de salida entre la aplicación de las herramientas. La SBR estéreo se realiza en dirección horizontal para preservar las relaciones izquierda-derecha de las frecuencias altas.

[0162] La Fig. 18 muestra una estructura topológica de un QCE. Cabe señalar que el QCE de la Fig. 18 es muy similar al QCE de la Fig. 11, de manera que se hace referencia a las explicaciones anteriores. Sin embargo, cabe señalar que en el QCE de la Fig. 18 no es necesario hacer uso del modelo psicoacústico cuando se realiza la predicción estéreo compleja (mientras que, dicho uso naturalmente, es opcionalmente posible). Por otra parte, se puede observar que la primera replicación del ancho de banda espectral estéreo (SBR estéreo) se lleva a cabo sobre la base del canal inferior izquierdo y el canal derecho inferior, y que esa segunda replicación de ancho de banda espectral estéreo (SBR estéreo) se lleva a cabo sobre la base del canal superior izquierdo y el canal superior derecho.

[0163] A continuación, se proporcionarán algunos términos y definiciones que se pueden aplicar en algunas realizaciones.

[0164] Un elemento de datos `qcelIndex` indica un modo QCE de un CPE. Respecto del significado de la variable de corriente de bits `qcelIndex`, se hace referencia a la Fig. 14b. Cabe señalar que `qcelIndex` describe si dos elementos posteriores de tipo `UsacChannelPairElement` () se tratan como un elemento de canal cuádruple (QCE). Los diferentes modos de QCE se proporcionan en la Fig. 14b. El `qcelIndex` será el mismo para los dos elementos posteriores que forman un QCE.

[0165] A continuación, se definirán algunos elementos de ayuda que se pueden utilizar en algunas realizaciones según la invención:

`cplx_out_dmx_L[]` primer canal del primer CPE después de decodificación estéreo de predicción compleja

`cplx_out_dmx_R[]` segundo canal del primer CPE después de la decodificación estéreo de predicción compleja

`cplx_out_res_L[]` segundo CPE después de la decodificación estéreo de predicción compleja (cero si `qcelIndex = 1`)

`cplx_out_res_R[]` segundo canal del segundo CPE después de la decodificación estéreo de predicción compleja (cero

si $qcelIndex = 1$)

$mps_out_L_1[]$ primer canal de salida de la primera caja MPS $mps_out_L_2[]$ segundo canal de salida de la primera caja MPS

5 $mps_out_R_1[]$ primer canal de salida de la segunda caja MPS

$mps_out_R_2[]$ segundo canal de salida de la segunda caja MPS

10 $sbr_out_L_1[]$ primer canal de salida de la primera caja SBR estéreo

$sbr_out_R_1[]$ segundo canal de salida de la primera caja SBR estéreo

$sbr_out_L_2[]$ primer canal de salida de la segunda caja SBR estéreo

15 $sbr_out_R_2[]$ segundo canal de salida de la segunda caja SBR estéreo

[0166] A continuación, se explicará un proceso de decodificación que se realiza en una realización según la invención.

20 **[0167]** El elemento de sintaxis (o elemento de la corriente de bits, o elemento de datos) $qcelIndex$ en `UsacChannelPairElementConfig()` indica si un CPE pertenece a un QCE y si se utiliza la codificación residual. En caso de que $qcelIndex$ sea diferente a 0, el CPE actual forma un QCE junto con su posterior elemento que será un CPE que tiene el mismo $qcelIndex$. SBR estéreo siempre se utiliza para el QCE, por lo tanto el elemento de sintaxis `stereoConfigIndex` será 3 y `bsStereoSbr` será 1.

[0168] En caso de $qcelIndex == 1$ solo las cargas útiles para Surround MPEG y SBR y ningún dato de la señal de audio significativo está contenido en el segundo CPE y el elemento de sintaxis `bsResidualCoding` se ajusta a 0.

30 **[0169]** La presencia de una señal residual en el segundo CPE se indica mediante $qcelIndex == 2$. En este caso el elemento de sintaxis `bsResidualCoding` se ajusta a 1.

[0170] Sin embargo, también se pueden utilizar algunos esquemas de señalización simplificados diferentes y posibles.

35 **[0171]** La decodificación de Joint Stereo con posibilidad de predicción estéreo compleja se realiza como se describe en ISO/IEC 23003-3 subcláusula 7.7. La salida resultante del primer CPE son las señales de la mezcla descendente MPS $cplx_out_dmx_L[]$ y $cplx_out_dmx_R[]$. Si se utiliza la codificación residual (es decir, $qcelIndex == 2$), la salida del segundo CPE son las señales residuales MPS $cplx_out_res_L[]$, $cplx_out_res_R[]$, si no se ha transmitido la señal residual (es decir $qcelIndex == 1$), se insertan las señales cero.

[0172] Antes de aplicar la decodificación Surround MPEG, el segundo canal del primer elemento ($cplx_out_dmx_R[]$) y el primer canal del segundo ($cplx_out_res_L[]$) se intercambian.

45 **[0173]** La decodificación de Surround MPEG se realiza como se describe en ISO/IEC 23003-3 subcláusula 7.11. Si se utiliza la codificación residual, la decodificación se puede, sin embargo, modificar cuando se compara con la decodificación de Surround MPEG convencional en algunas realizaciones. La decodificación de Surround MPEG sin residual que utiliza SBR como se define en ISO/IEC 23003-3 subcláusula 7.11.2.7 (figure 23), se modifica de manera que la SBR estéreo también se utiliza para `bsResidualCoding == 1`, lo que da como resultado el esquema del decodificador que se muestra en la Fig. 19. La Fig. 19 muestra un diagrama esquemático en bloques de un codificador de audio para `bsResidualCoding == 0` y `bsStereoSbr == 1`.

50 **[0174]** Como puede verse en la Fig. 19, un decodificador de núcleo USAC 2010 proporciona una señal de mezcla descendente (DMX) 2012 a un decodificador MPS (Surround MPEG) 2020, que proporciona una primera señal de audio decodificada 2022 y una segunda señal de audio decodificada 2024. Un decodificador de SBR estéreo 2030 recibe la primera señal de audio decodificada 2022 y la segunda señal de audio decodificada 2024 y proporciona, sobre la base de las mismas una señal de audio extendida de ancho de banda izquierda 2032 y una señal de audio extendida de ancho de banda derecha 2034.

[0175] Antes de aplicar la SBR estéreo, el segundo canal del primer elemento (mps_out_L_2[]) y el primer canal del segundo elemento (mps_out_R_1[]) se intercambian para permitir la SBR estéreo. Después de la aplicación de la SBR estéreo, el segundo canal de salida del primer elemento (sbr_out_R_1[]) y el primer canal del segundo elemento (sbr_out_L_2[]) se intercambian nuevamente para restaurar el orden del canal de entrada.

5

[0176] Una estructura del decodificador QCE se ilustra en la Fig. 20, que muestra un esquema del decodificador de QCE.

[0177] Cabe señalar que el QCE del diagrama esquemático en bloques de la Fig. 20 es muy similar al diagrama esquemático en bloques de la Fig. 13, de manera que también se hace referencia a las explicaciones anteriores. Además, cabe mencionar que se ha añadido alguna marcación de señal en la Fig. 20, en la que se hace referencia a las definiciones de esta sección. Además, se muestra la reordenación de los canales que se realiza después de la SBR estéreo.

[0178] 21 muestra un diagrama esquemático en bloques de un codificador de canal cuádruple 2200, según una realización de la presente invención. En otras palabras, a codificador de canal cuádruple (elemento de canal cuádruple), que se puede tener en cuenta como una herramienta Codificador Core, se ilustra en la Fig. 21.

[0179] El codificador de canal cuádruple 2200 comprende una primera SBR estéreo 2210, que recibe una primera señal de entrada del canal izquierdo 2212 y una segunda señal de entrada del canal izquierdo 2214, y que proporciona, sobre la base de las mismas, una primera carga útil SBR 2215, una primera señal de salida de SBR del canal izquierdo 2216 y una primera señal de salida de SBR del canal derecho 2218. Además, el codificador de canal cuádruple 2200 comprende una segunda SBR estéreo, que recibe una segunda señal de entrada del canal izquierdo 2222 y una segunda señal de entrada del canal derecho 2224, y que proporciona, sobre la base de las mismas, una primera carga útil de SBR carga 2225, una primera señal de salida de SBR del canal izquierdo 2226 y una primera señal de salida de SBR del canal derecho 2228.

[0180] El codificador de canal cuádruple 2200 comprende un primer codificador multicanal tipo Surround MPEG (MPS 2-1-2 o estéreo unificado) 2230 que recibe la primera señal de salida de SBR del canal izquierdo 2216 y la segunda señal de salida de SBR del canal izquierdo 2226, y que proporciona, sobre la base de las mismas, una primera carga útil MPS 2232, una señal de la mezcla descendente Surround MPEG del canal izquierdo 2234 y, opcionalmente, una señal residual de Surround MPEG del canal izquierdo 2236. El codificador de canal cuádruple 2200 también comprende un segundo codificador multicanal tipo Surround MPEG (MPS 2-1-2 o estéreo unificado) 2240 que recibe la primera señal de salida de SBR del canal derecho 2218 y la segunda señal de salida de SBR del canal derecho 2228, y que proporciona, sobre la base de las mismas, una primer carga útil MPS 2242, una señal de la mezcla descendente Surround MPEG del canal derecho 2244 y, opcionalmente, una señal residual de Surround MPEG del canal derecho 2246.

[0181] El codificador de canal cuádruple 2200 comprende una primera codificación estéreo de predicción compleja 2250, que recibe la señal de la mezcla descendente Surround MPEG del canal izquierdo 2234 y la señal de la mezcla descendente Surround MPEG del canal derecho 2244, y que proporciona, sobre la base de las mismas, una carga útil de predicción compleja 2252 y una representación codificada en forma conjunta 2254 de la señal de la mezcla descendente Surround MPEG del canal izquierdo 2234 y la señal de la mezcla descendente Surround MPEG del canal derecho 2244. El codificador de canal cuádruple 2200 comprende una segunda codificación estéreo de predicción compleja 2260, que recibe la señal residual de Surround MPEG del canal izquierdo 2236 y la señal residual de Surround MPEG del canal derecho 2246, y que proporciona, sobre la base de las mismas, una carga útil de predicción compleja 2262 y una representación codificada en forma conjunta 2264 de la señal de la mezcla descendente Surround MPEG del canal izquierdo 2236 y la señal de la mezcla descendente Surround MPEG del canal derecho 2246.

50

[0182] El codificador de canal cuádruple también comprende un primer codificador de la corriente de bits 2270, que recibe la representación codificada en forma conjunta 2254, la carga útil de predicción compleja 2252, la carga útil MPS 2232 y la carga útil de SBR 2215 y proporciona, sobre la base de las mismas, una parte de la corriente de bits que representa un primer elemento del par de canales. El codificador de canal cuádruple también comprende un segundo codificador de la corriente de bits 2280, que recibe la representación codificada en forma conjunta 2264, la carga útil de predicción compleja 2262, la carga útil MPS 2242 y la carga útil de SBR 2225 y proporciona, sobre la base de las mismas, una parte de la corriente de bits que representa un primer elemento del par de canales.

55

14. Alternativas de implementación

- [0183]** Si bien algunos aspectos se han descrito en el contexto de un aparato, está claro que estos aspectos también representan una descripción del procedimiento correspondiente, en el que un bloque o dispositivo corresponde a una etapa del procedimiento o una función de una etapa del procedimiento. Análogamente, los aspectos descritos en el contexto de una etapa del procedimiento también representan una descripción de un correspondiente bloque o elemento o función de un aparato correspondiente. Algunas o todas las etapas del procedimiento se pueden ejecutar por (o mediante) un aparato de hardware como, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, algunas o más de las etapas más importantes del procedimiento se pueden ejecutar mediante tal aparato.
- [0184]** La señal de audio codificada de la invención se puede almacenar en un medio de almacenamiento digital o se puede transmitir en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cable, tal como Internet.
- [0185]** Según ciertos requerimientos de la implementación, las realizaciones de la invención se pueden implementar en el hardware o en el software. La implementación se puede realizar utilizando un medio de almacenamiento digital, por ejemplo un disquete, un DVD, un disco Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene señales de control legibles electrónicamente almacenadas, que cooperan (o son capaces de cooperar) con un sistema informático programable de manera que se lleva a cabo el procedimiento respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.
- [0186]** Algunas realizaciones según la invención comprenden un soporte de datos que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema informático programable, de manera que se lleva a cabo uno de los procedimientos descritos en esta invención.
- [0187]** En general, las realizaciones de la presente invención se pueden implementar como un producto de programa informático con un código del programa, el código del programa es operativo para realizar uno de los procedimientos, cuando el producto de programa informático se ejecuta en un ordenador. El código del programa, por ejemplo, se puede almacenar en un soporte legible por máquina.
- [0188]** Otras realizaciones comprenden el programa informático para realizar uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por máquina.
- [0189]** En otras palabras, una realización del procedimiento de la invención es, por lo tanto, un programa informático que tiene un código del programa para realizar uno de los procedimientos descritos en esta invención, cuando el programa informático se ejecuta en un ordenador.
- [0190]** Una realización adicional de los procedimientos de la invención es, por lo tanto, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para realizar uno de los procedimientos descritos en esta invención. El soporte de datos, el medio de almacenamiento digital o el medio grabado normalmente son tangibles y/o no transitorios.
- [0191]** Una realización adicional del procedimiento de la invención, en consecuencia, es una corriente de datos o una secuencia de señales que representan el programa informático para realizar uno de los procedimientos descritos en esta invención. La corriente de datos o la secuencia de señales, por ejemplo, se puede configurar para ser transferidos a través de una conexión de comunicación de datos, por ejemplo a través de Internet.
- [0192]** Una realización comprende además un medio de procesamiento, por ejemplo un ordenador, o un dispositivo lógico programable, configurado para o adaptado para realizar uno de los procedimientos descritos en esta invención.
- [0193]** Una realización adicional comprende además un ordenador que tiene instalado el programa informático para realizar uno de los procedimientos descritos en esta invención.
- [0194]** Una realización adicional según la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, en forma electrónica u óptica) un programa informático para realizar uno de los procedimientos descritos en esta invención a un receptor. El receptor, por ejemplo, puede ser un ordenador, un dispositivo móvil, un dispositivo de memoria o similares. El aparato o sistema, por ejemplo, pueden comprender un servidor de archivos para transferir el programa informático al receptor.

