

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 650 544**

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: **11.07.2014 PCT/EP2014/064915**

87 Fecha y número de publicación internacional: **29.01.2015 WO15010926**

96 Fecha de presentación y número de la solicitud europea: **11.07.2014 E 14739141 (1)**

97 Fecha y número de publicación de la concesión europea: **06.09.2017 EP 3022735**

54 Título: **Codificador de audio, decodificador de audio, procedimientos y programa informático que utiliza señales residuales codificadas conjuntamente**

30 Prioridad:

22.07.2013 EP 13177376

18.10.2013 EP 13189305

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

19.01.2018

73 Titular/es:

**FRAUNHOFER GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.0%)**

**Hansastraße 27 c
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 2 650 544 T3

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

DESCRIPCIÓN

Codificador de audio, decodificador de audio, procedimientos y programa informático que utiliza señales residuales codificadas conjuntamente

5

Campo técnico

[0001] Las realizaciones según la invención se refieren a un decodificador de audio para proporcionar al menos cuatro señales de canal de audio sobre la base de una representación codificada.

10

[0002] Otras realizaciones según la invención se refieren a un codificador de audio para proporcionar una representación codificada sobre la base de al menos cuatro señales de canal de audio.

[0003] Otras realizaciones según la invención se refieren a un procedimiento para proporcionar al menos cuatro señales de canal de audio sobre la base de una representación codificada y a un procedimiento para proporcionar una representación codificada sobre la base de al menos cuatro señales de canal de audio.

[0004] Otras realizaciones según la invención se refieren a un programa informático para llevar a cabo uno de dichos procedimientos.

20

[0005] Generalmente hablando, las realizaciones según la invención se refieren a una codificación conjunta de n canales.

Antecedentes de la invención

25

[0006] En los años recientes, se ha ido incrementando de modo constante una demanda por el almacenamiento y la transmisión de contenidos de audio. Además, los requisitos de calidad para el almacenamiento y la transmisión de contenidos de audio también se han 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. Además, se han creado algunas extensiones espaciales como, por ejemplo, el así llamado concepto de "MPEG Surround" que se describe, por ejemplo, en el estándar internacional ISO/IEC 23003-1:2007. Además, 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 del objeto de audio espacial (SAOC).

30

[0007] Además, 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 y para manipular señales de audio multicanal, se define en el estándar internacional ISO/IEC 23003-3:2012, que describe el así llamado concepto de "codificación de voz y audio unificados" (USAC).

40

[0008] En MPEG USAC [1], la codificación estéreo conjunta de dos canales se lleva a cabo mediante el uso de predicción compleja, MPS 2-1-1 o estéreo unificado con señales residuales de banda limitada o de banda completa. MPEG surround [2] combina jerárquicamente las cajas OTT y TTT para codificación conjunta de audio multicanal con o sin transmisión de señales residuales.

45

[0009] La codificación y decodificación de audio multicanal se describen por ejemplo también en el documento EP2194526 A1. Sin embargo, hay un deseo de proporcionar un concepto incluso más avanzado para una codificación y decodificación eficaz de escenas de audio tridimensionales.

50

Resumen de la invención

[0010] Una realización según la invención crea un decodificador de audio para proporcionar al menos cuatro señales de canal de audio sobre la base de una representación codificada. El decodificador de audio se configura para proporcionar una primera señal residual y una segunda señal residual sobre la base de una representación codificada conjuntamente de la primera señal residual y de la segunda señal residual mediante el uso de una decodificación multicanal que explota las similitudes y/o dependencias entre las señales residuales. El decodificador de audio también se configura para proporcionar 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 mediante el uso de una decodificación multicanal asistida por señales residuales. El decodificador de audio también se configura para proporcionar una tercera señal de canal de audio y una cuarta señal de canal de audio sobre la base de una

55

60

segunda señal de mezcla descendente y la segunda señal residual mediante el uso de una decodificación multicanal asistida por señales residuales.

[0011] Esta realización según la invención se basa en el hallazgo de que las dependencias entre cuatro o incluso más señales de canal de audio se pueden explotar derivando dos señales residuales, cada una de las cuales se usa para proporcionar dos o más señales de canal de audio mediante el uso de una decodificación multicanal asistida por señales residuales, de una representación codificada conjuntamente de las señales residuales. En otras palabras, se ha encontrado que hay típicamente algunas similitudes de dichas señales residuales, de tal forma que una tasa de bits para codificar dichas señales residuales, que ayudan a mejorar una calidad de audio cuando se decodifican las al menos cuatro señales de canal de audio, se puede reducir derivando las dos señales residuales de una representación codificada conjuntamente mediante el uso de una decodificación multicanal, que explota similitudes y/o dependencias entre las señales residuales.

[0012] En una realización preferida, el decodificador de audio se configura 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 mediante el uso de una decodificación multicanal. Conforme a ello, se crea una estructura jerárquica de un decodificador de audio, en la que tanto las señales de mezcla descendente como las señales residuales, que se usan en la decodificación multicanal asistida por señales residuales para proporcionar las al menos cuatro señales de canal de audio, se derivan mediante el uso de decodificación multicanal separada. Tal concepto es particularmente eficiente, dado que las dos señales de mezcla descendente comprenden típicamente similitudes, que se pueden explotar en una codificación/decodificación multicanal y dado que las dos señales residuales comprenden también típicamente similitudes, que se pueden explotar en una codificación/decodificación multicanal. De esta manera, se puede obtener típicamente una buena eficacia de codificación mediante el uso de este concepto.

[0013] En una realización preferida, el decodificador de audio se configura para proporcionar la primera señal residual y la segunda señal residual sobre la base de la representación codificada conjuntamente de la primera señal residual y de la segunda señal residual mediante el uso de una decodificación multicanal a base de predicciones. El uso de una decodificación multicanal a base de predicciones trae consigo típicamente una calidad de reconstrucción comparativamente buena para las señales residuales. Esto es ventajoso, por ejemplo, si la primera señal residual representa un lado izquierdo de una escena de audio y la segunda señal residual representa un lado derecho de la escena de audio, porque el oído humano es típicamente sensible en comparación para diferencias entre los lados izquierdo y derecho de la escena de audio.

[0014] En una realización preferida, el decodificador de audio se configura para proporcionar la primera señal residual y la segunda señal residual sobre la base de la representación codificada conjuntamente de la primera señal residual y de la segunda señal residual mediante el uso de una decodificación multicanal asistida por señales residuales. Se ha encontrado que una calidad particularmente buena de la primera y la segunda señal residual se puede lograr si la primera señal residual y la segunda señal residual se proporcionan mediante el uso de una decodificación multicanal, que, a su vez, recibe una señal residual (y típicamente también una señal de mezcla descendente, que combina la primera señal residual y la segunda señal residual). De esta manera, hay una cascada de etapas de decodificación, en la que dos señales residuales (la primera señal residual, que se usa para proporcionar la primera señal del canal de audio y la segunda señal del canal de audio y la segunda señal residual, que se usa para proporcionar la tercera señal del canal de audio y la cuarta señal del canal de audio), se proporcionan sobre la base de una señal de mezcla descendente de entrada y una señal residual de entrada, en la que la última también se puede designar como una señal residual común) de la primera señal residual y la segunda señal residual). De esta manera, la primera señal residual y la segunda señal residual son realmente señales residuales "intermediarias", que se derivan mediante el uso de una decodificación multicanal de una señal de mezcla descendente correspondiente y una señal residual "común" correspondiente.

[0015] En una realización preferida, la decodificación multicanal a base de predicciones se configura para evaluar un parámetro de predicción que describe una contribución de un componente de señal, que se deriva mediante el uso de un componente de señal de un marco previo, para proporcionar las señales residuales (es decir, la primera señal residual y la segunda señal residual) de un marco corriente. El uso de tal decodificación multicanal a base de predicciones trae consigo una calidad particularmente buena de las señales residuales (primera señal residual y segunda señal residual).

[0016] En una realización preferida, la decodificación multicanal a base de predicciones se configura para obtener la primera señal residual y la segunda señal residual sobre la base de una señal de mezcla descendente (correspondiente) y una señal residual "común" (correspondiente), en la que la decodificación multicanal a base de predicciones se configura para aplicar la señal residual común con un primer signo para obtener la primera señal

residual y para aplicar la señal residual común con un segundo signo, que se opone al primer signo, para obtener la segunda señal residual. Se ha encontrado que tal decodificación multicanal a base de predicciones trae consigo una buena eficacia para reconstruir la primera señal residual y la segunda señal residual.

5 **[0017]** En una realización preferida, el decodificador de audio se configura para proporcionar la primera señal residual y la segunda señal residual sobre la base de la representación codificada conjuntamente de la primera señal residual y de la segunda señal residual mediante el uso de una decodificación multicanal que es operativa en el dominio de transformada de coseno discreto modificado (dominio MDCT). Se ha encontrado que tal concepto se puede implementar de una manera eficaz, dado que una decodificación de audio, que se puede usar para
10 proporcionar la representación codificada conjuntamente de la primera señal residual y de la segunda señal residual, opera preferentemente en el dominio MDCT. Conforme a ello, se pueden evitar transformaciones intermedias por aplicación de la decodificación multicanal para proporcionar la primera señal residual y la segunda señal residual en el dominio MDCT.

15 **[0018]** En una realización preferida, el decodificador de audio se configura para proporcionar la primera señal residual y la segunda señal residual sobre la base de la representación codificada conjuntamente de la primera señal residual y de la segunda señal residual mediante el uso de una predicción estéreo compleja USAC (por ejemplo, tal como se mencionó en el estándar USAC antes mencionado). Se ha encontrado que tal predicción estéreo compleja USAC trae consigo buenos resultados para la decodificación de la primera señal residual y de la segunda señal
20 residual. Además, el uso de la predicción estéreo compleja USAC para la decodificación de la primera señal residual y la segunda señal residual permite también una simple implementación del concepto mediante el uso de bloques decodificantes que están ya disponibles en la codificación de voz y audio unificados (USAC). Conforme a ello, un decodificador de voz y audio unificados se puede reconfigurar fácilmente para realizar el concepto de decodificación tratado aquí.