[0195] En algunas realizaciones, un dispositivo lógico programable (por ejemplo, una matriz de puertas programables por campo) se puede utilizar para realizar todas o algunas de las funcionalidades de los procedimientos descritos en esta invención. En algunas realizaciones, un matriz de puertas programables por campo puede cooperar con un microprocesador para realizar uno de los procedimientos descritos en esta invención. En general, los procedimientos se llevan a cabo preferentemente en cualquier aparato de hardware.

[0196] Las realizaciones anteriormente descritas son meramente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en esta invención serán evidentes para otros expertos en la técnica. Es la intención, en consecuencia, de que la invención solo está limitada por el alcance de las reivindicaciones de patentes inminentes y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones de la presente.

15. Conclusiones

15

[0197] A continuación, se proporcionarán algunas conclusiones.

[0198] Las realizaciones según la invención se basan en la consideración de que, para tener en cuenta las dependencias de señal entre los canales distribuidos en forma vertical y horizontal, cuatro canales se pueden codificar en forma conjunta mediante la combinación jerárquica de las herramientas de codificación estéreo conjuntas. Por ejemplo, los pares de canales verticales se combinan utilizando MPS 2–1–2 y/o estéreo unificado con la codificación residual de banda limitada o de banda completa. Con el fin de satisfacer los requisitos perceptivos para desenmascaramiento binaural, las mezclas descendentes de salida, por ejemplo, se codifican conjuntamente mediante el uso de la predicción compleja en el dominio de MCDT, que incluye la posibilidad de codificación izquierda–derecha y semilateral. Si las señales residuales están presentes se combinan horizontalmente utilizando el mismo procedimiento.

[0199] Por otra parte, cabe mencionar que las realizaciones según la invención superan algunas o todas las desventajas de la técnica anterior. Las realizaciones según la invención están adaptadas al contexto de audio 3D, en el que los canales de altavoces se distribuyen varias capas de altura, lo que produce pares de canales horizontales y verticales. Se ha encontrado que la codificación conjunta de solo dos canales como se define en USAC no es suficiente para tener en cuenta las relaciones espaciales y perceptuales entre los canales. Sin embargo, este problema se supera mediante las realizaciones según la invención.

[0200] Por otra parte, el Surround MPEG convencional se aplica en una etapa de pre/posprocesamiento adicional, de manera que las señales residuales se transmiten de forma individual y sin la posibilidad de la codificación estéreo conjunta, por ejemplo, para explorar las dependencias entre señales residuales radicales izquierda y derecha. En contraste, las realizaciones según la invención permiten una codificación/decodificación eficaz mediante el uso de dichas dependencias.

40

[0201] Para concluir adicionalmente, las realizaciones según la invención crean un aparato, un procedimiento o un programa informático para la codificación y decodificación como se describe en esta invención.

Referencias:

45

[0202]

[1] ISO/IEC 23003-3: 2012 - information Technology - MPEG Audio Technologies, Part 3: Unified Speech and Audio Coding;

50

[2] ISO/IEC 23003-1: 2007 - Information Technology - MPEG Audio Technologies, Part 1: Surround MPEG

REIVINDICACIONES

1. Un decodificador de audio (500; 600; 1300; 1600; 2000) para proporcionar al menos cuatro señales de canales con ancho de banda extendido (520, 522, 524, 526) sobre la base de una representación codificada (510; 5 610, 682; 1310, 1312),
 en el que el decodificador de audio está configurado para proporcionar una primera señal de mezcla descendente (532; 632; 1342) y una segunda señal de mezcla descendente (534; 634; 1344) sobre la base de una representación codificada conjuntamente (510; 610; 1310) de la primera señal de mezcla descendente y la segunda señal de mezcla descendente que utiliza una decodificación multicanal (530; 630; 1340);
- 10 en el que el decodificador de audio está configurado para proporcionar al menos una primera señal del canal de audio (542; 642; 1372) y una segunda señal del canal de audio (544; 644; 1374) sobre la base de la primera señal de mezcla descendente que utiliza una decodificación multicanal (540; 640; 1370);
 en el que el decodificador de audio está configurado para proporcionar al menos una tercera señal de canal de audio (556; 656; 1382) y una cuarta señal de canal de audio (558; 658; 1384) sobre la base de la segunda señal de mezcla
- 15 descendente que utiliza una decodificación multicanal (550; 650; 1380);
 en el que el decodificador de audio está configurado para realizar una primera extensión de ancho de banda multicanal conjunta (560; 660; 1390) sobre la base de la primera señal del canal de audio y la tercera señal del canal de audio con el fin de obtener una primera señal del canal con ancho de banda extendido (520; 620; 1320) y una tercera señal del canal con ancho de banda extendido (524; 624; 1324), en el que la extensión de ancho de banda multicanal utiliza
- 20 una relación entre la primera señal del canal de audio y la tercera señal del canal de audio; y
 en el que el decodificador de audio está configurado para realizar una segunda extensión de ancho de banda multicanal conjunta (570; 670; 1394) sobre la base de la segunda señal del canal de audio y la cuarta señal del canal de audio con el fin de obtener una segunda señal del canal con ancho de banda extendido (522; 622; 1322) y una cuarta señal del canal con ancho de banda extendido (526; 626; 1326).
- 25
2. El decodificador de audio según la reivindicación 1 en el que la primera señal de mezcla descendente y la segunda señal de mezcla descendente se asocian con diferentes posiciones horizontales o posiciones azimutales de una escena de audio.
- 30 3. El decodificador de audio según la reivindicación 1 o la reivindicación 2 en el que la primera señal de mezcla descendente está asociada con un lado izquierdo de una escena de audio y en el que la segunda señal de mezcla descendente está asociada con un lado derecho de la escena de audio.
4. El decodificador de audio según una de las reivindicaciones 1 a 3 en el que la primera señal del canal de audio y la segunda señal del canal de audio se asocian con posiciones vecinas verticalmente de una escena de audio y
- 35 en el que la tercera señal del canal de audio y la cuarta señal del canal de audio se asocian con posiciones vecinas verticalmente de la escena de audio.
- 40 5. El decodificador de audio según una de las reivindicaciones 1 a 4 en el que la primera señal del canal de audio y la tercera señal del canal de audio se asocian con un primer plano horizontal común o una primera elevación común de una escena de audio pero diferentes posiciones horizontales o posiciones azimutales de la escena de audio, en el que la segunda señal del canal de audio y la cuarta señal del canal de audio se asocian con un segundo plano horizontal común o una segunda elevación común de la escena de audio pero diferentes posiciones horizontales o
- 45 posiciones azimutales de la escena de audio,
 en el que el primer plano horizontal común o la primera elevación común es diferente del segundo plano horizontal común o la segunda elevación común.
6. El decodificador de audio según la reivindicación 5 en el que la primera señal del canal de audio y la
- 50 segunda señal del canal de audio se asocian con un primer plano vertical común o una primera posición azimutal común de la escena de audio pero diferentes posiciones o elevaciones verticales de la escena de audio, y
 en el que la tercera señal del canal de audio y la cuarta señal del canal de audio se asocian con un segundo plano vertical común o una segunda posición azimutal común de la escena de audio pero diferentes posiciones o elevaciones verticales de la escena de audio,
- 55 en el que el primer plano vertical común o primera posición azimutal es diferente del segundo plano vertical común o segunda posición azimutal.
7. El decodificador de audio según una de las reivindicaciones 1 a 6 en el que la primera señal del canal de audio y la segunda señal del canal de audio se asocian con un lado izquierdo de una escena de audio y

en el que la tercera señal del canal de audio y la cuarta señal del canal de audio se asocian con un lado derecho de la escena de audio.

8. El decodificador de audio según una de las reivindicaciones 1 a 7 en el que la primera señal del canal de audio y la tercera señal del canal de audio se asocian con una parte inferior de una escena de audio y en el que la segunda señal del canal de audio y la cuarta señal del canal de audio se asocian con una parte superior de la escena de audio.

9. El decodificador de audio según una de las reivindicaciones 1 a 8 en el que el decodificador de audio está configurado para realizar una división horizontal cuando se proporciona la primera señal de mezcla descendente y la segunda señal de mezcla descendente sobre la base de la representación codificada conjuntamente de la primera señal de mezcla descendente y la segunda señal de mezcla descendente que utiliza la decodificación multicanal.

10. El decodificador de audio según una de las reivindicaciones 1 a 9 en el que el decodificador de audio está configurado para realizar una división vertical cuando se proporciona al menos la primera señal del canal de audio y la segunda señal del canal de audio sobre la base de la primera señal de mezcla descendente que utiliza la decodificación multicanal; y en el que el decodificador de audio está configurado para realizar una división vertical cuando se proporciona al menos la tercera señal del canal de audio y la cuarta señal del canal de audio sobre la base de la segunda señal de mezcla descendente que utiliza la decodificación multicanal.

11. El decodificador de audio según una de las reivindicaciones 1 a 10 en el que el decodificador de audio está configurado para realizar una extensión de ancho de banda estéreo sobre la base de la primera señal del canal de audio y la tercera señal del canal de audio con el fin de obtener la primera señal del canal de ancho de banda extendida y la tercera señal del canal de ancho de banda extendida, en el que la primera señal del canal de audio y la tercera señal del canal de audio representan un primer par de canales izquierdo/derecho; y en el que el decodificador de audio está configurado para realizar una extensión de ancho de banda estéreo sobre la base de la segunda señal del canal de audio y la cuarta señal del canal de audio con el fin de obtener la segunda señal del canal con ancho de banda extendido y la tercera señal del canal con ancho de banda extendido, en el que la segunda señal del canal de audio y la cuarta señal del canal de audio representan un segundo par de canales izquierdo/derecho.

12. El decodificador de audio según una de las reivindicaciones 1 a 11 en el que el decodificador de audio está configurado para proporcionar la primera señal de mezcla descendente y la segunda señal de mezcla descendente sobre la base de una representación codificada conjuntamente de la primera señal de mezcla descendente y la segunda señal de mezcla descendente que utilizan una decodificación multicanal basada en predicciones.

13. El decodificador de audio según una de las reivindicaciones 1 a 12 en el que el decodificador de audio está configurado para proporcionar la primera señal de mezcla descendente y la segunda señal de mezcla descendente sobre la base de una representación codificada conjuntamente de la primera señal de mezcla descendente y la segunda señal de mezcla descendente que utilizan una decodificación multicanal asistida con señales residuales.

14. El decodificador de audio según una de las reivindicaciones 1 a 13 en el que el decodificador de audio está configurado para proporcionar al menos la primera señal del canal de audio y la segunda señal del canal de audio sobre la base de la primera señal de mezcla descendente que utiliza una decodificación multicanal basada en parámetros; en el que el decodificador de audio está configurado para proporcionar al menos la tercera señal del canal de audio y la cuarta señal del canal de audio sobre la base de la segunda señal de mezcla descendente que utiliza una decodificación multicanal basada en parámetros.

15. El decodificador de audio según la reivindicación 14 en el que la decodificación multicanal basada en parámetros está configurada para evaluar uno o varios parámetros que describen una correlación deseada entre dos canales y/o diferencias de niveles entre dos canales a fin de proporcionar las dos o más señales de canales de audio sobre la base de una respectiva señal de mezcla descendente.

16. El decodificador de audio según una de las reivindicaciones 1 a 15 en el que el decodificador de audio

está configurado para proporcionar al menos la primera señal del canal de audio y la segunda señal del canal de audio sobre la base de la primera señal de mezcla descendente que utiliza una decodificación multicanal asistida con señales residuales; y

5 en el que el decodificador de audio está configurado para proporcionar al menos la tercera señal del canal de audio y la cuarta señal del canal de audio sobre la base de la segunda señal de mezcla descendente que utiliza una decodificación multicanal asistida con señales residuales.

17. El decodificador de audio según una de las reivindicaciones 1 a 16 en el que el decodificador de audio está configurado para proporcionar una primera señal residual, que se utiliza para proporcionar al menos la primera
10 señal del canal de audio y la segunda señal del canal de audio y una segunda señal residual, que se utiliza para proporcionar al menos la tercera señal del canal de audio y la cuarta señal del canal de audio, sobre la base de una representación codificada conjuntamente de la primera señal residual y la segunda señal residual que utiliza una decodificación multicanal.

15 18. El decodificador de audio según la reivindicación 17 en el que la primera señal residual y la segunda señal residual se asocian con diferentes posiciones horizontales o posiciones azimutales de una escena de audio.

19. El decodificador de audio según la reivindicación 17 o la reivindicación 18 en el que la primera señal residual está asociada con un lado izquierdo de una escena de audio y en el que la segunda señal residual está
20 asociada con un lado derecho de la escena de audio.

20. Un codificador de audio (400; 1500; 2200) que proporciona una representación codificada (420; 1532; 2272, 2282) sobre la base de al menos cuatro señales de canales de audio (410, 412; 1512, 1514; 2212, 2222, 2214, 2224),

25 en el que el codificador de audio está configurado para obtener un primer grupo (2215) de parámetros de extensión de ancho de banda comunes sobre la base de una primera señal del canal de audio (410; 2212) y una tercera señal de canal de audio (414, 2214);

en el que el codificador de audio está configurado para obtener un segundo grupo (2225) de parámetros de extensión de ancho de banda comunes sobre la base de una segunda señal del canal de audio (412; 2222) y una cuarta señal
30 de canal de audio (416; 2224);

en el que el codificador de audio está configurado para codificar conjuntamente al menos la primera señal del canal de audio y la segunda señal del canal de audio utilizando una codificación multicanal (450; 2230) con el fin de obtener una primera señal de mezcla descendente (452; 2234);

35 en el que el codificador de audio está configurado para codificar conjuntamente al menos la tercera señal del canal de audio y la cuarta señal del canal de audio utilizando una codificación multicanal (460; 2240) con el fin de obtener una segunda señal de mezcla descendente (462; 2244); y

en el que el codificador de audio está configurado para codificar conjuntamente la primera señal de mezcla descendente y la segunda señal de mezcla descendente utilizando una codificación multicanal (470; 2250) con el fin de obtener una representación codificada de las señales de mezcla descendente.
40

21. El codificador de audio según la reivindicación 20 en el que la primera señal de mezcla descendente y la segunda señal de mezcla descendente se asocian con diferentes posiciones horizontales o posiciones azimutales de una escena de audio.

45 22. El codificador de audio según una de las reivindicaciones 20 o 21 en el que la primera señal de mezcla descendente está asociada con un lado izquierdo de una escena de audio y en el que la segunda señal de mezcla descendente está asociada con un lado derecho de la escena de audio.

23. El codificador de audio según una de las reivindicaciones 20 a 22 en el que la primera señal del canal de audio y la segunda señal del canal de audio se asocian con posiciones vecinas verticalmente de una escena de
50 audio y en el que la tercera señal del canal de audio y la cuarta señal del canal de audio se asocian con posiciones vecinas verticalmente de la escena de audio.

24. El codificador de audio según una de las reivindicaciones 20 a 23 en el que la primera señal del canal de audio y la tercera señal del canal de audio se asocian con un primer plano horizontal común o una primera elevación
55 de una escena de audio pero diferentes posiciones horizontales o posiciones azimutales de la escena de audio, en el que la segunda señal del canal de audio y la cuarta señal del canal de audio se asocian con un segundo plano horizontal común o una segunda elevación de la escena de audio pero diferentes posiciones horizontales o posiciones azimutales de la escena de audio,

en el que el primer plano horizontal común o la primera elevación es diferente del segundo plano horizontal común o la segunda elevación.

25. El codificador de audio según la reivindicación 24 en el que la primera señal del canal de audio y la segunda señal del canal de audio se asocian con un primer plano vertical común o a primera posición azimutal de la escena de audio pero diferentes posiciones o elevaciones verticales de la escena de audio, y en el que la tercera señal del canal de audio y la cuarta señal del canal de audio se asocian con un segundo plano vertical común o una segunda posición azimutal de la escena de audio pero diferentes posiciones o elevaciones verticales de la escena de audio, en el que el primer plano vertical común o la primera posición azimutal es diferente del segundo plano vertical común o la segunda posición azimutal.

26. El codificador de audio según una de las reivindicaciones 20 a 25 en el que la primera señal del canal de audio y la segunda señal del canal de audio se asocian con un lado izquierdo de una escena de audio y en el que la tercera señal del canal de audio y la cuarta señal del canal de audio se asocian con un lado derecho de la escena de audio.

27. El codificador de audio según una de las reivindicaciones 20 a 26 en el que la primera señal del canal de audio y la tercera señal del canal de audio se asocian con una parte inferior de una escena de audio y en el que la segunda señal del canal de audio y la cuarta señal del canal de audio se asocian con una parte superior de la escena de audio.

28. El codificador de audio según una de las reivindicaciones 20 a 27 en el que el codificador de audio está configurado para realizar una combinación horizontal cuando se proporciona la representación codificada de las señales de mezcla descendente sobre la base de la primera señal de mezcla descendente y la segunda señal de mezcla descendente que utilizan la codificación multicanal.

29. El codificador de audio según una de las reivindicaciones 20 a 28 en el que el decodificador de audio está configurado para realizar una combinación vertical cuando se proporciona la primera señal de mezcla descendente sobre la base de la primera señal del canal de audio y la segunda señal del canal de audio que utilizan la codificación multicanal; y en el que el codificador de audio está configurado para realizar una combinación vertical cuando se proporciona la segunda señal de mezcla descendente sobre la base de la tercera señal del canal de audio y la cuarta señal del canal de audio que utilizan la codificación multicanal.

32. El codificador de audio según una de las reivindicaciones 20 a 31 en el que el codificador de audio está configurado para proporcionar la primera señal de mezcla descendente sobre la base de la primera señal del canal de audio y la segunda señal del canal de audio que utilizan una codificación multicanal basada en parámetros; y en el que el codificador de audio está configurado para proporcionar la segunda señal de mezcla descendente sobre la base de la tercera señal del canal de audio y la cuarta señal del canal de audio que utilizan una codificación multicanal basada en parámetros.

33. El codificador de audio según la reivindicación 32 en el que la codificación multicanal basada en parámetros está configurada para proporcionar uno o varios parámetros que describen una correlación deseada entre dos canales y/o diferencias de niveles entre dos canales.

34. El codificador de audio según una de las reivindicaciones 20 a 33 en el que el codificador de audio está configurado para proporcionar la primera señal de mezcla descendente sobre la base de la primera señal del canal de audio y la segunda señal del canal de audio que utilizan una codificación multicanal asistida con señales residuales; y en el que el codificador de audio está configurado para proporcionar la segunda señal de mezcla descendente sobre la base de la tercera señal del canal de audio y la cuarta señal del canal de audio que utilizan una codificación multicanal asistida con señales residuales.

35. El codificador de audio según una de las reivindicaciones 20 a 34 en el que el codificador de audio está configurado para proporcionar una representación codificada conjuntamente de una primera señal residual, que se obtiene cuando se codifica conjuntamente al menos la primera señal del canal de audio y la segunda señal del canal de audio y de una segunda residual, que se obtiene cuando se codifica conjuntamente al menos la tercera señal del canal de audio y la cuarta señal del canal de audio que utilizan una codificación multicanal.

36. El codificador de audio según la reivindicación 35 en el que la primera señal residual y la segunda señal residual se asocian con diferentes posiciones horizontales o posiciones azimutales de una escena de audio.

37. El codificador de audio según la reivindicación 35 o la reivindicación 36 en el que la primera señal residual está asociada con un lado izquierdo de una escena de audio y en el que la segunda señal residual está asociada con un lado derecho de la escena de audio.

38. Un procedimiento (1000) para proporcionar al menos cuatro señales de canales de audio sobre la base de una representación codificada en el que el procedimiento comprende:

10 proporcionar (1010) una primera señal de mezcla descendente y una segunda señal de mezcla descendente sobre la base de una representación codificada conjuntamente de la primera señal de mezcla descendente y la segunda señal de mezcla descendente que utilizan una decodificación multicanal;

proporcionar (1020) al menos una primera señal del canal de audio y una segunda señal del canal de audio sobre la base de la primera señal de mezcla descendente que utiliza una decodificación multicanal;

15 proporcionar (1030) al menos una tercera señal de canal de audio y una cuarta señal de canal de audio sobre la base de la segunda señal de mezcla descendente que utiliza una decodificación multicanal;

realizar (1040) una primera extensión de ancho de banda multicanal conjunta sobre la base de la primera señal del canal de audio y la tercera señal del canal de audio con el fin de obtener una primera señal del canal con ancho de banda extendido y una tercera señal del canal con ancho de banda extendido, en la que la extensión de ancho de banda multicanal utiliza un relación entre la primera señal del canal de audio y la tercera señal del canal de audio; y

realizar (1050) una segunda extensión de ancho de banda multicanal conjunta sobre la base de la segunda señal del canal de audio y la cuarta señal del canal de audio con el fin de obtener la segunda señal del canal con ancho de banda extendido y la tercera señal del canal con ancho de banda extendido.

25

39. Un procedimiento (900) para proporcionar una representación codificada sobre la base de al menos cuatro señales de canales de audio, el procedimiento que comprende:

la obtención (920) de un primer grupo de parámetros de extensión de ancho de banda comunes sobre la base de una primera señal del canal de audio y una tercera señal de canal de audio;

30 la obtención (930) un segundo grupo de parámetros de extensión de ancho de banda comunes sobre la base de una segunda señal del canal de audio y una cuarta señal de canal de audio.

la codificación conjunta (930) de al menos la primera señal del canal de audio y la segunda señal del canal de audio que utilizan una codificación multicanal con el fin de obtener una primera señal de mezcla descendente;

35 la codificación conjunta (940) de al menos la tercera señal del canal de audio y la cuarta señal del canal de audio que utilizan una codificación multicanal con el fin de obtener una segunda señal de mezcla descendente; y

la codificación conjunta (950) de la primera señal de mezcla descendente y la segunda señal de mezcla descendente que utilizan una codificación multicanal con el fin de obtener una representación codificada de las señales de mezcla descendente.

40 40. Un programa informático configurado para realizar el procedimiento según la reivindicación 38 o 39 cuando el programa informático se ejecuta en un ordenador.