25 **[0019]** En una realización preferida, el decodificador de audio se configura para proporcionar 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 y la primera señal residual mediante el uso de una decodificación multicanal asistida por señales residuales en base a parámetros. De modo similar, el decodificador de audio se configura para proporcionar 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 y la segunda
30 señal residual mediante el uso de una decodificación multicanal asistida por señales residuales en base a parámetros. Se ha encontrado que tal decodificación multicanal es bien apropiada para la derivación de las señales de canal de audio sobre la base de la primera señal de mezcla descendente, la primera señal residual, la segunda señal de mezcla descendente y la segunda señal residual. Además, se ha encontrado que tal decodificación
35 multicanal asistida por señales residuales en base a parámetros se puede implementar con un pequeño esfuerzo mediante el uso de bloques de procesamiento que ya están presentes en decodificadores de audio multicanal típicos.

[0020] En una realización preferida, la decodificación multicanal asistida por señales residuales en base a
40 parámetros se configura 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 canal de audio sobre la base de una respectiva señal de mezcla descendente y una respectiva señal residual correspondiente. Se ha encontrado que tal decodificación multicanal asistida por señales residuales en base a parámetros está bien adaptada para la segunda etapa de una decodificación multicanal en cascada (en la que, preferentemente, la
45 primera y la segunda señal de mezcla descendentes y la primera y la segunda señales residuales se proporcionan mediante el uso de una decodificación multicanal a base de predicciones).

[0021] En una realización preferida, el decodificador de audio se configura para proporcionar 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 y
50 la primera señal residual mediante el uso de una decodificación multicanal asistida por señales residuales que es operativa en el dominio QMF. De modo similar, el decodificador de audio se configura preferentemente para proporcionar 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 y la segunda señal residual mediante el uso de una decodificación multicanal asistida por señales residuales que es operativa en el dominio QMF. Conforme a ello, la segunda etapa de la decodificación
55 multicanal jerárquica es operativa en el dominio QMF, que se adapta bien a un típico post-procesamiento, que también se realiza en el dominio QMF, de modo tal que se puedan evitar conversiones intermedias.

[0022] En una realización preferida, el decodificador de audio se configura para proporcionar 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 y
60 la primera señal residual mediante el uso de una decodificación MPEG Surround 2–1–2 o una decodificación estéreo unificada. De modo similar, el decodificador de audio se configura preferentemente para proporcionar 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 y la segunda señal residual mediante el uso de una decodificación MPEG Surround 2–1–2 o una decodificación estéreo unificada. Se ha encontrado que tales conceptos de decodificación son particularmente muy apropiados para la segunda etapa de una decodificación jerárquica.

5

[0023] En una realización preferida, la primera señal residual y la segunda señal residual se asocian con diferentes posiciones horizontales (o, de modo equivalente, posiciones azimutales) de una escena de audio. Se ha encontrado que es particularmente ventajoso separar señales residuales, que se asocian con diferentes posiciones horizontales (o posiciones azimutales), en una primera etapa del procesamiento multicanal jerárquico porque se puede obtener una impresión auditiva particularmente si la separación izquierda/derecha perceptualmente importante se lleva a cabo en una primera etapa de la decodificación multicanal jerárquica.

10

[0024] En una realización preferida, la primera señal del canal de audio y la segunda señal del canal se asocian con posiciones vecinas de forma vertical de la escena de audio (o, de modo equivalente, con posiciones de elevación vecinas de la escena de audio). Además, la tercera señal del canal de audio y la cuarta señal del canal de audio están asociadas, preferentemente, con posiciones vecinas de forma vertical de la escena de audio (o, de modo equivalente, con posiciones de elevación vecinas de la escena de audio). Se ha encontrado que se pueden lograr buenos resultados de decodificación si la separación entre señales superior e inferior se lleva a cabo en una segunda etapa de la decodificación de audio jerárquica (que comprende típicamente una precisión de separación algo menor que la primera etapa), dado que el sistema auditivo humano es menos sensible con respecto a una posición vertical de una fuente de audio cuando se compara con una posición horizontal de la fuente de audio.

15

20

[0025] En una realización preferida, la primera señal del canal de audio y la segunda señal del canal de audio se asocian con una primera posición horizontal de una escena de audio (o, de modo equivalente, posición azimutal) y la tercera señal del canal de audio y la cuarta señal del canal de audio se asocian con una segunda posición horizontal de la escena de audio (o, de modo equivalente, posición azimutal), que es diferente de la primera posición horizontal (o, de modo equivalente, posición azimutal).

25

[0026] Preferentemente, 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 separación izquierda–derecha se lleva a cabo en una primera etapa de la decodificación de audio jerárquica.

30

[0027] En una realización preferida, la primera señal del canal de audio y la segunda señal del canal de audio se asocian con el 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.

35

[0028] En otra realización preferida, la primera señal del canal de audio está asociada con un lado izquierdo inferior de la escena de audio, la segunda señal del canal de audio está asociada con un lado izquierdo superior de la escena de audio, la tercera señal del canal de audio está asociada con un lado derecho inferior de la escena de audio y la cuarta señal del canal de audio está asociada con un lado derecho superior de la escena de audio. Tal asociación de las señales de canal de audio trae consigo particularmente buenos resultados de codificación.

40

[0029] En una realización preferida, el decodificador de audio se configura 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 mediante el uso de una decodificación multicanal, en la que la primera señal de mezcla descendente está asociada con el lado izquierdo de una escena de audio y la segunda señal de mezcla descendente está asociada con el lado derecho de la escena de audio. Se ha encontrado que las señales de mezcla descendente también se pueden codificar con buena eficacia de codificación mediante el uso de una codificación multicanal, incluso si las señales de mezcla descendente se asocian con diferentes lados de la escena de audio.

45

50

[0030] En una realización preferida, el decodificador de audio se configura para proporcionar 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 de la segunda señal de mezcla descendente mediante el uso de una decodificación multicanal a base de predicciones o incluso mediante el uso de una decodificación multicanal en base a predicciones asistida por señales residuales. Se ha encontrado que el uso de tales conceptos de decodificación multicanal proporciona un resultado de decodificación particularmente bueno. Además, las funciones de decodificación existentes se pueden reutilizar en algunos decodificadores de audio.

55

[0031] En una realización preferida, el decodificador de audio se configura para llevar a cabo una primera extensión del ancho de banda multicanal sobre la base de la primera señal del canal de audio y la tercera señal del

60

canal de audio. Además, el decodificador de audio se puede configurar para llevar a cabo una segunda extensión de ancho de banda multicanal (típicamente separado) sobre la base de la segunda señal del canal de audio y la cuarta señal del canal de audio. Se ha encontrado que es ventajoso llevar a cabo una posible extensión del ancho de banda sobre la base de dos señales de canal de audio que se asocian con diferentes lados de una escena de audio (en la que diferentes señales residuales están típicamente asociadas con diferentes lados de la escena de audio).

[0032] En una realización preferida, el decodificador de audio se configura para llevar a cabo la primera extensión del ancho de banda multicanal a fin de obtener dos o más señales de canal de audio de ancho de banda extendido asociados con un primer plano horizontal común (o, de modo equivalente, con una primera elevación común) de una escena de audio sobre la base de la primera señal del canal de audio y la tercera señal del canal de audio y uno o más parámetros de extensión de ancho de banda. Además, el decodificador de audio se configura preferentemente para llevar a cabo la segunda extensión del ancho de banda multicanal a fin de obtener dos o más señales de canal de audio de ancho de banda extendido asociada con un segundo plano horizontal común (o, de modo equivalente, una segunda elevación común) de la escena de audio sobre la base de la segunda señal del canal de audio y la cuarta señal del canal de audio y uno o más parámetros de extensión de ancho de banda. Se ha encontrado que tal esquema de decodificación resulta en una buena calidad de audio, dado que la extensión del ancho de banda multicanal puede considerar características estéreo, que son importantes para la impresión auditiva, en tal disposición.

[0033] En una realización preferida, la representación codificada conjuntamente de la primera señal residual y de la segunda señal residual comprende un elemento del par de canales que comprende una señal de mezcla descendente de la primera y la segunda señal residual y una señal residual común de la primera y la segunda señal residual. Se ha encontrado que la codificación de la señal de mezcla descendente de la primera y la segunda señal residual y de la señal residual común de la primera y la segunda señal residual mediante el uso de un elemento del par de canales es ventajosa dado que la señal de mezcla descendente de la primera y la segunda señal residual y la señal residual común de la primera y la segunda señal residual comparten típicamente una cantidad de características. Conforme a ello, el uso de un elemento del par de canales típicamente reduce una señalización superior y, en consecuencia, permite una codificación eficaz.

[0034] En otra realización preferida, el decodificador de audio se configura 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 mediante el uso de una decodificación multicanal, en la que la representación codificada conjuntamente de la primera señal de mezcla descendente y de la segunda señal de mezcla descendente comprende un elemento del par de canales, comprendiendo el elemento del par de canales una señal de mezcla descendente de la primera y la segunda señal de mezcla descendente y una señal residual común de la primera y la segunda señal de mezcla descendente. Esta realización se basa en las mismas consideraciones que la realización descrita con anterioridad.

[0035] Otra realización según la invención crea un codificador de audio para proporcionar una representación codificada sobre la base de al menos cuatro señales de canal de audio. El codificador de audio se configura para codificar conjuntamente al menos una primera señal del canal de audio y una segunda señal del canal de audio mediante el uso de una codificación multicanal asistida por señales residuales, para obtener una primera señal de mezcla descendente y una primera señal residual. El codificador de audio se configura para codificar conjuntamente al menos una tercera señal de canal de audio y una cuarta señal de canal de audio mediante el uso de una codificación multicanal asistida por señales residuales, para obtener una segunda señal de mezcla descendente y una segunda señal residual. Además, el codificador de audio se configura para codificar conjuntamente la primera señal residual y la segunda señal residual mediante el uso de una codificación multicanal que explota similitudes y/o dependencias entre las señales residuales, para obtener una representación codificada conjuntamente de las señales residuales. Este codificador de audio se basa en las mismas consideraciones que el decodificador de audio antes descrito.