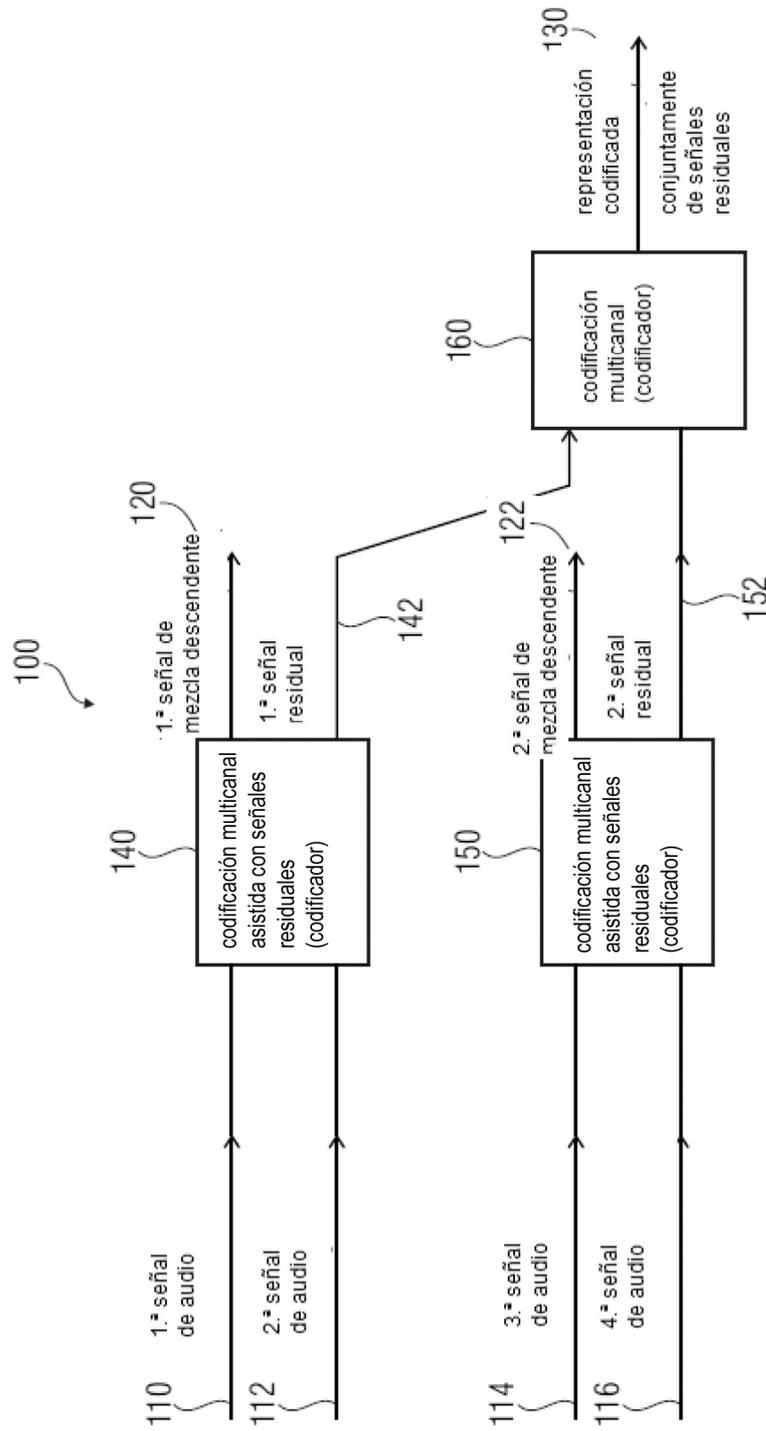


FIG. 1

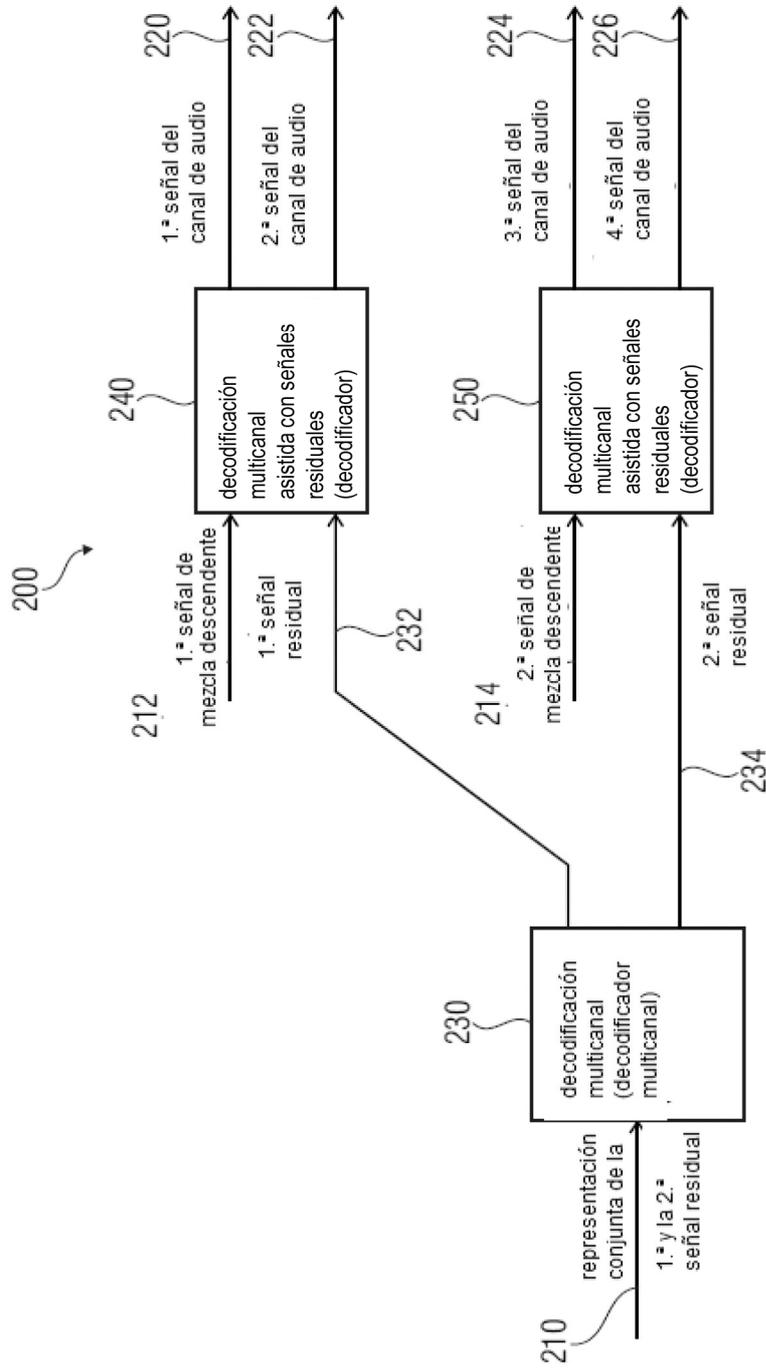


FIG. 2

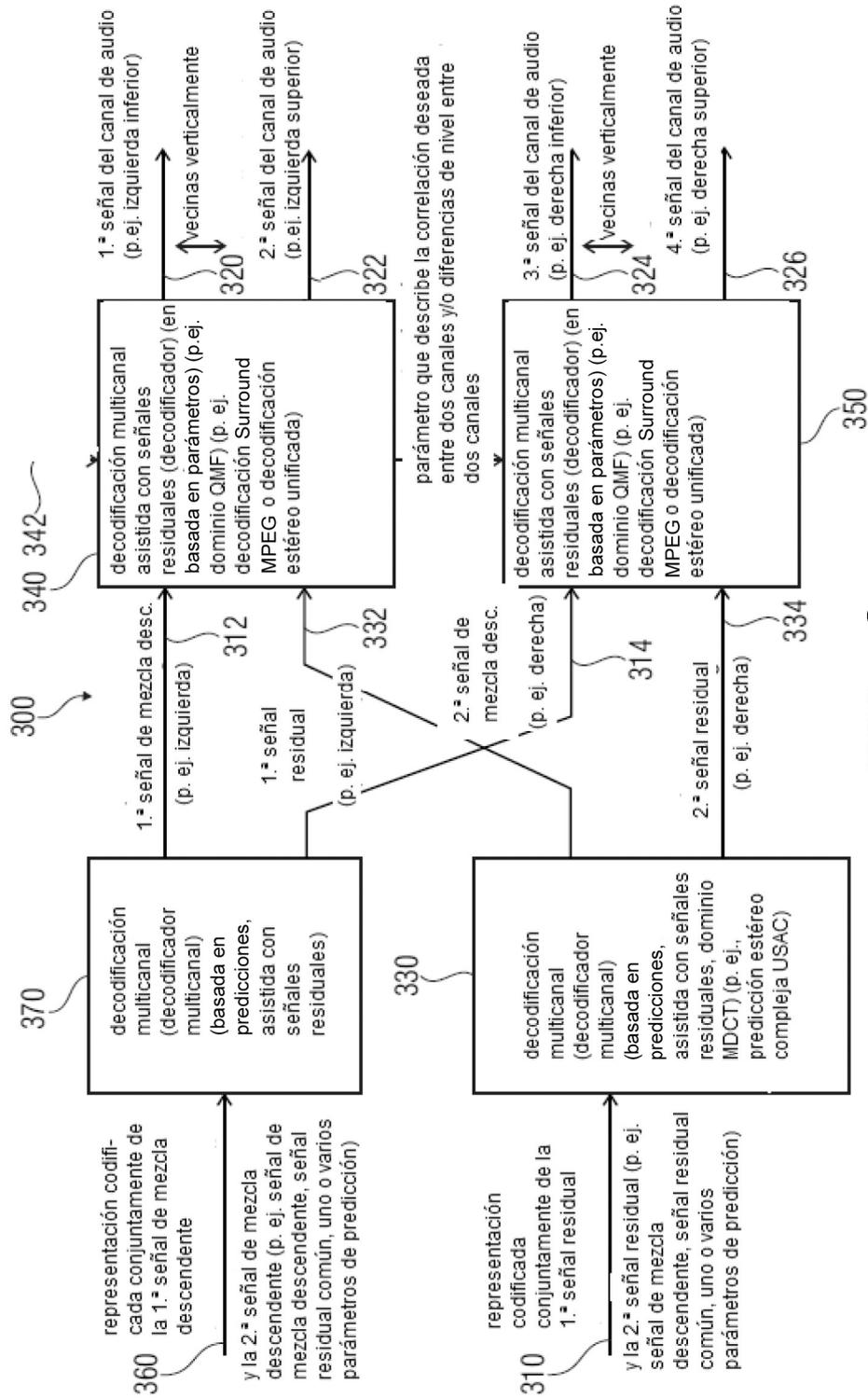


FIG. 3

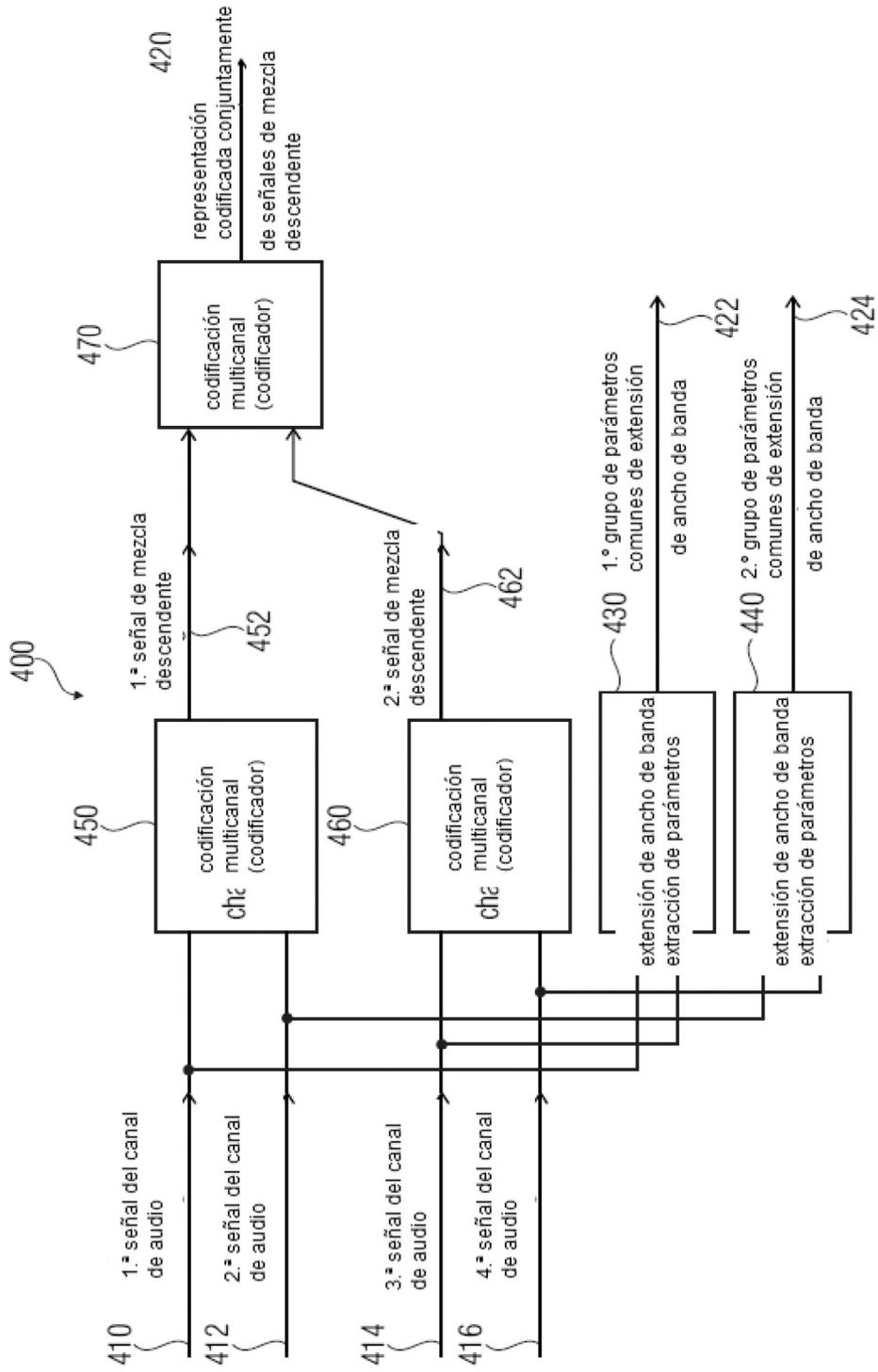


FIG. 4

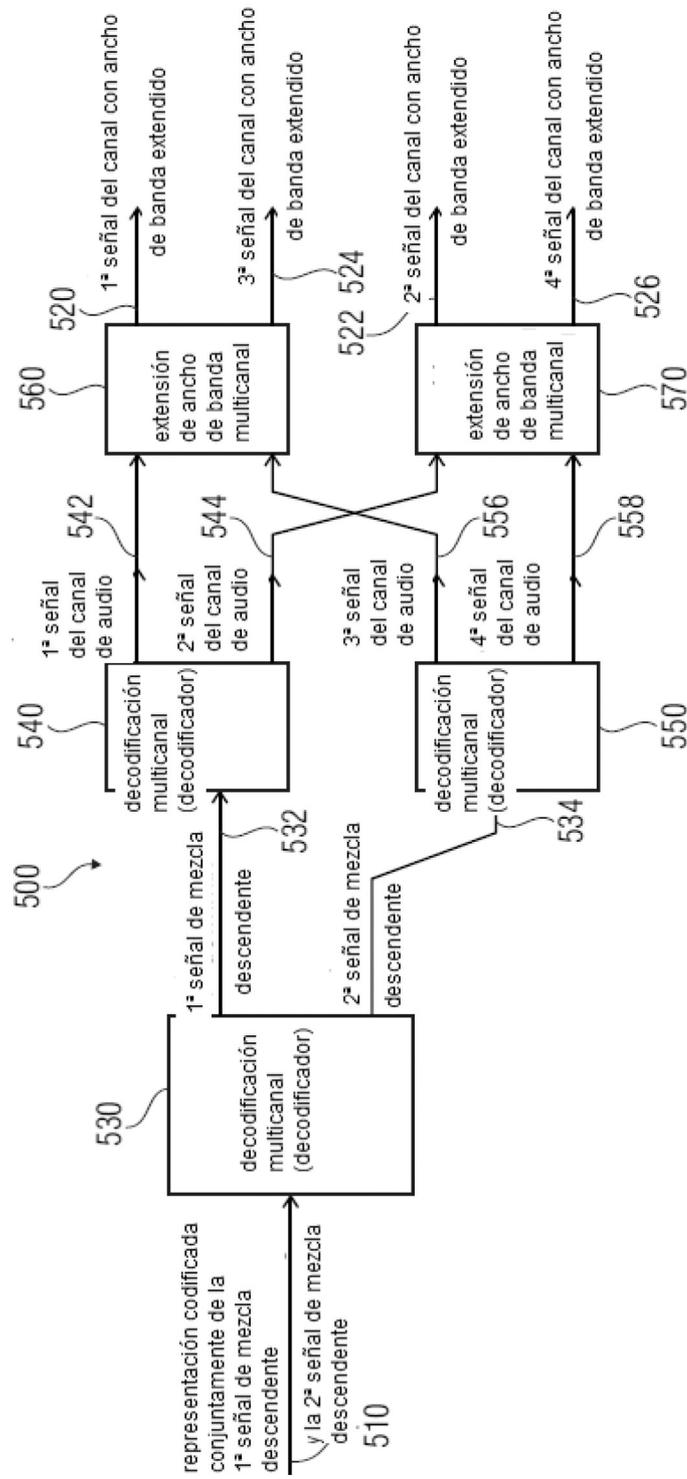


FIG. 5

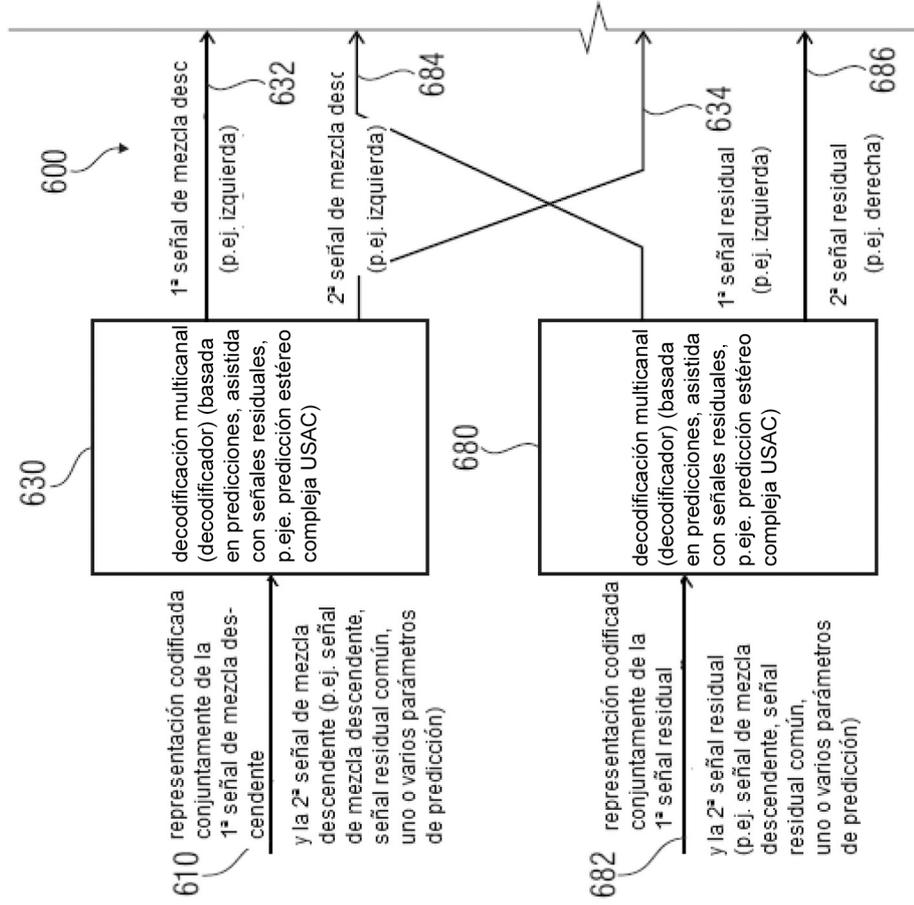


FIG. 6A

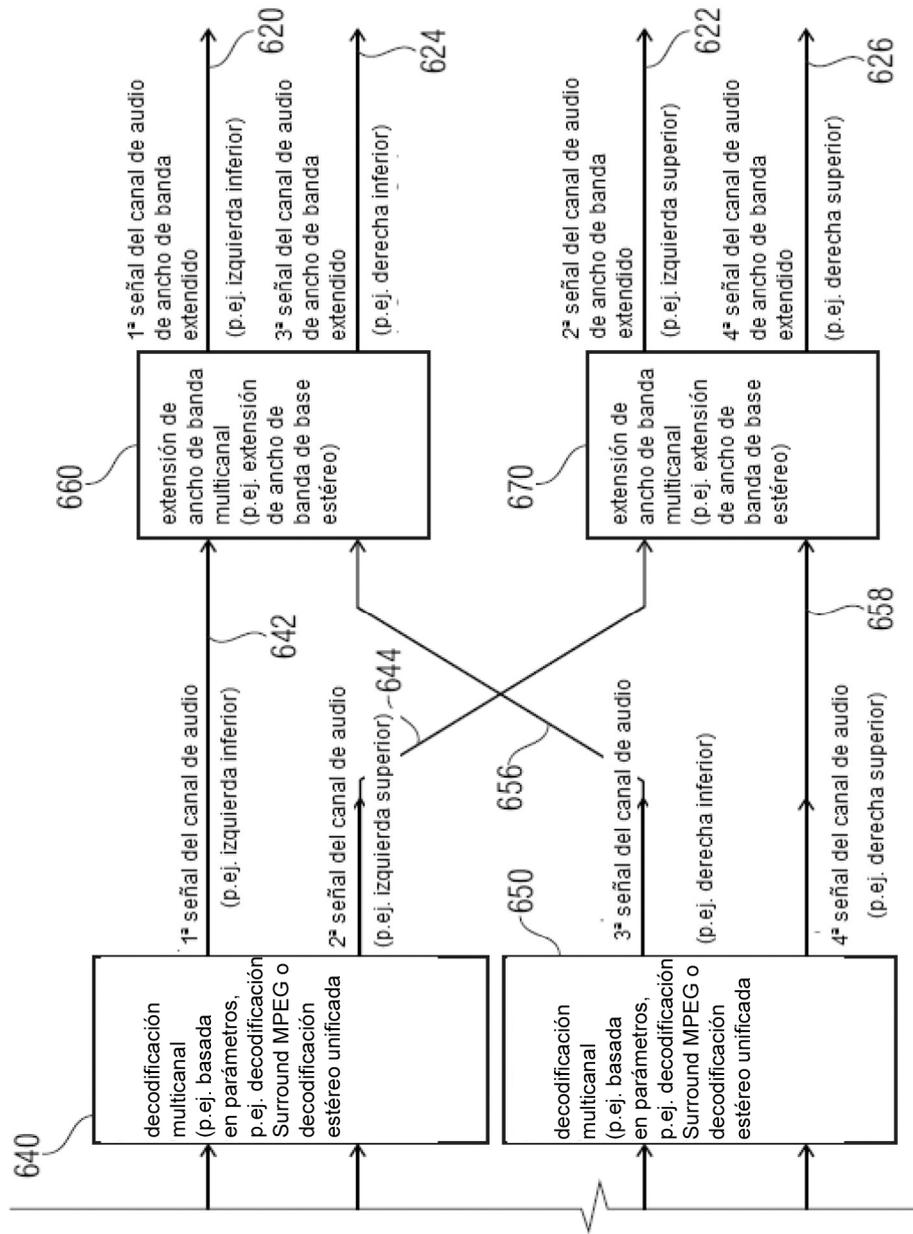


FIG. 6B

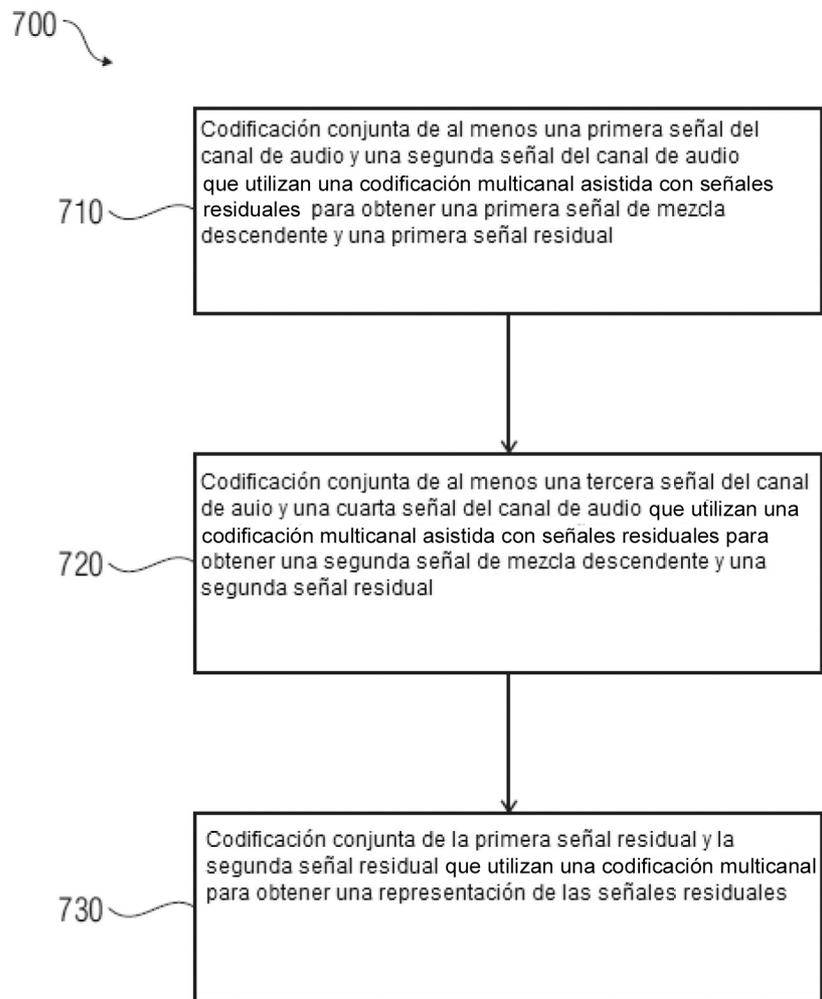


FIG. 7

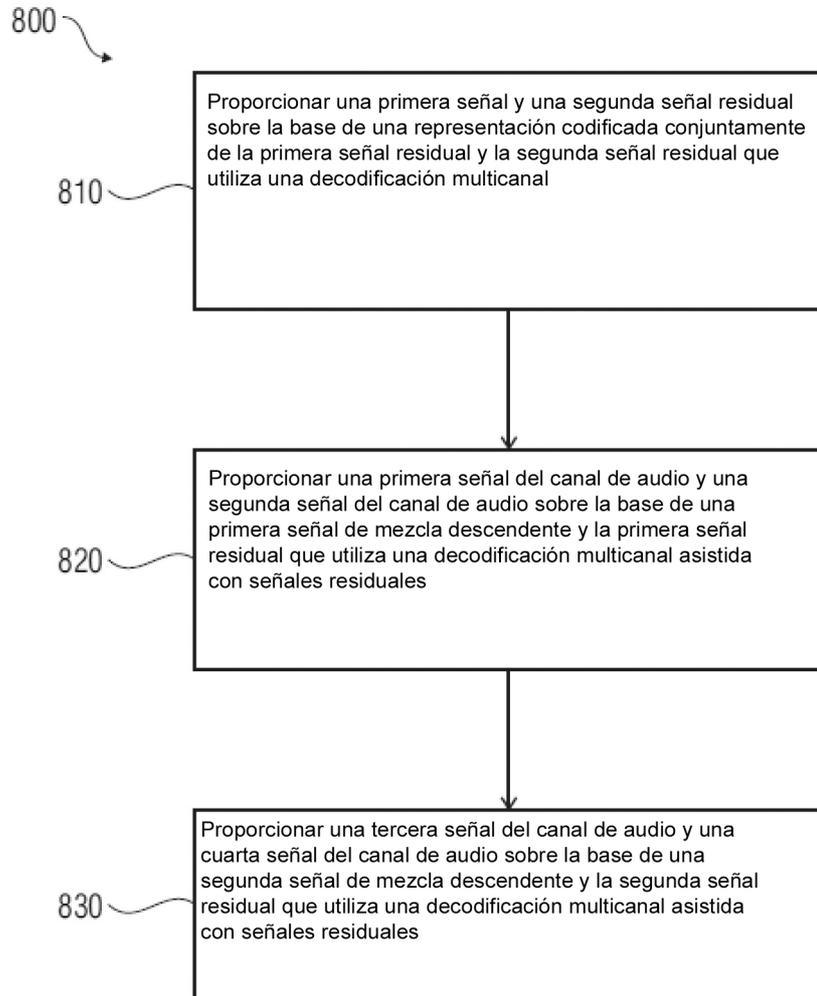


FIG. 8

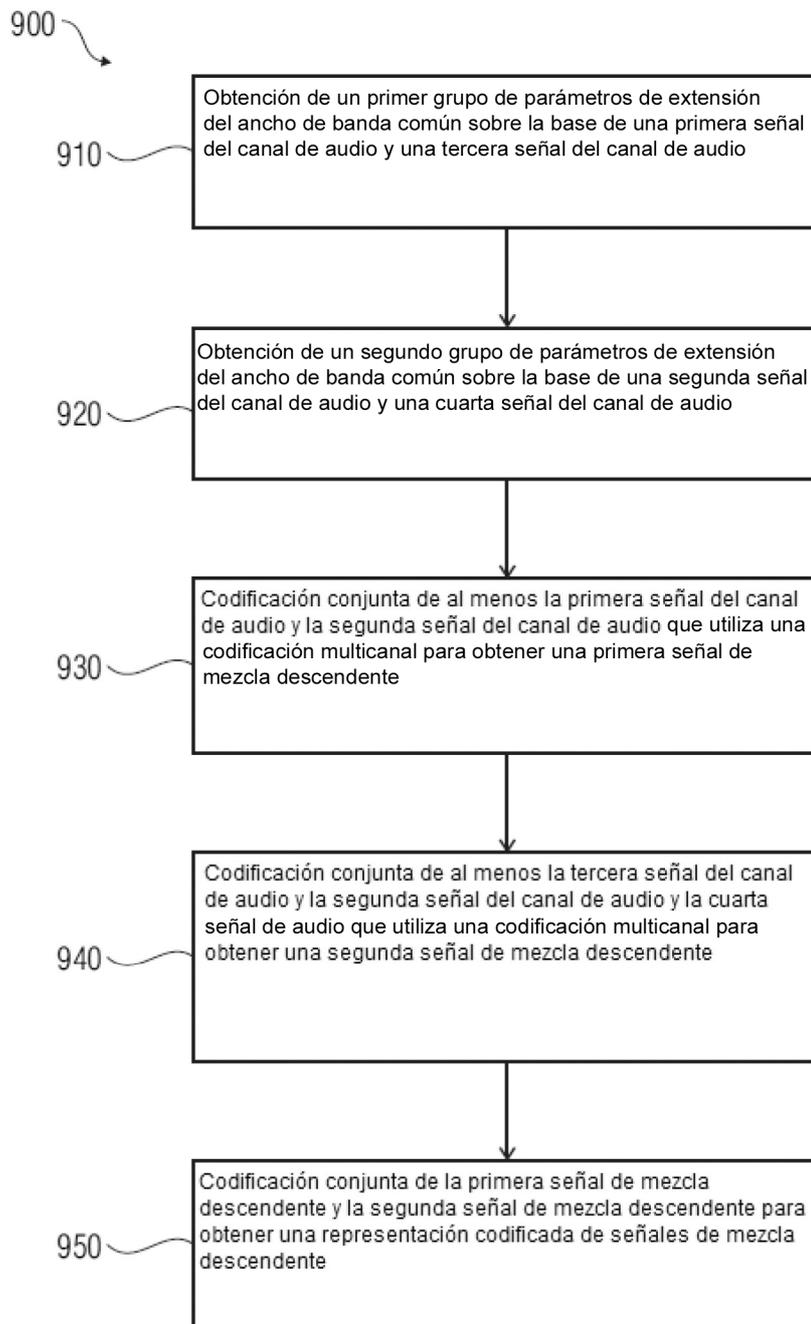


FIG. 9

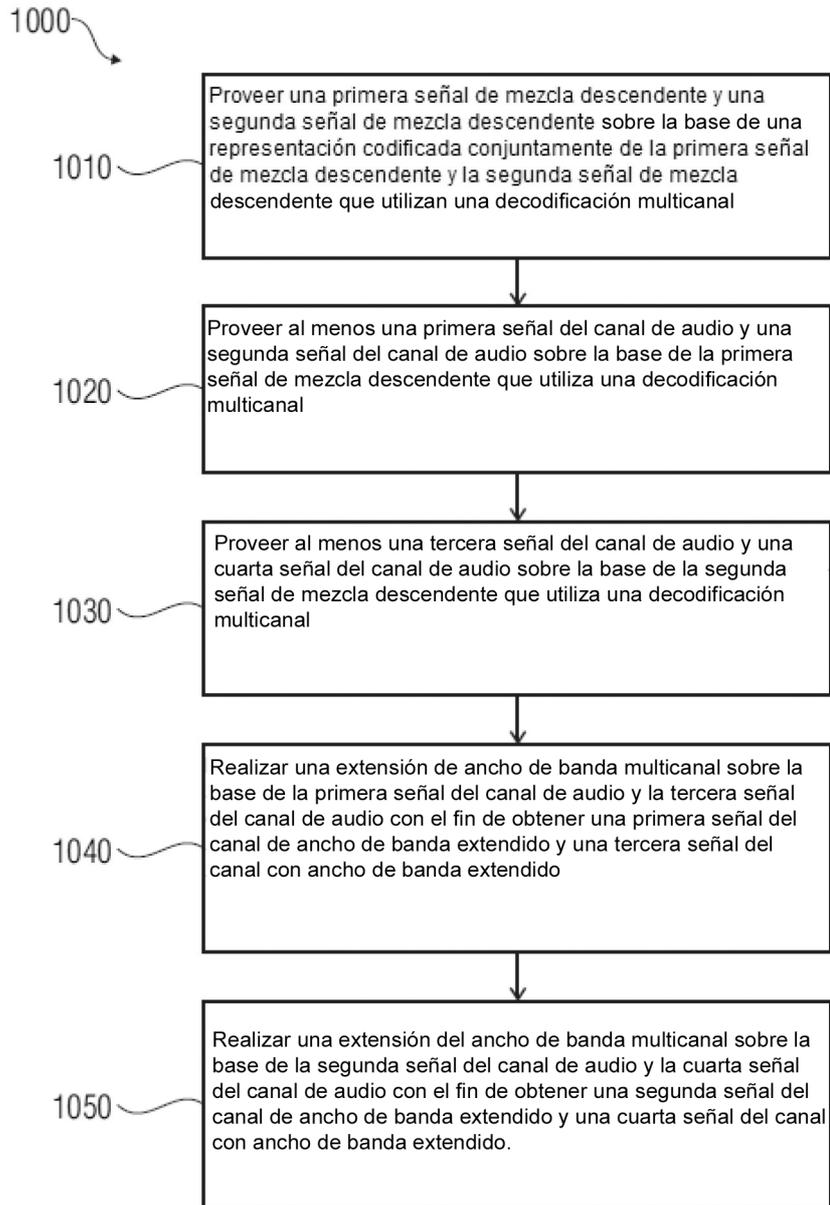


FIG. 10

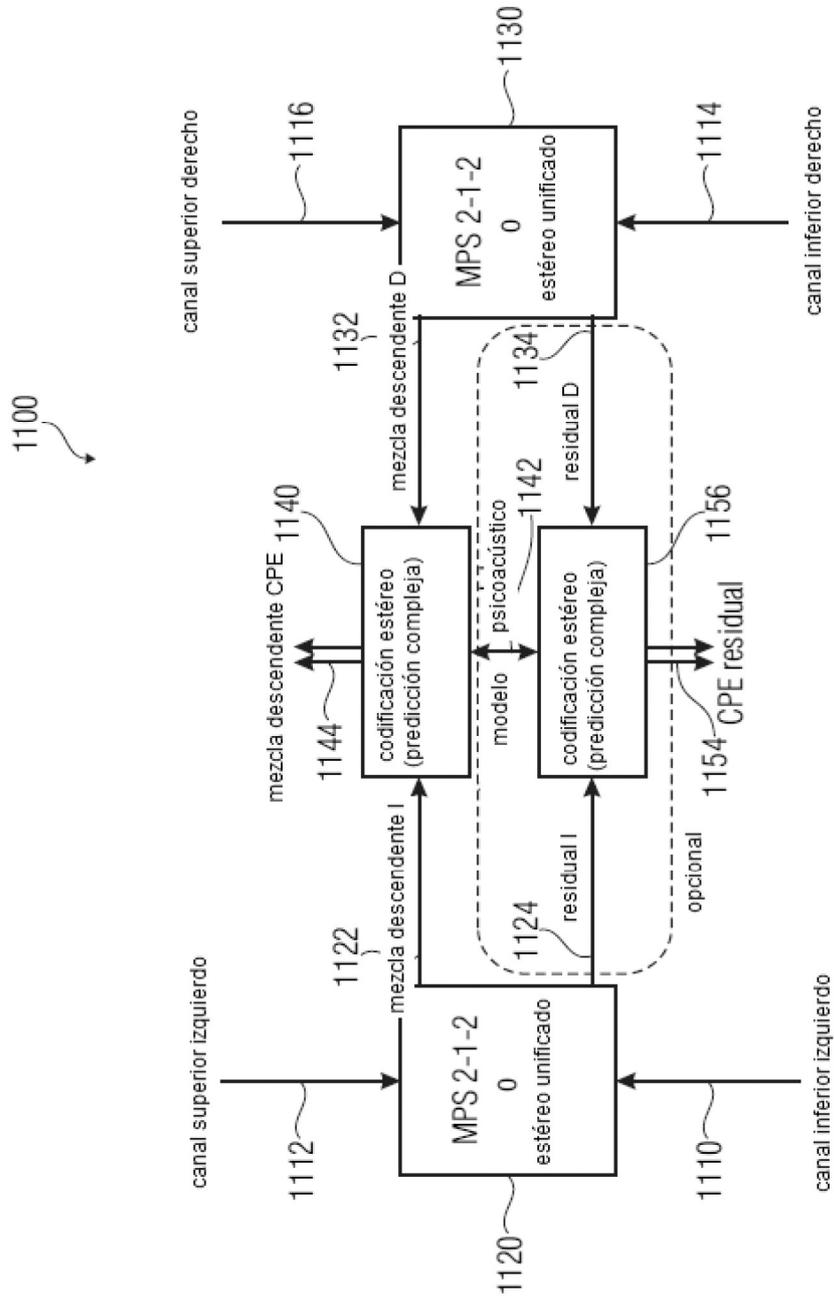


FIG. 11

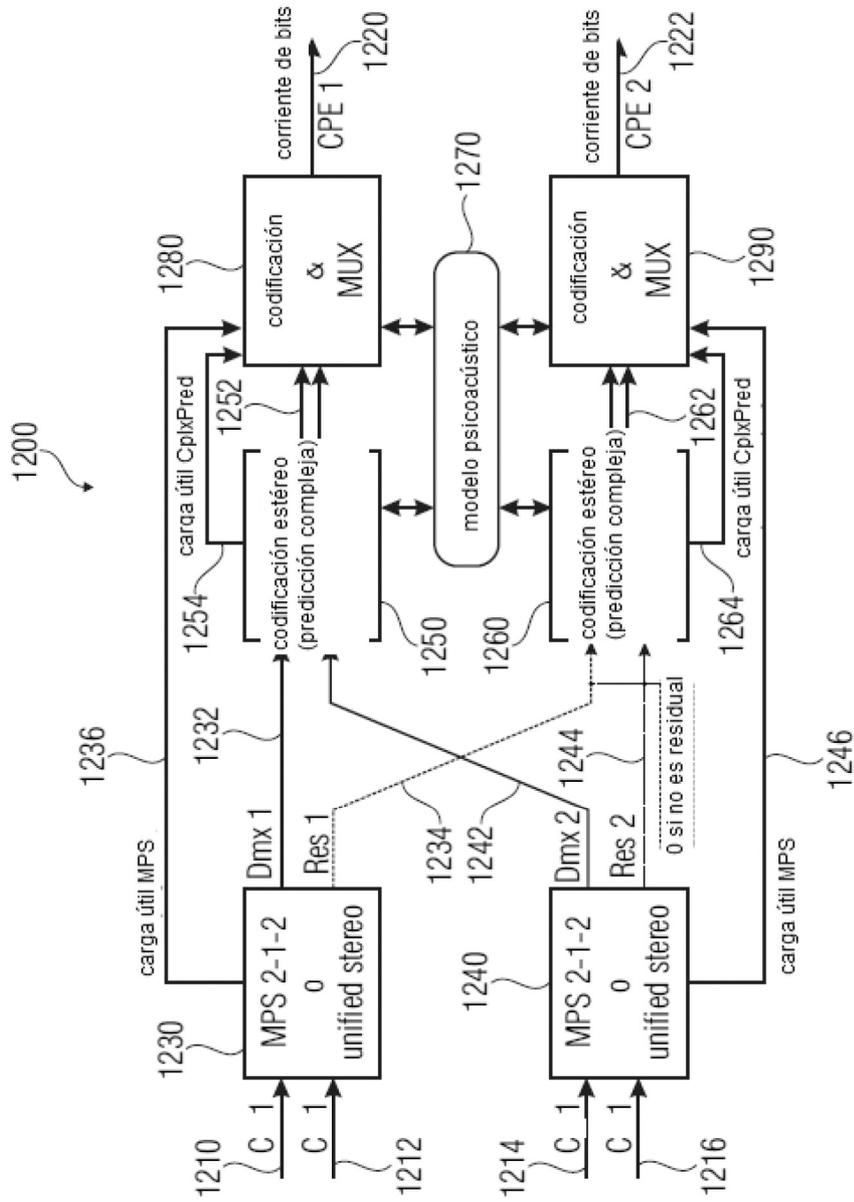