[0036] Además, mejoras opcionales de este codificador de audio y configuraciones preferidas del codificador de audio, están sustancialmente en paralelo con mejoras y configuraciones preferidas del decodificador de audio tratado con anterioridad. Conforme a ello, se hace referencia a la discusión anterior.

[0037] Otra realización según la invención crea un procedimiento para proporcionar al menos cuatro señales de canal de audio sobre la base de una representación codificada, que realiza sustancialmente la funcionalidad del codificador de audio descrito con anterioridad y que se puede suplementar por cualquiera de las características y funcionalidades tratadas con anterioridad.

[0038] Otra realización según la invención crea un procedimiento para proporcionar una representación

codificada sobre la base de al menos cuatro señales de canal de audio, que sustancialmente satisface la funcionalidad del decodificador de audio descrito con anterioridad.

[0039] Otra realización según la invención crea un programa informático para llevar a cabo los procedimientos mencionados con anterioridad.

Breve descripción de las figuras

[0040] Las realizaciones según la presente invención se describirán posteriormente haciendo referencia a las figuras adjuntas, en las que:

- Fig. 1 muestra un diagrama esquemático de bloque de un codificador de audio, según una realización de la presente invención;
- Fig. 2 muestra un diagrama esquemático de bloque de un decodificador de audio, según una realización de la presente invención;
- Fig. 3 muestra un diagrama esquemático de bloque de un decodificador de audio, según otra realización de la presente invención;
- Fig. 4 muestra un diagrama esquemático de bloque de un codificador de audio, según una realización de la presente invención;
- Fig. 5 muestra un diagrama esquemático de bloque de un decodificador de audio, según una realización de la presente invención;
- Fig. 6 muestra un diagrama esquemático de bloque de un decodificador de audio, según otra realización de la presente invención;
- Fig. 7 muestra un diagrama de flujo de un procedimiento para proporcionar una representación codificada sobre la base de al menos cuatro señales de canal de audio, según una realización de la presente invención;
- Fig. 8 muestra un diagrama de flujo de un procedimiento para proporcionar al menos cuatro señales de canal de audio sobre la base de una representación codificada, según una realización de la invención;
- Fig. 9 muestra un diagrama de flujo de un procedimiento para proporcionar una representación codificada sobre la base de al menos cuatro señales de canal de audio, según una realización de la invención; y
- Fig. 10 muestra un diagrama de flujo de un procedimiento para proporcionar al menos cuatro señales de canal de audio sobre la base de una representación codificada, según una realización de la invención;
- Fig. 11 muestra un diagrama esquemático de bloque de un codificador de audio, según una realización de la invención;
- Fig. 12 muestra un diagrama esquemático de bloque de un codificador de audio, según otra realización de la invención;
- Fig. 13 muestra un diagrama esquemático de bloque de un decodificador de audio, según una realización de la invención;
- Fig. 14a muestra una representación de sintaxis de una corriente de bits, que se puede usar con el codificador de audio según la Fig. 13;
- Fig. 14b muestra una representación en tabla de diferentes valores del parámetro $q_{celIndex}$;
- Fig. 15 muestra un diagrama esquemático de bloque de un codificador de audio 3D en el que se pueden usar los conceptos según la presente invención;
- Fig. 16 muestra un diagrama esquemático de bloque de un decodificador de audio 3D, en el que se usan los conceptos según la presente invención; y
- Fig. 17 muestra un diagrama esquemático de bloque de un convertidor de formato.
- Fig. 18 muestra una representación gráfica de una estructura topológica de un elemento de canal cuádruple (QCE), según una realización de la presente invención;
- Fig. 19 muestra un diagrama esquemático de bloque de un decodificador de audio, según una realización de la presente invención;
- Fig. 20 muestra un diagrama esquemático de bloques detallado de un decodificador QCE, según una realización de la presente invención; y
- Fig. 21 muestra un diagrama esquemático de bloques detallado de un codificador de canal cuádruple, según una realización de la presente invención.

55 Descripción detallada de las realizaciones

1. Codificador de audio según la Fig. 1

[0041] La Fig. 1 muestra un diagrama esquemático de bloque de un codificador de audio, que se designa en su totalidad con 100. El codificador de audio 100 se configura para proporcionar una representación codificada sobre la base de al menos cuatro señales de canal de audio. El codificador de audio 100 se configura 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. Además, el codificador de audio 100 se configura 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 se configura para codificar conjuntamente la primera señal del canal de audio 110 y la segunda señal del canal de audio 112 mediante el uso de una codificación multicanal asistida por señales residuales, para obtener la primera señal de mezcla descendente 120 y una primera señal residual 142. El codificador de señales de audio 100 comprende también un codificador multicanal asistido con señales residuales 150, que se configura para codificar conjuntamente al menos la tercera señal del canal de audio 114 y la cuarta señal del canal de audio 116 mediante el uso de una codificación multicanal asistida por señales residuales, para obtener la segunda señal de mezcla descendente 122 y una segunda señal residual 152. El decodificador de audio 100 comprende también un codificador multicanal 160, que se configura para codificar conjuntamente la primera señal residual 142 y la segunda señal residual 152 mediante el uso de una codificación multicanal, para obtener la representación codificada conjuntamente 130 de las señales residuales 142, 152.

[0042] Con respecto a la funcionalidad del codificador de audio 100, se debería observar que el codificador de audio 100 realiza una codificación jerárquica, en la que la primera señal del canal de audio 110 y la segunda señal del canal de audio 112 se codifican conjuntamente mediante el uso de 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 puede ser representada por la primera señal de mezcla descendente 120 y parámetros opcionales, que se pueden proporcionar por el codificador multicanal asistido por 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 por el codificador multicanal asistido por señales residuales 140. Por ejemplo, la primera señal residual 142 puede permitir al menos una reconstrucción de forma ondulada parcial de la primera señal del canal de audio 110 y de la segunda señal del canal de audio 112 en el 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 por 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 modo tal 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 en el 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 canal de audio 110, 112, 114, 116 comprenden la misma correlación, la primera señal residual 142 y la segunda señal residual 152 también se correlacionan típicamente hasta cierto grado. Conforme a ello, la codificación conjunta de la primera señal residual 142 y de la segunda señal residual 152 mediante el uso del codificador multicanal 160 comprende típicamente una alta eficiencia dado que una codificación multicanal de señales correlacionadas típicamente reduce la tasa de bits explotando las 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 pequeñas.

[0043] Para resumir, la realización según la Fig. 1 proporciona una codificación multicanal jerárquica, en la que una buena calidad de reproducción se puede lograr mediante el uso de los codificadores multicanal asistidos por señales residuales 140, 150 y en la que una demanda de tasa de bits se puede mantener moderada por codificación conjunta de una primera señal residual 142 y una segunda señal residual 152.

[0044] 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 la que la funcionalidad del codificador de audio es típicamente inversa a la funcionalidad del decodificador de audio.

2. Decodificador de audio según la Fig. 2

[0045] La Fig. 2 muestra un diagrama esquemático de bloque de un decodificador de audio, que se designa en su totalidad con 200.

[0046] El decodificador de audio 200 se configura 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 recibe también una representación de una primera señal de mezcla descendente 212 y de una segunda señal de mezcla descendente 214. El decodificador de audio 200 se configura para proporcionar una primera señal del canal de audio 220, una segunda señal del canal de audio 222, una tercera señal de canal de audio 224 y una cuarta señal de canal de audio 226.

[0047] El decodificador de audio 200 comprende un decodificador multicanal 230, que se configura 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 de audio 200 comprende también un (primer) decodificador multicanal asistido por señales residuales 240 que se configura 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 mediante el uso de una decodificación multicanal. El decodificador de audio 200 comprende también un (segundo) decodificador multicanal asistido por señales residuales 250, que se configura 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.

[0048] Con respecto a 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 222 sobre la base de una (primera) decodificación multicanal asistida por señales residuales comunes 240, en la 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 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 por 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 ondulada parcial de la primera señal del canal de audio 220 y de la segunda señal del canal de audio 222.

[0049] De modo similar, el (segundo) decodificador multicanal asistido por 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 la 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. Además, se pueden describir diferencias entre la tercera señal del canal de audio 224 y la cuarta señal del canal de audio 226, por ejemplo, por parámetros (opcionales), que pueden ser recibidos por el (segundo) decodificador multicanal asistido por 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 ondulada 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.

[0050] Sin embargo, la primera señal residual 232 y la segunda señal residual 234 se derivan de una representación codificada conjuntamente 210 de la primera señal residual y de la segunda señal residual. Tal 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 típicamente similares o "correlacionadas". Conforme a ello, la primera señal residual 232 y la segunda señal residual 234 también son típicamente similares o "correlacionadas", que se pueden explotar derivando la primera señal residual 232 y la segunda señal residual 234 de una representación codificada conjuntamente 210 mediante el uso de una decodificación multicanal.

[0051] En consecuencia, es posible obtener una alta calidad de decodificación con tasa de bits moderada por decodificación las señales residuales 232, 234 sobre la base de una representación codificada conjuntamente 210 y mediante el uso de cada una de las señales residuales para la decodificación de dos o más señales de canal de audio.

[0052] Para concluir, el decodificador de audio 200 permite una elevada eficacia de codificación al proporcionar señales de canal de audio de alta calidad 220, 222, 224, 226.

[0053] Se debería observar que 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

5

[0054] La Fig. 3 muestra un diagrama esquemático de bloque de un decodificador de audio según otra realización de la presente invención. El decodificador de audio de la Fig. 3 se designó en su totalidad con 300. El decodificador de audio 300 es similar al decodificador de audio 200 según la Fig. 2, de tal forma que se aplican también las explicaciones anteriores. Sin embargo, el decodificador de audio 300 se suplementa con características y funcionalidades adicionales cuando se compara con el decodificador de audio 200, como se explicará a continuación.