FIG. 12

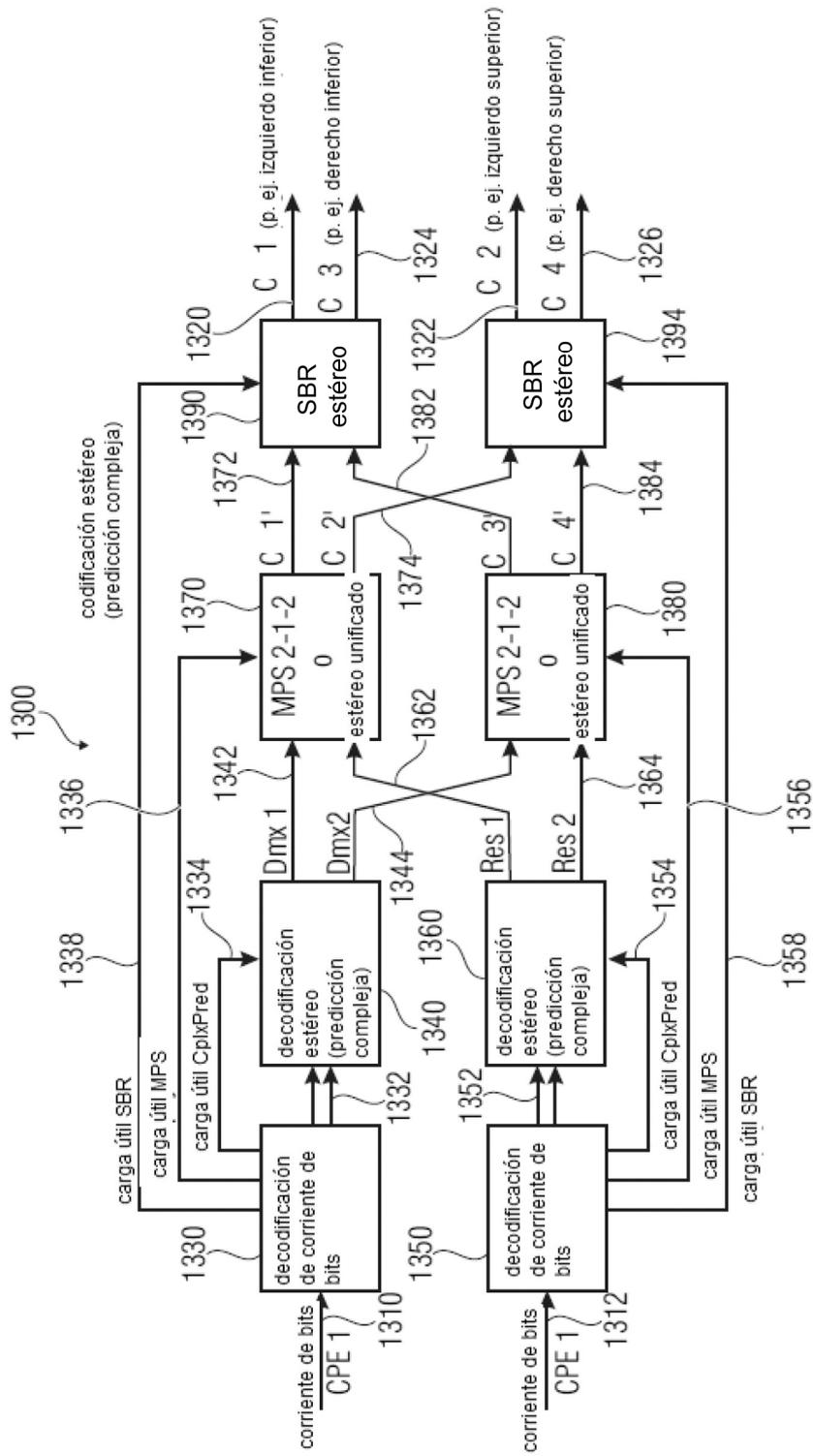


FIG. 13

```

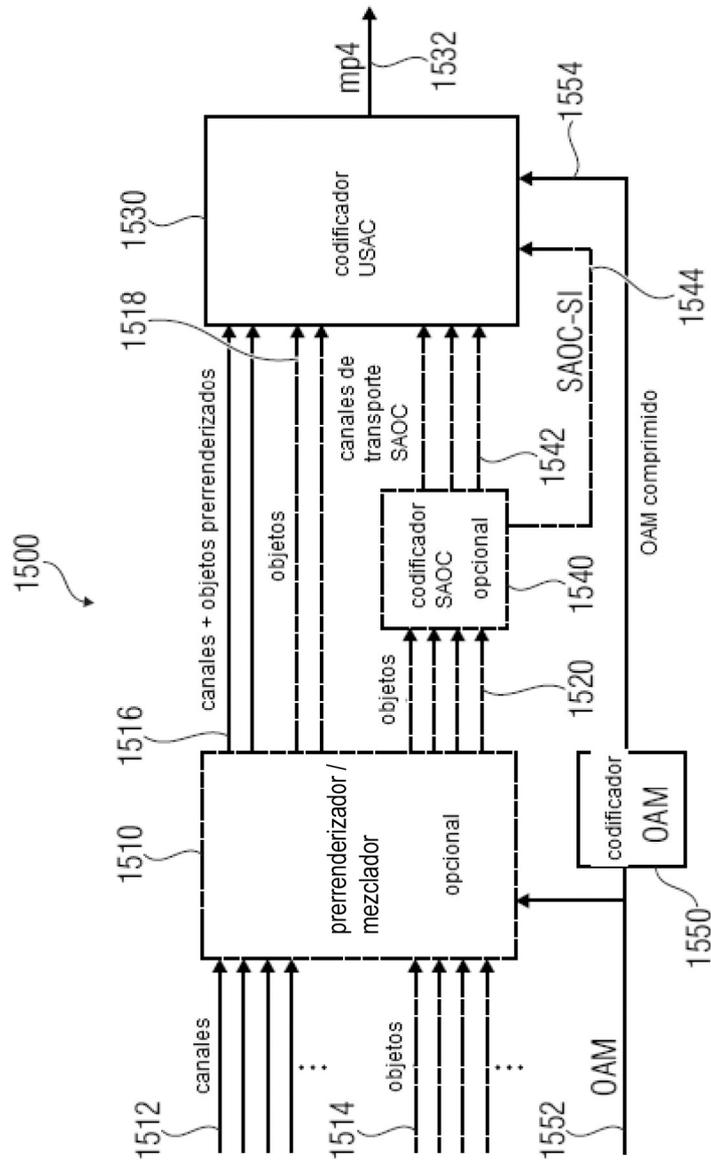
UsacChannelPairElementConfig (sbrRatioIndex)
{
    UsacCoreConfig ();
    if (sbrRatioIndex > 0) {
        SbrConfig ();
        stereoConfigIndex;                2        uimsbf
    } else {