10

[0055] El decodificador de audio 300 se configura para recibir una representación codificada conjuntamente 310 de una primera señal residual y de una segunda señal residual. Además, el decodificador de audio 300 se configura para recibir una representación codificada conjuntamente 360 de una primera señal de mezcla descendente y de una segunda señal de mezcla descendente. Además, el decodificador de audio 300 se configura 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 se configura para recibir la representación codificada conjuntamente 310 de la primera señal residual y de la segunda señal residual y para proporcionar, en base a ello, una primera señal residual 332 y una segunda señal residual 334. El decodificador de audio 300 comprende también 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 comprende también una (segunda) decodificación multicanal asistida con señales residuales 350, que se configura 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.

20

25

[0056] El decodificador de audio 300 también comprende otro decodificador multicanal 370, que se configura 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, en base a ello, la primera señal de mezcla descendente 312 y la segunda señal de mezcla descendente 314.

30

[0057] 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).

35

[0058] En una realización preferida, 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 por señales residuales a base de 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 deriva mediante el uso de un componente de señal de un marco previo, a una provisión de la primera señal residual 332 y la segunda señal residual 334 para un marco corriente. Además, 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 se opone al primer signo, para obtener la segunda señal residual 334. De esta manera, la señal residual común puede describir, al menos en parte, 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, para 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. Además, se debería observar que la primera señal residual 332 puede estar asociada con una primera posición horizontal (o posición azimutal), por ejemplo, una posición

40

45

50

55

60

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.

[0059] 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 por señales residuales a base de 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 tal forma que también se pueden aplicar las explicaciones y referencias anteriores. Además, 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 como el decodificador multicanal 330 pueden realizar una división horizontal (o separación horizontal o distribución horizontal).

[0060] El decodificador multicanal asistido por señales residuales 340 puede estar basado preferentemente 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 por 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 por 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 de forma vertical 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 modo 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 por señales residuales 340 puede llevar a cabo una división vertical (o distribución o separación).

[0061] La funcionalidad del decodificador multicanal asistido por señales residuales 350 puede ser idéntica a la funcionalidad del decodificador multicanal asistido por señales residuales 340, en el 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 el 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 de forma vertical 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 por señales residuales 350 realiza una división vertical (o separación o distribución).

[0062] Para resumir, el decodificador de audio 300 según la Fig. 3 realiza una decodificación de audio jerárquica, en la que una división izquierda–derecha se lleva a cabo en la primera etapas (decodificador multicanal 330, decodificador multicanal 370) y en la que una división superior–inferior se lleva a cabo en la segunda etapa (decodificadores multicanal asistidos por señales residuales 340, 350). Además, las señales residuales 332, 334 se codifican también mediante el uso de 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 explotan tanto para la codificación (y 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 logra una elevada eficacia de codificación y las correlaciones entre las señales se explotan adecuadamente.

4. Codificador de audio según la Fig. 4

5

[0063] La Fig. 4 muestra un diagrama esquemático de bloque de un codificador de audio, según otra realización de la presente invención. El codificador de audio según la Fig. 4 se designa en su totalidad con 400. El codificador de audio 400 se configura para recibir cuatro señales de canal de audio, a saber, 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. Además, el codificador de audio 400 se configura para proporcionar una representación codificada sobre la base de las señales de canal 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 se configura 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 se configura 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.

[0064] Además, el codificador de audio 400 comprende un (primer) codificador multicanal 450, que se configura para codificar conjuntamente al menos la primera señal del canal de audio 410 y la segunda señal del canal de audio 412 mediante el uso de una codificación multicanal, para obtener una primera señal de mezcla descendente 452. Además, el codificador de audio 400 también comprende un (segundo) codificador multicanal 460, que se configura para codificar conjuntamente al menos la tercera señal del canal de audio 414 y la cuarta señal del canal de audio 416 mediante el uso de una codificación multicanal, para obtener una segunda señal de mezcla descendente 462. Además, el codificador de audio 400 comprende también un (tercer) codificador multicanal 470, que se configura para codificar conjuntamente la primera señal de mezcla descendente 452 y la segunda señal de mezcla descendente 462 mediante el uso de una codificación multicanal, para obtener la representación codificada conjuntamente 420 de las señales de mezcla descendente.

[0065] Con respecto a la funcionalidad del codificador de audio 400, se debería observar que el codificador de audio 400 realiza una codificación multicanal jerárquica, en la 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 la que la tercera señal del canal de audio 414 y la cuarta señal del canal de audio 416 se combinan también en la primera etapa, para obtener así 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 canal de audio 410, 414 que se manipulan por diferentes codificadores multicanal 450, 460 en la primera 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 canal de audio 412, 416, que se manipulan por diferentes codificadores multicanal 450, 460 en la primera etapa de procesamiento. Este orden específico de procesamiento trae consigo la ventaja de que los grupos 422, 424 de parámetros de extensión de ancho de banda se basan en canales 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 tales canales de audio en la primera etapa de la codificación jerárquica, cuya relación no es muy relevante con respecto a una percepción de la 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 una percepción de ubicación de la fuente de sonido, porque la relación entre 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 se basa en dos canales de audio (señales de canal 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 proporciona sobre la base de señales de canal 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 canal de audio antes descrito en la codificación multicanal jerárquica. En consecuencia, el primer grupo 422 de parámetros de extensión de ancho de banda comunes se basa 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 462, en la que la última domina típicamente la impresión espacial generada en el lado de un decodificador de audio. Conforme a ello, la provisión del primer grupo 422 de parámetros de extensión de ancho de banda y también la provisión del segundo grupo 424 de parámetros de extensión de ancho de banda se adapta bien a una impresión acústica espacial que se genera en el lado de un
5 decodificador de audio.

5. Decodificador de audio según la Fig. 5

[0066] La Fig. 5 muestra un diagrama esquemático de bloque de un decodificador de audio, según otra
10 realización de la presente invención. El decodificador de audio según la Fig. 5 se designa en su totalidad con 500.

[0067] El decodificador de audio 500 se configura para recibir una representación codificada conjuntamente 510 de una primera señal de mezcla descendente y una segunda señal de mezcla descendente. Además, el decodificador de audio 500 se configura para proporcionar una primera señal del canal de ancho de banda extendido
15 520, una segunda señal del canal de ancho de banda extendido 522, una tercera señal del canal de ancho de banda extendido 524 y una cuarta señal del canal de ancho de banda extendido 526.

[0068] El decodificador de audio 500 comprende un (primer) decodificador multicanal 530, que se configura 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 mediante el uso de una decodificación multicanal. El decodificador de audio 500 comprende también un (segundo) decodificador multicanal 540, que se configura 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 mediante el uso de una decodificación multicanal. El decodificador de audio 500
20 comprende también un (tercer) decodificador multicanal 550, que se configura 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 mediante el uso de una decodificación multicanal. Además, el decodificador de audio 500 comprende una (primera) extensión de ancho de banda multicanal 560, que se configura para llevar a cabo 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 ancho de banda extendido 520 y la tercera
30 señal del canal de ancho de banda extendido 524. Además, el decodificador de audio comprende una (segunda) extensión de ancho de banda multicanal 570, que se configura para llevar a cabo 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, para obtener la segunda señal del canal de ancho de banda extendido 522 y la cuarta señal del canal de ancho de banda
35 extendido 526.

[0069] Con respecto a la funcionalidad del decodificador de audio 500, se debería observar que el decodificador de audio 500 realiza una decodificación multicanal jerárquica, en la 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
40 etapa de la decodificación jerárquica y en la que la primera señal del canal de audio 542 y la segunda señal del canal de audio 544 se derivan de la primera señal de mezcla descendente 532 en una segunda etapa de la decodificación jerárquica y en la que la tercera señal del canal de audio 556 y la cuarta señal del canal de audio 558 se derivan 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
45 banda multicanal 570 reciben cada una, 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 logra típicamente 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
50 señales de entrada que están bien separadas (porque se originan de la primera señal de mezcla descendente 532 y la segunda señal de mezcla descendente 534, que están bien separadas por canales). De esta manera, la extensión del ancho de banda multicanal 560, 570 pueden considerar 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.

[0070] En otras palabras, la estructura "cruzada" del decodificador de audio, en la 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
55 estéreo entre los canales.

[0071] Sin embargo, se debería observar que el decodificador de audio 500 se puede suplementar por
60

cualquiera de las características y funcionalidades descritas en esta invención con respecto a los decodificadores de audio según las Figs. 2, 3, 6 y 13, en donde es posible introducir características individuales en el decodificador de audio 500 para mejorar de forma gradual el rendimiento del decodificador de audio.

5 6. Decodificador de audio según la Fig. 6

[0072] La Fig. 6 muestra un diagrama esquemático de bloque de un decodificador de audio según otra realización de la presente invención. El decodificador de audio según la Fig. 6 se designa 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 tal modo 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.

[0073] El decodificador de audio 600 se configura 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 de ancho de banda extendido 620, una segunda señal de ancho de banda extendido 622, una tercera señal de ancho de banda extendido 624 y una cuarta señal de ancho de banda extendido 626. El decodificador de audio 600 comprende un decodificador multicanal 630, que se configura 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, en base a ello, la primera señal de mezcla descendente 632 y la segunda señal de mezcla descendente 634. El decodificador de audio 600 comprende también un decodificador multicanal 640, que se configura para recibir la primera señal de mezcla descendente 632 y para proporcionar, en base a ello, una primera señal del canal de audio 542 y una segunda señal del canal de audio 544. El decodificador de audio 600 comprende también un decodificador multicanal 650, que se configura 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 comprende también una (primera) extensión de ancho de banda multicanal 660, que se configura para recibir la primera señal del canal de audio 642 y la tercera señal del canal de audio 656 y para proporcionar, en base a ello, la primera señal del canal de ancho de banda extendido 620 y la tercera señal del canal de 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, en base a ello, la segunda señal del canal de ancho de banda extendido 622 y la cuarta señal del canal de ancho de banda extendido 626.

[0074] El decodificador de audio 600 también comprende otro decodificador multicanal 680, que se configura para recibir una representación codificada conjuntamente 682 de una primera señal residual y de una segunda señal residual y que proporciona, en base a ello, 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.

[0075] El decodificador multicanal 630 es preferentemente un decodificador multicanal asistido por señales residuales en base a predicciones. Por ejemplo, el decodificador multicanal 630 puede ser sustancialmente idéntico al decodificador multicanal 370 descrito con anterioridad. Por ejemplo, el decodificador multicanal 630 puede ser un decodificador de predicción estéreo complejo USAC, como se mencionó con anterioridad y tal como se describe en el estándar USAC mencionado con anterioridad. Conforme a ello, la representación codificada conjuntamente 610 de la 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 evaluados por el decodificador multicanal 630.

[0076] Además, 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, una 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.

[0077] Además, el decodificador multicanal 680 puede ser, por ejemplo, un decodificador multicanal asociado con señales residuales en base a predicciones. El decodificador multicanal 680 puede ser sustancialmente idéntico al decodificador multicanal 330 descrito con anterioridad. Por ejemplo, el decodificador multicanal 680 puede ser un decodificador USAC de predicción estéreo compleja, como se mencionó con anterioridad. En consecuencia, la 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 evaluados por el decodificador multicanal 680. Además, se debería observar que la primera señal residual 684 puede estar asociada con una primera posición horizontal o posición azimutal (por ejemplo, una posición 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.

[0078] El decodificador multicanal 640 puede ser, por ejemplo, una decodificación multicanal a base de parámetros como, por ejemplo, una decodificación multicanal envolvente MPEG, tal como se describió con anterioridad y en el estándar de referencia. Sin embargo, en presencia del decodificador multicanal (opcional) 680 y la primera señal residual (opcional) 684, el decodificador multicanal 640 puede ser un decodificador multicanal asistido por señales residuales en base a parámetros como, por ejemplo, un decodificador estéreo unificado. De esta manera, el decodificador multicanal 640 puede ser sustancialmente idéntico al decodificador multicanal 340 descrito con anterioridad y el decodificador multicanal 640 puede recibir, por ejemplo, los parámetros 342 descritos con anterioridad.

[0079] 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).

[0080] Además, 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 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 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).

[0081] 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 canal 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 considerar 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 considerar características estéreo, dado que la segunda extensión del ancho de banda multicanal opera en señales de canal 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.

[0082] 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 opera en un par de señales izquierda / derecha (extensión de ancho de banda multicanal 660, 670). Este "cruce" de las rutas de decodificación permite que se pueda llevar a cabo 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 llevar a cabo en un par de señales de canal 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 canal de audio (o señales del canal de ancho de banda extendido) sin degradar de forma significativa la impresión auditiva.

7. Procedimiento según la Fig. 7

[0083] 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 canal de audio.

5

[0084] 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 mediante el uso de una codificación multicanal asistida por señales residuales, para obtener una primera señal de mezcla descendente y una primera señal residual. El procedimiento comprende también la codificación conjunta 720 de al menos una tercera señal de canal de audio y una cuarta señal de canal de audio mediante el uso de una codificación multicanal asistida por señales residuales, para obtener una segunda señal de mezcla descendente y una segunda señal residual. El procedimiento comprende también la codificación conjunta 730 de la primera señal residual y la segunda señal residual mediante el uso de una codificación multicanal, para obtener una representación codificada de las señales residuales. Sin embargo, se debería observar que el procedimiento 700 puede ser suplementado 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

[0085] La Fig. 8 muestra un diagrama de flujo de un procedimiento 800 para proporcionar al menos cuatro señales de canal de audio sobre la base de una representación codificada.

[0086] El procedimiento 800 comprende la provisión 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 mediante el uso de una decodificación multicanal. El procedimiento 800 comprende también la provisión 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 mediante el uso de una decodificación multicanal asistida por señales residuales. El procedimiento comprende también la provisión 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 mediante el uso de una decodificación multicanal asistida por señales residuales.

30

[0087] Además, 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

[0088] La Fig. 9 muestra un diagrama de flujo de un procedimiento 900 para proporcionar una representación codificada sobre la base de al menos cuatro señales del canal de audio.

[0089] 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 comprende también la obtención 920 de 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 comprende también la codificación conjunta de al menos la primera señal del canal de audio y la segunda señal del canal de audio mediante el uso de una codificación multicanal, para 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 mediante el uso de una codificación multicanal para obtener una segunda señal de mezcla descendente. El procedimiento comprende también la codificación conjunta 950 de la primera señal de mezcla descendente y la segunda señal de mezcla descendente mediante el uso de una codificación multicanal, para obtener una representación codificada de las señales de mezcla descendente.

[0090] Se debería observar que algunas de las etapas del procedimiento 900, que no comprenden interdependencias específicas, se pueden llevar a cabo por orden arbitrario o en paralelo. Además, 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

[0091] La Fig. 10 muestra un diagrama de flujo de un procedimiento 1000 para proporcionar al menos cuatro señales de canal de audio sobre la base de una representación codificada.

60

[0092] El procedimiento 1000 comprende la provisión 1010 de 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 mediante el uso de una decodificación multicanal, la provisión 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 mediante el uso de una decodificación multicanal, la provisión 1030 de 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 mediante el uso de 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, para obtener una primera señal del canal de ancho de banda extendido y una tercera señal del canal de 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, para obtener una segunda señal del canal de ancho de banda extendido y una cuarta señal del canal de ancho de banda extendido.

[0093] Se debería observar que algunas de las etapas del procedimiento 1000 se pueden llevar a cabo en paralelo o en un orden diferente. Además, 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 al decodificador de audio.

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

[0094] A continuación, se describirán algunas realizaciones adicionales según la presente invención y las consideraciones subyacentes.

[0095] La Fig. 11 muestra un diagrama esquemático de bloque de un codificador de audio 1100 según una realización de la invención. El codificador de audio 1100 se configura 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 canal superior derecho 1116.

[0096] El codificador de audio 1100 comprende un primer codificador de audio multicanal (o codificación) 1120, que es un codificador de audio MPEG surround 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 de audio multicanal 1120 proporciona una señal de mezcla descendente 1122 y, opcionalmente, una señal residual izquierda 1124. Además, 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 derecha 1132 y, opcionalmente, una señal residual derecha 1134. El codificador de audio 1100 comprende también un codificador estéreo (o codificación) 1140, que recibe la señal de mezcla descendente izquierda 1122 y la señal de mezcla descendente derecha 1132. Además, 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 designa con 1144 y que describe la señal de mezcla descendente izquierda 1122 y la señal de mezcla descendente derecha 1132 en una forma conjuntamente codificada. Además, el codificador de audio 1100 comprende opcionalmente un segundo codificador estéreo (o codificación) 1150, que se configura 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, se configura para proporcionar un elemento del par de canales (CPE) “residual”, que representa la señal residual izquierda 1124 y la señal residual derecha 1134 en una forma conjuntamente codificada.

[0097] El codificador 1100 (así como los otros codificadores de audio descritos en esta invención) se basa en la idea de que las dependencias de señales horizontales y verticales se explotan combinando jerárquicamente herramientas estéreo USAC disponibles (es decir, conceptos de codificación que están disponibles en la codificación USAC). Los pares de canales vecinos de forma vertical se combinan mediante el uso de MPEG surround 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 requisitos perceptuales para 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.

- [0098]** La estructura jerárquica explicada con referencia a la Fig. 11 se puede lograr permitiendo 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-/post-procesamiento 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 resulta en la estructura del codificador mostrada en la Fig. 12.
- 10 **[0099]** La Fig. 12 muestra un diagrama esquemático de bloque de un codificador de audio 1200, según una realización de la invención. El codificador de audio 1200 se configura 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 codificador de audio 1200 se configura para proporcionar una corriente de bits 1220 para un primer elemento del par de canales y una corriente de bits 1222 para un segundo elemento del par de canales.
- 15 **[0100]** 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. Además, el primer codificador multicanal 1230 proporciona una primera señal de mezcla descendente 1232, una carga útil MPEG surround 1236 y, opcionalmente, una primera señal residual 1234. El
- 20 codificador de audio 1200 comprende también un segundo codificador multicanal 1240 que es un codificador MPEG surround 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.
- 25 **[0101]** El codificador de audio 1200 comprende también 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
- 30 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). Además, la (primera) codificación estéreo de predicción compleja 1250 proporciona una carga útil de predicción compleja 1254, que comprende típicamente uno o varios coeficientes de predicción compleja. Además, el codificador de audio 1200
- 35 comprende también 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
- 40 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). Además, la codificación estéreo de predicción compleja 1260 proporciona una carga útil de predicción compleja 1264 que comprende típicamente uno o varios coeficientes de predicción.
- 45 **[0102]** Además, 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
- 50 proporcionada por el modelo psicoacústico 1270 es opcional.
- [0103]** Además, 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
- 55 carga útil envolvente MPEG 1236 del primer codificador multicanal de audio 1230. Además, la primera codificación y multiplexación 1280 puede recibir información del modelo psicoacústico 1270, que describe, por ejemplo, qué precisión de codificación se debería 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.
- 60 **[0104]** Además, el codificador de audio 1200 comprende una segunda codificación y multiplexación 1290,

que se configura 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 envolvente MPEG 1246 proporcionada por el segundo codificador multicanal de audio 1240. Además, 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.

[0105] Con respecto a 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.

[0106] Además, se debería observar que este concepto se puede extender al uso de múltiples cajas envolventes 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.

[0107] A continuación, se describirá la implementación de un elemento del canal cuádruple. En un sistema de codificación de auto tridimensional, se usa la combinación jerárquica de cuatro canales para formar un elemento del canal cuádruple (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 mediante el uso de 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 utilizan 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.

[0108] Una posible estructura de decodificador se describirá tomando como referencia la Fig. 13 que muestra un diagrama esquemático de 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.