        stereoConfigIndex = 0;
    }
    if (stereoConfigIndex > 0) {
        Mps212Config(stereoConfigIndex);
    }
+ qcelIndex                            2        uimsbf
}

```

FIG. 14A

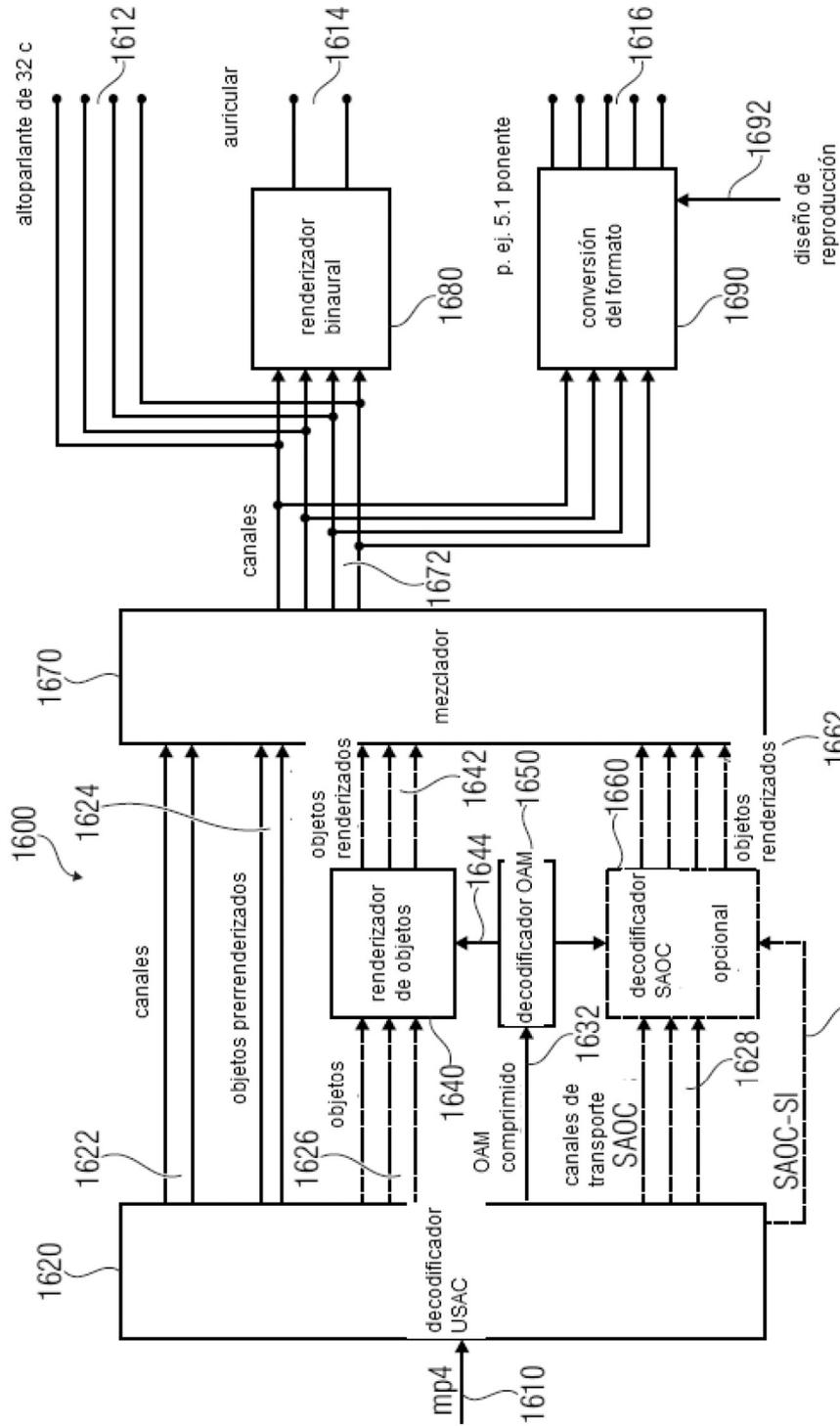
qcelIndex	Significado
0	CPE estéreo
1	QCE sin residuo
2	QCE con residuo
3	-reservado-