[0109] El decodificador de audio 1300 está configurado para proporcionar una primera señal de canal de ancho de banda extendido 1320, que puede, por ejemplo, representar una posición izquierda inferior de una escena de audio, una segunda señal de canal de ancho de banda extendido 1322, que puede, por ejemplo, representar una posición izquierda superior de la escena de audio, una tercera señal de canal de 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 de ancho de banda extendido 1326, que puede, por ejemplo, estar asociado con una posición derecha superior de la escena de audio.

[0110] El decodificador de audio 1300 comprende una primera corriente de bits decodificadora 1330, que se configura para recibir la corriente de bits 1310 para el primer elemento del par de canales y para proporcionar, sobre la base de esta, una representación codificada de forma conjunta de dos señales de la mezcla descendente, una carga útil de predicción compleja 1334, una carga útil de MPEG surround 1336 y una carga útil de replicación de ancho de banda espectral 1338. El decodificador de audio 1300 comprende también un primer decodificador estéreo para predicción compleja 1340, que se configura para recibir la representación codificada de forma conjunta 1332 y la carga útil de predicción compleja 1334 y para proporcionar, sobre esta base, 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 se configura para recibir la corriente de bits 1312 para el segundo elemento del canal y para proporcionar, sobre esta base, una representación codificada de forma conjunta 1352 de dos señales residuales, una carga útil de predicción compleja 1354, una carga útil de MPEG surround 1356 y una carga de bits de replicación del ancho de banda espectral 1358. El decodificador de audio comprende también un segundo decodificador estéreo para predicción compleja 1360, que proporciona una primera señal residual 1362 y una segunda señal residual 1364 sobre la base de la representación codificada de forma conjunta 1352 y la carga útil de predicción compleja 1354.

[0111] Además, el decodificador de audio 1300 comprende un primer decodificador multicanal tipo MPEG

surround 1370, que es un decodificador 2–1–2 MPEG surround o un decodificador estéreo unificado. El primer decodificador multicanal tipo MPEG surround 1370 recibe la primera señal de la mezcla descendente 1342, la primera señal residual 1362 (opcional) y la carga útil de MPEG surround 1336 y proporciona, sobre esta base, 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 MPEG surround 1380, que es un decodificador multicanal 2–1–2 MPEG surround o un decodificador multicanal estéreo unificado. El segundo decodificador multicanal tipo MPEG surround 1380 recibe la segunda señal de la mezcla descendente 1344 y la segunda señal residual 1364 (opcional), así como la carga útil de MPEG surround 1356, y proporciona, sobre esta base, una tercera señal del canal de audio 1382 y la cuarta señal del canal de audio 1384. El decodificador de audio 1300 comprende también una primera replicación del ancho de banda espectral estéreo 1390, que se configura 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 esta base, la primera señal de canal de ancho de banda extendido 1320 y la tercera señal de canal de ancho de banda extendido 1324. Además, el decodificador de audio comprende una segunda replicación del ancho de banda espectral estéreo 1394, que se configura 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 esta base, la segunda señal de canal de ancho de banda extendido 1322 y la cuarta señal de canal de ancho de banda extendido 1326.

[0112] 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.

[0113] A continuación, se describirá un ejemplo de una corriente de bits que se puede usar para la codificación/decodificación de audio descrita en esta invención con referencia a las Figs. 14a y 14b. Se debería observar que la corriente de bits, por ejemplo, puede ser una extensión de la corriente de bits usada en la 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 MPEG surround 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.

[0114] 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 y la carga útil MPS para 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.

[0115] 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.

[0116] Sin embargo, también se pueden usar naturalmente diferentes formatos de corriente de bits.

12. Entorno de codificación/decodificación

[0117] A continuación, se describirá un entorno de codificación/decodificación de audio en el que se pueden aplicar los conceptos según la presente invención.

[0118] Sistema de códec de audio en 3D, en el que se pueden utilizar los conceptos según la presente invención, se basa en un códec de MPEG–D USAC para la decodificación de señales de canal y objeto. Para aumentar la eficiencia de 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 mediante el uso de SAOC, la información de metadatos de objeto correspondiente se comprime y multiplexa en la corriente de bits de audio en 3D.

[0119] La Fig. 15 muestra un diagrama esquemático de bloques de tal codificador de audio, y la Fig. 16 muestra un diagrama esquemático de bloques de tal decodificador de audio. En otras palabras, las Figs. 15 y 16 muestran los diferentes bloques algorítmicos del sistema de audio en 3D.

[0120] Tomando como referencia la Fig. 15, que muestra un diagrama esquemático de bloques de un codificador de audio en 3D 1500, se explicarán algunos detalles. El codificador 1500 comprende un pre-renderizador/mezclador 1510 opcional, que recibe una o más señales del canal 1512 y una o más señales del objeto 1514 y proporciona, sobre esta base, una o más señales del canal 1516 así como una o más señales del objeto 1518, 1520. El codificador de audio comprende también 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 proporcionados al codificador SAOC. Además, el codificador USAC 1530 está configurado para recibir las señales del canal 1516 que comprende canales y objetos pre-renderizados desde el pre-renderizador/mezclador, para recibir una o más señales del objeto 1518 desde el pre-renderizador/mezclador y para recibir uno o más canales de transporte SAOC 1542 e información secundaria SAOC 1544, y proporciona, sobre esta base, una representación codificada 1532. Además, el codificador de audio 1500 comprende también un codificador de metadatos del objeto 1550 que se configura para recibir los metadatos del objeto 1552 (que se puede evaluar mediante el pre-renderizador/mezclador 1510) y para codificar los metadatos del objeto para obtener metadatos del objeto codificado 1554. Los metadatos codificados son recibidos también por el codificador USAC 1530 y utilizados para proporcionar la representación codificada 1532.

[0121] Algunos detalles con respecto a los componentes individuales del codificador de audio 1500 se describirán a continuación.

[0122] Tomando como referencia 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 esta base, 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).

[0123] El decodificador de audio 1600 comprende un decodificador USAC 1620, y proporciona una o más señales del canal 1622, una o más señales del objeto pre-renderizadas 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 comprende también un renderizador del objeto 1640 que está configurado para proporcionar una o más señales del objeto renderizadas 1642 sobre la base de la señal del objeto 1626 y una información de metadatos del objeto 1644, en la que la información de metadatos del objeto 1644 es proporcionada por un decodificador de metadatos del objeto 1650 sobre la base de la información de metadatos del objeto comprimida 1632. El decodificador de audio 1600 comprende también, opcionalmente, un decodificador de SAOC 1660, que se configura para recibir el canal de transporte de SAOC 1628 y la información secundaria SAOC 1630, y para proporcionar, sobre esta base, una o más señales del objeto renderizadas 1662. El decodificador de audio 1600 comprende también un mezclador 1670, que se configura para recibir las señales del canal 1622, las señales del objeto pre-renderizadas 1624, las señales del objeto renderizadas 1642, y las señales del objeto renderizadas 1662, y para proporcionar, sobre esta base, 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, puede comprender también un renderizador binaural 1680, que se configura para recibir las señales del canal mixtas 1672 y para proporcionar, sobre esta base, las señales para los auriculares 1614. Además, el decodificador de audio 1600 puede comprender una conversión de formato 1690, que se configura para recibir las señales del canal mixtas 1672 y una información del diseño de reproducción 1692 y para proporcionar, sobre esta base, una señal del altavoz 1616 para una configuración de altavoz alternativa.

[0124] A continuación, se describirán algunos detalles de los componentes del codificador de audio 1500 y del decodificador de audio 1600.

Pre-renderizador/mezclador

[0125] El pre-renderizador/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 pre-renderización 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 pre-renderización de los 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 utilizar. Los pesos de los objetos para canal se obtienen de los metadatos del objeto asociados (OAM) 1552.

[0126] Códec del núcleo USAC

[0126] 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 pre-renderizadas se basa 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 asignación de objetos basada en la información geométrica y semántica de canal de la entrada y la asignación de objetos. Esta información de asignación describe cómo los canales de entrada y los objetos se asignan a elementos del canal USAC (CPE, SCE, LFE) y la información correspondiente se transmite al decodificador. Todas las cargas útiles adicionales como datos SAOC o metadatos de objetos han pasado a través de los elementos de extensión y se han considerado en el control de la frecuencia de los codificadores.

10 **[0127]** La codificación de objetos es posible en diferentes formas, según los requisitos de frecuencia/distorsión y los requisitos de interactividad para el renderizador. Las siguientes variantes de codificación del objeto son posibles:

1. Objetos pre-renderizados: las señales del objeto se pre-renderizan y mezclan con las señales del canal 22.2 antes de codificar. La subsiguiente cadena de codificación ve señales del canal 22.2.

15 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 a lo largo.

20 3. Formas de onda del objeto paramétrico: las propiedades del objeto y la relación entre sí se describen por medio de los parámetros de SAOC. La mezcla descendente de las señales del objeto se codifica con USAC. La información paramétrica se transmite a lo largo. El número de canales de mezcla descendente se elige dependiendo del número de objetos y la tasa de datos global. La información de metadatos del objeto comprimida se transmite al renderizador SAOC.

25

SAOC

[0128] El codificador SAOC 1540 y el decodificador de SAOC 1660 para señales de objeto se basan en la tecnología MPEG SAOC. El sistema es capaz de volver a crear, modificar y renderizar una serie de objetos de audio sobre la base de un número menor de canales transmitidos y datos paramétricos adicionales (diferencias de nivel de objeto OLD, correlaciones inter-objetos IOC, ganancias de la mezcla descendente DMG). Los datos paramétricos adicionales exhiben una velocidad de datos significativamente inferior que la requerida para la transmisión de todos los objetos individualmente, lo que produce una codificación muy eficiente. 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 en 3D 1532, 1610) y los canales de transporte SAOC (que se codifican mediante el uso de elementos de canal único y se transmiten).

[0129] El decodificador de 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 sobre la base del diseño de reproducción, la información de metadatos del objeto descomprimida y, opcionalmente, de la información de la interacción del usuario.

Códec de metadatos del objeto

45 **[0130]** Para cada objeto, los metadatos asociados que especifican la posición geométrica y el volumen del objeto en el espacio en 3D se codifica de manera eficiente 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.

50 Mezclador/renderizador del objeto

[0131] El renderizador del objeto utiliza los metadatos de objeto comprimidos para generar 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 decodifican el contenido basado en el canal, así como los objetos discretos / paramétricos, las formas de onda basadas en 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 un módulo de post-procesador como el renderizador binaural o el módulo renderizador del altavoz).

Renderizador binaural

60

[0132] El módulo del renderizador binaural 1680 produce una mezcla descendente binaural del material de

audio multicanal, de modo que cada canal de entrada está representado por una fuente de sonido virtual. El procesamiento se realiza por cuadro en un dominio QMF. La binauralización se basa en las respuestas de impulso del espacio binaural medido.

5 Renderizador de altavoz/conversión de formato

[0133] 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” a continuación. 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 estándar.

[0134] La Fig. 17 muestra un diagrama esquemático de 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 altavoz 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 del formato de salida del mezclador 1732 y una información de formato de reproducción 1734.

[0135] Además, se debería observar que los conceptos descritos anteriormente, por ejemplo el codificador de audio 100, el decodificador de audio 200 ó 300, el codificador de audio 400, el decodificador de audio 500 ó 600, los procedimientos 700, 800, 900, ó 1000, el codificador de audio 1100 ó 1200 y el decodificador de audio 1300 se pueden usar dentro del codificador de audio 1500 y/o dentro del decodificador de audio 1600. Por ejemplo, los codificadores/decodificadores de audio mencionados anteriormente se pueden usar para codificar o decodificar las señales del canal que se asocian con posiciones espaciales diferentes.

30 13. Realizaciones alternativas

[0136] A continuación se describirán algunas realizaciones adicionales.

[0137] Tomando como referencia las Figs. 18 a 21, se explicarán realizaciones adicionales según la invención.

[0138] Se debería observar que un llamado “elemento del canal cuádruple” (QCE) se puede considerar como una herramienta de un decodificador de audio, que se puede usar, por ejemplo, para decodificar el contenido de audio tridimensional.

[0139] En otras palabras, el elemento de canal cuádruple (QCE) es un procedimiento para la codificación conjunta de cuatro canales para la codificación más eficiente 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 conjunta con la posibilidad de la herramienta de predicción de estéreo compleja en dirección horizontal y la herramienta de estéreo basada en MPEG Surround en dirección vertical. Esto se obtiene al permitir ambas herramientas de estéreo e intercambiar los canales de salida entre la aplicación de las herramientas. SBR estéreo se realiza en dirección horizontal para preservar las relaciones izquierda–derecha de las altas frecuencias.

[0140] La Fig. 18 muestra una estructura topológica de un QCE. Se debería observar que el QCE de la Fig. 18 es muy similar al QCE de la Fig. 11, de modo tal que se hace referencia a las explicaciones anteriores. Sin embargo, se debería observar 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, tal 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 inferior derecho 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.

[0141] A continuación, se proporcionarán algunos términos y definiciones, que se pueden aplicar en algunas realizaciones.

[0142] Un elemento de datos qcelIndex indica un modo QCE de un CPE. Con respecto al significado de la

variable de corriente de bits `qcelIndex`, se hace referencia a la Fig. 14b. Se debería observar 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 dan en la Fig. 14b. El `qcelIndex` será el mismo para los dos elementos posteriores que forman una QCE.

5 **[0143]** A continuación, se definirán algunos elementos de ayuda, que se pueden usar en algunas realizaciones según la invención:

| | | |
|----|-------------------------------|---|
| 10 | <code>cplx_out_dmx_L[]</code> | primer canal del primer CPE después de la decodificación estéreo para predicción compleja |
| | <code>cplx_out_dmx_R[]</code> | segundo canal del primer CPE después de la decodificación estéreo para predicción compleja |
| | <code>cplx_out_res_L[]</code> | segundo CPE después de la decodificación estéreo para predicción compleja (cero si <code>qcelIndex = 1</code>) |
| 15 | <code>cplx_out_res_R[]</code> | segundo canal del segundo CPE después de la decodificación estéreo para predicción compleja (cero si <code>qcelIndex = 1</code>) |
| | <code>mps_out_L_1[]</code> | primer canal de salida de la primera caja MPS |
| | <code>mps_out_L_2[]</code> | segundo canal de salida de la primera caja MPS |
| | <code>mps_out_R_1[]</code> | primer canal de salida de la segunda caja MPS |
| 20 | <code>mps_out_R_2[]</code> | segundo canal de salida de la segunda caja MPS |
| | <code>sbr_out_L_1[]</code> | primer canal de salida de la primera caja SBR estéreo |
| | <code>sbr_out_R_1[]</code> | segundo canal de salida de la primera caja SBR estéreo |
| | <code>sbr_out_L_2[]</code> | primer canal de salida de la segunda caja SBR estéreo |
| | <code>sbr_out_R_2[]</code> | segundo canal de salida de la segunda caja SBR estéreo |

25 **[0144]** A continuación, se explicará un proceso de decodificación, que se realiza en una realización según la invención.

30 **[0145]** 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 se utiliza siempre para el QCE, por lo tanto el elemento de sintaxis `stereoConfigIndex` será 3 y `bsStereoSbr` será 1.

35 **[0146]** En caso de `qcelIndex == 1` solo las cargas útiles para MPEG Surround y SBR y ningún dato de la señal de audio relevante está contenido en el segundo CPE y el elemento de sintaxis `bsResidualCoding` se ajusta a 0.

[0147] 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.

40 **[0148]** Sin embargo, también se pueden usar algunos esquemas de señalización simplificados diferentes y posibles.

45 **[0149]** La decodificación de Joint Stereo con posibilidad de predicción estéreo compleja se realiza como se describe en ISO/IEC 23003-3, sub-cláusula 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 usa 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.

50 **[0150]** Antes de aplicar la decodificación MPEG Surround, el segundo canal del primer elemento (`cplx_out_dmx_R[]`) y el primer canal del segundo (`cplx_out_res_L[]`) se intercambian.

55 **[0151]** La decodificación de MPEG Surround se realiza como se describe en ISO/IEC 23003-3, se usa la sub-cláusula 7.11. Si se usa la codificación residual, la decodificación, sin embargo, se puede modificar cuando se compara con la decodificación de MPEG Surround convencional en algunas realizaciones. La decodificación de MPEG Surround sin la residual mediante el uso de SBR como se define en ISO/IEC 23003-3, sub-cláusula 7.11.2.7 (figura 23), se modifican de modo que SBR estéreo también se usa para `bsResidualCoding == 1`, que resulta en el esquema del decodificador mostrado en la Fig. 19. La Fig. 19 muestra un diagrama esquemático de bloques de un codificador de audio para `bsResidualCoding == 0` y `bsStereoSbr == 1`.

60 **[0152]** Como se puede observar en la Fig. 19, un decodificador de núcleo USAC 2010 proporciona una señal

de la mezcla descendente (DMX) 2012 a un decodificador MPS (MPEG Surround) 2020, que proporciona una primera señal de audio decodificada 2022 y una segunda señal de audio decodificada 2024. Un decodificador SBR estéreo 2030 recibe la primera señal de audio decodificada 2022 y la segunda señal de audio decodificada 2024 y proporciona, sobre esta base una señal de audio extendida de ancho de banda izquierda 2032 y una señal de audio extendida de ancho de banda derecha 2034.

[0153] Antes de aplicar 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 de nuevo para restaurar el orden del canal de entrada.

[0154] Una estructura del decodificador QCE se ilustra en la Fig. 20, que muestra un esquema del decodificador de QCE.

[0155] Se debería observar que el diagrama esquemático de bloques de la Fig. 20 es muy similar al diagrama esquemático de bloques de la Fig. 13, de modo que también se hace referencia a las explicaciones anteriores. Además, se debería observar 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.

[0156] La Fig. 21 muestra un diagrama esquemático de bloques de un codificador de canal cuádruple 2200, según una realización de la presente invención. En otras palabras, un codificador de canal cuádruple (elemento del canal cuádruple), que se puede considerar como una herramienta Codificador Core, se ilustra en la Fig. 21.

[0157] 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 esta base, 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 un segundo 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 esta base, una primera carga útil de SBR 2225, una primera señal de salida de SBR del canal izquierdo 2226 y una primera señal de salida de SBR del canal derecho 2228.

[0158] El codificador de canal cuádruple 2200 comprende un primer codificador multicanal tipo MPEG-Surround (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 esta base, una primera carga útil MPS 2232, una señal de la mezcla descendente MPEG Surround del canal izquierdo 2234 y, opcionalmente, una señal residual de MPEG Surround del canal izquierdo 2236. El codificador de canal cuádruple 2200 comprende también un segundo codificador multicanal tipo MPEG-Surround (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 esta base, una primer carga útil MPS 2242, una señal de la mezcla descendente MPEG Surround del canal derecho 2244 y, opcionalmente, una señal residual de MPEG Surround del canal derecho 2246.

[0159] El codificador de canal cuádruple 2200 comprende un primer codificador estéreo de predicción compleja 2250, que recibe la señal de la mezcla descendente MPEG Surround del canal izquierdo 2234 y la señal de la mezcla descendente MPEG Surround del canal derecho 2244, y que proporciona, sobre esta base, una carga útil de predicción compleja 2252 y una representación codificada de forma conjunta 2254 de la señal de la mezcla descendente MPEG Surround del canal izquierdo 2234 y la señal de la mezcla descendente MPEG Surround del canal derecho 2244. El codificador de canal cuádruple 2200 comprende un segundo codificador estéreo de predicción compleja 2260, que recibe la señal residual de MPEG Surround del canal izquierdo 2236 y la señal residual de MPEG Surround del canal derecho 2246, y que proporciona, sobre esta base, una carga útil de predicción compleja 2262 y una representación codificada de forma conjunta 2264 de la señal de la mezcla descendente MPEG Surround del canal izquierdo 2236 y la señal de la mezcla descendente MPEG Surround del canal derecho 2246.