FIG. 14B



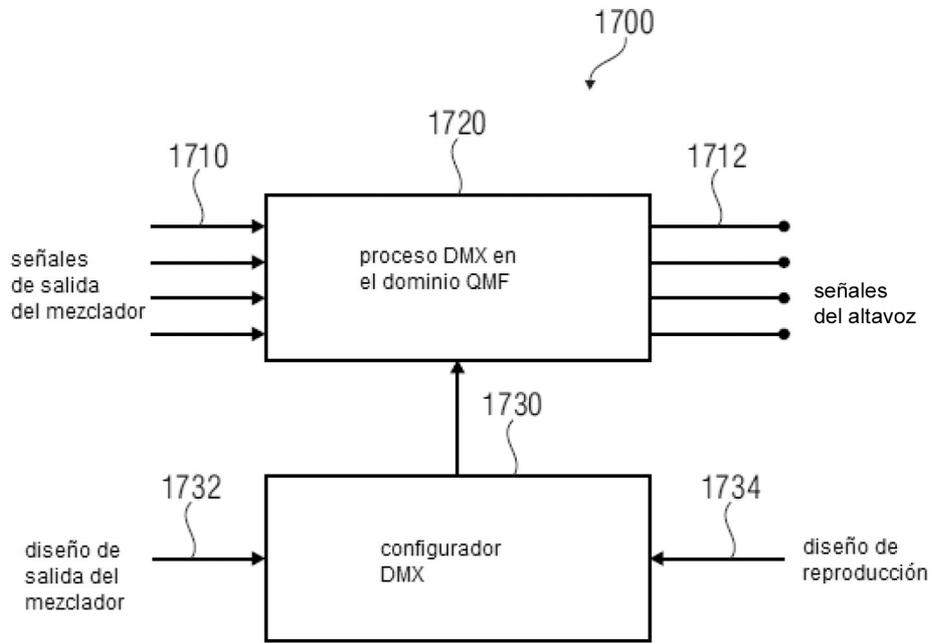
RESEÑA DEL CODIFICADOR DE AUDIO 3D

FIG. 15



RESEÑA DEL DECODIFICADOR DE AUDIO 3D

FIG. 16



ESTRUCTURA DEL CONVERTIDOR DE FORMATO

FIG. 17

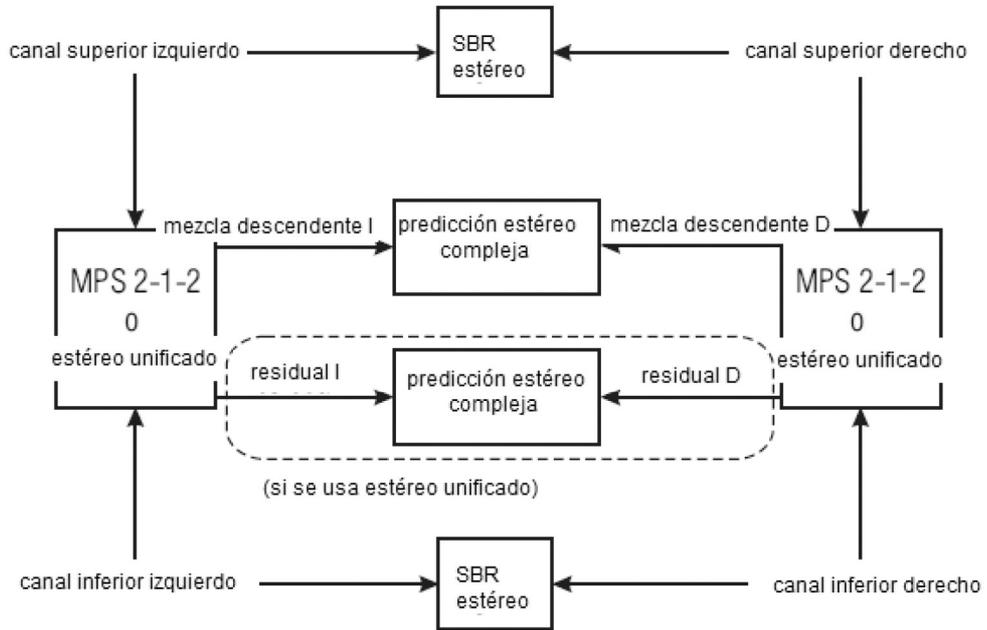


FIG. 18

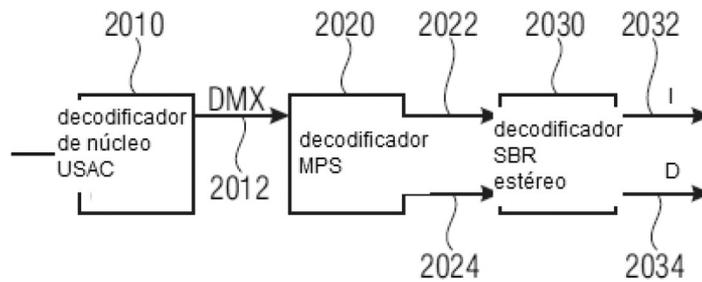
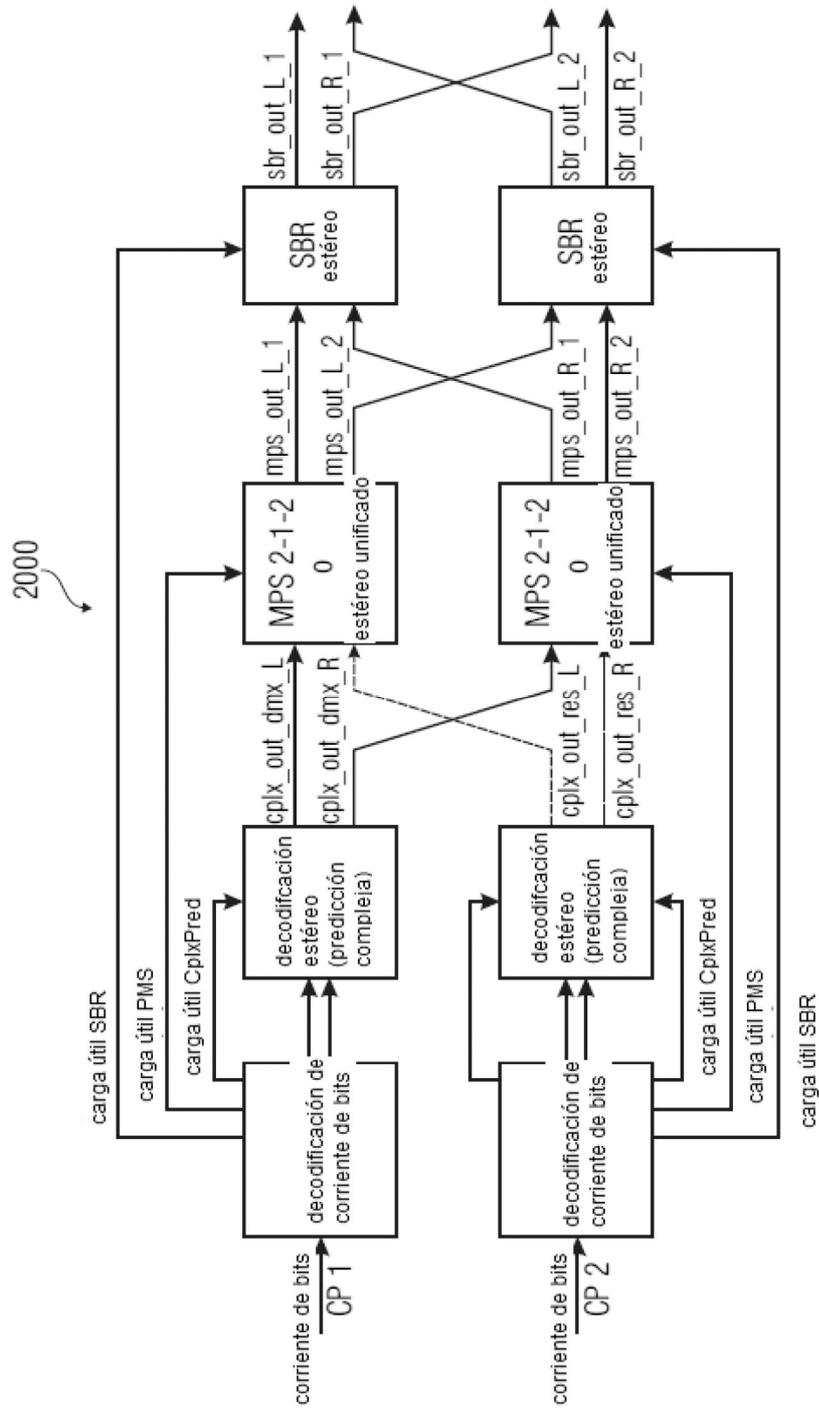


FIG. 19



ESQUEMAS DEL DECODIFICADOR QCE

FIG. 20

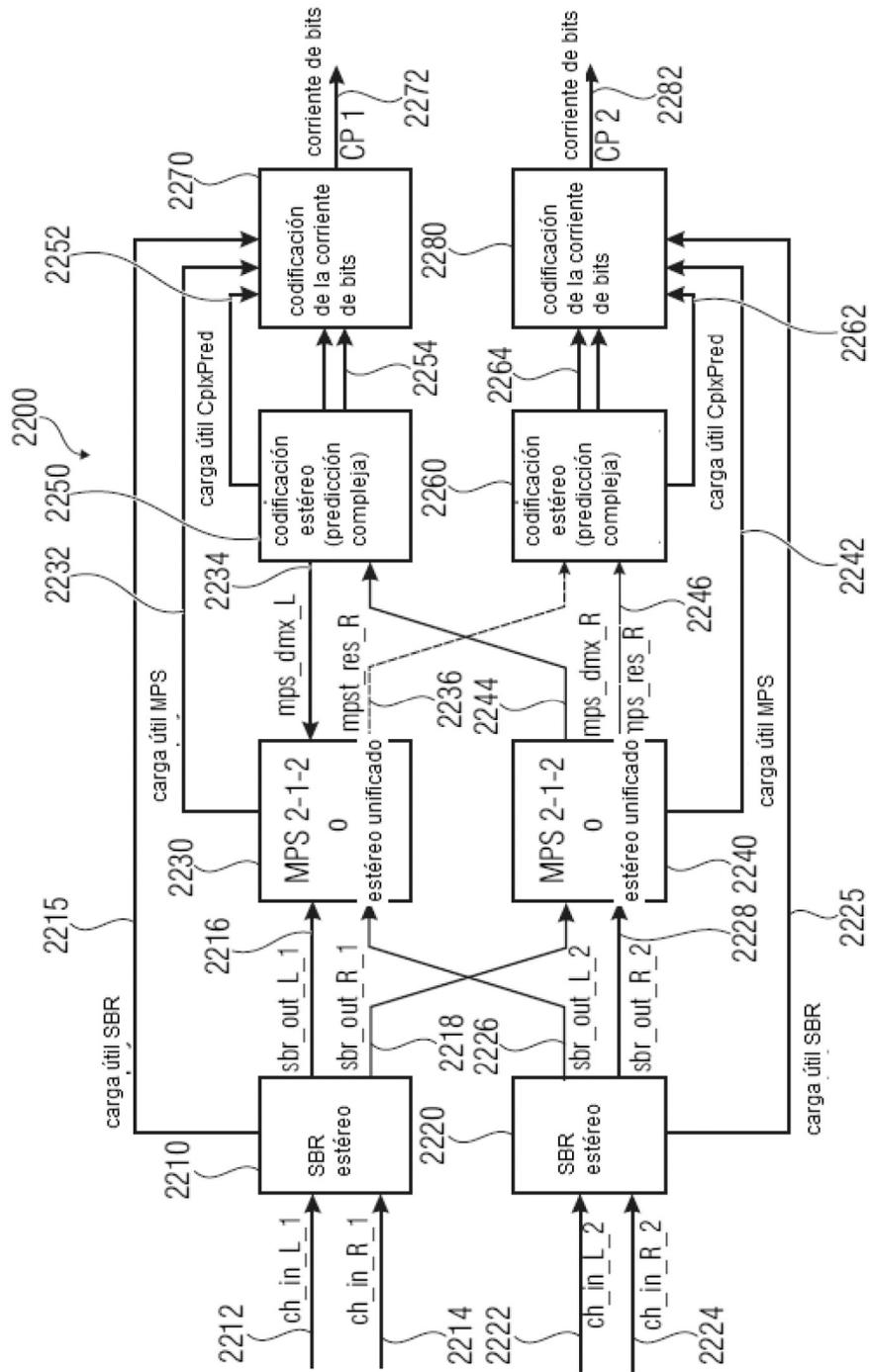


FIG. 21

ESQUEMAS DEL CODIFICADOR DEL CANAL QUAD