[0160] El codificador de canal cuádruple también comprende un primer codificador de la corriente de bits 2270, que recibe la representación codificada de forma conjunta 2254, la carga útil de predicción compleja 2252m, la carga útil MPS 2232 y la carga útil de SBR 2215 y proporciona, sobre esta base, una porción de la corriente de bits que representa un primer elemento del par de canales. El codificador de canal cuádruple comprende también un segundo codificador de la corriente de bits 2280, que recibe la representación codificada de 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 esta base, una porción de la corriente de bits que representa un primer elemento del par de canales.

14. Alternativas de implementación

5

[0161] Aunque algunos aspectos se han descrito en el contexto de un aparato, está claro que estos aspectos también representan una descripción del procedimiento correspondiente, donde 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.

10

[0162] 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.

15

[0163] Dependiendo de ciertos requisitos 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 tal manera que se lleva a cabo el procedimiento respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

20

25

[0164] 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 tal manera que se lleva a cabo uno de los procedimientos descritos en esta invención.

30

[0165] En general, las realizaciones de la presente invención se pueden implementar como un producto del programa informático con un código del programa, el código del programa es operativo para llevar a cabo uno de los procedimientos, cuando el producto del programa informático se ejecuta en un ordenador. El código del programa por ejemplo, se puede almacenar en un soporte legible por máquina.

35

[0166] Otras realizaciones comprenden el programa informático para realizar uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por máquina.

40

[0167] En otras palabras, una realización del procedimiento de la invención es, por 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.

45

[0168] Una realización adicional de los procedimientos de la invención es, por 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 de grabado son típicamente tangibles y / o no transitorios.

50

[0169] Una realización adicional del procedimiento de la invención es, por tanto, 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 transferida a través de una conexión de comunicación de datos, por ejemplo a través de Internet.

55

[0170] Una realización adicional comprende un medio de procesamiento, por ejemplo un ordenador, o un dispositivo lógico programable, configurado para o adaptado para llevar a cabo uno de los procedimientos descritos en esta invención.

[0171] Una realización adicional comprende un ordenador que tiene instalado el programa informático para realizar uno de los procedimientos descritos en esta invención.

60

[0172] Una realización adicional según la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, de 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.

5 **[0173]** En algunas realizaciones, un dispositivo lógico programable (por ejemplo, una matriz de puertas programable de campo) se puede usar para llevar a cabo todas o algunas de las funcionalidades de los procedimientos descritos en esta invención. En algunas realizaciones, un campo de matriz de puertas programable de campo puede cooperar con un microprocesador para llevar a cabo uno de los procedimientos descritos en esta invención. En general, los procedimientos se llevan a cabo preferentemente en cualquier aparato de hardware.

10

[0174] Las realizaciones descritas anteriormente 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. Se pretende, por tanto, que la invención solo esté 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 esta invención.

15

15. Conclusiones

[0175] A continuación, se proporcionarán algunas conclusiones.

20

[0176] Las realizaciones según la invención se basan en la consideración de que, para tener en cuenta las dependencias de señal entre los canales distribuidos de forma vertical y horizontal, cuatro canales se pueden codificar de 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 MDCT, que incluye la posibilidad de codificación izquierda-derecha y semi-lateral. Si las señales residuales están presentes, se combinan horizontalmente utilizando el mismo procedimiento.

25

[0177] Por otra parte, se debería observar 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 en 3D, en el que los canales de altavoces se distribuyen en 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 considerar las relaciones espaciales y perceptuales entre los canales. Sin embargo, este problema se supera mediante las realizaciones según la invención.

30

[0178] Por otra parte, el MPEG surround convencional se aplica en una etapa de pre-/post-procesamiento adicional, de tal manera que las señales residuales se transmiten de forma individual 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 eficiente, mediante el uso de dichas dependencias.

35

[0179] 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.

40

Referencias:

[0180]

45

- [1] ISO/IEC 23003-3: 2012 – Information Technology – MPEG Audio Technologies, Part 3: Unified Speech and Audio Coding;
- [2] ISO/IEC 23003-1: 2007 – Information Technology – MPEG Audio Technologies, Part 1: MPEG Surround

REIVINDICACIONES

1. Un decodificador de audio (200; 300; 600; 1300; 1600; 2000) para proporcionar al menos cuatro señales de canal de audio (220, 222, 224, 226; 320, 322, 324, 326; 620, 622, 624, 626; 1320, 1322, 1324, 1326) sobre la base de una representación codificada (210; 310, 360; 610, 682; 1310,1312; 1610),
 5 en la que el decodificador de audio se configura para proporcionar una primera señal residual (232; 332; 684; 1362) y una segunda señal residual (234; 334; 686; 1364) sobre la base de una representación codificada conjuntamente (210; 310; 682; 1312) de la primera señal residual y de la segunda señal residual mediante el uso de una
 10 decodificación multicanal (230; 330; 680; 1360) que explota las similitudes y/o dependencias entre las señales residuales;
 en la que el decodificador de audio se configura para proporcionar una primera señal del canal de audio (220; 320; 642; 1372) y una segunda señal del canal de audio (222; 322; 644; 1374) sobre la base de una primera señal de mezcla descendente (212; 312; 632; 1342) y la primera señal residual mediante el uso de una decodificación
 15 multicanal asistida por señales residuales (240; 340; 640; 1370); y
 en la que el decodificador de audio se configura para proporcionar una tercera señal de canal de audio (224; 324; 656; 1382) y una cuarta señal de canal de audio (226; 326; 658; 1384) sobre la base de una segunda señal de mezcla descendente (214; 314; 634; 1344) y la segunda señal residual mediante el uso de una decodificación
 20 multicanal asistida por señales residuales (250; 350; 650; 1380).
2. El decodificador de audio según la reivindicación 1, en el que el decodificador de audio se configura para proporcionar la primera señal de mezcla descendente (212; 312; 632; 1342) y la segunda señal de mezcla descendente (214; 314; 634; 1344) sobre la base de una representación codificada conjuntamente (360; 610; 1310) de la primera señal de mezcla descendente y la segunda señal de mezcla descendente mediante el uso de una
 25 decodificación multicanal (370; 630; 1340).
3. El decodificador de audio según la reivindicación 1 o la reivindicación 2, en el que el decodificador de audio se configura para proporcionar la primera señal residual y la segunda señal residual sobre la base de la representación codificada conjuntamente de la primera señal residual y de la segunda señal residual mediante el uso
 30 de una decodificación multicanal a base de predicciones.
4. El decodificador de audio según una de las reivindicaciones 1 a 3, en el que el decodificador de audio se configura para proporcionar la primera señal residual y la segunda señal residual sobre la base de la representación codificada conjuntamente de la primera señal residual y de la segunda señal residual mediante el uso
 35 de una decodificación multicanal asistida por señales residuales.
5. El decodificador de audio según la reivindicación 3, en el que la decodificación multicanal a base de predicciones se configura para evaluar un parámetro de predicción que describe una contribución de un componente de señal, que se deriva mediante el uso de un componente de señal de un marco previo, a la provisión de las
 40 señales residuales del marco actual.
6. El decodificador de audio según una de las reivindicaciones 3, reivindicación 4, en cuanto a dependiente de la reivindicación 3 y la reivindicación 5, en el que la decodificación multicanal a base de predicciones se configura para obtener la primera señal residual y la segunda señal residual sobre la base de una señal de
 45 mezcla descendente de la primera señal residual y de la segunda señal residual y sobre la base de una señal residual común de la primera señal residual y la segunda señal residual.
7. El decodificador de audio según la reivindicación 6, en el que la decodificación multicanal a base de predicciones se configura para aplicar la señal residual común con un primer signo para obtener la primera señal residual y para aplicar la señal residual común con un segundo signo, que se opone al primer signo, para obtener la
 50 segunda señal residual.
8. El decodificador de audio según una de las reivindicaciones 1 a 7, en el que el decodificador de audio se configura para proporcionar la primera señal residual y la segunda señal residual sobre la base de la representación codificada conjuntamente de la primera señal residual y de la segunda señal residual mediante el uso
 55 de una decodificación multicanal que es operativa en un dominio MDCT.
9. El decodificador de audio según una de las reivindicaciones 1 a 8, en el que el decodificador de audio se configura para proporcionar la primera señal residual y la segunda señal residual sobre la base de la representación codificada conjuntamente de la primera señal residual y de la segunda señal residual mediante el uso
 60 de una predicción estéreo del complejo USAC, significando USAC discurso unificado y codificación de audio.

10. El decodificador de audio según una de las reivindicaciones 1 a 9,
 en el que el decodificador de audio se configura para proporcionar la primera señal del canal de audio y la segunda
 5 señal del canal de audio sobre la base de la primera señal de mezcla descendente y la primera señal residual
 mediante el uso de una decodificación multicanal asistida por señales residuales en base a parámetros; y
 en el que el decodificador de audio se configura para proporcionar 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 y la segunda señal residual
 mediante el uso de una decodificación multicanal asistida por señales residuales en base a parámetros.
- 10 11. El decodificador de audio según la reivindicación 10, en el que la decodificación multicanal asistida por
 señales residuales en base a parámetros se configura 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
 15 más señales de canal de audio sobre la base de una respectiva de las señales de mezcla descendente y una
 correspondiente de las señales residuales.
12. El decodificador de audio según una de las reivindicaciones 1 a 11, en el que el decodificador de audio
 se configura para proporcionar 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 y la primera señal residual mediante el uso de una decodificación
 20 multicanal asistida por señales residuales que es operativa en un dominio QMF; y en el que el decodificador de
 audio se configura para proporcionar 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 y la segunda señal residual mediante el uso de una decodificación
 multicanal asistida por señales residuales que es operativa en el dominio QMF.
- 25 13. El decodificador de audio según una de las reivindicaciones 1 a 12, en el que el decodificador de audio
 se configura para proporcionar 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 y la primera señal residual mediante el uso de una decodificación
 MPEG Surround 2-1-2 o una decodificación estéreo unificada; y
- 30 en el que el decodificador de audio se configura para proporcionar 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 y la segunda señal residual
 mediante el uso de una decodificación MPEG Surround 2-1-2 o una decodificación estéreo unificada.
14. El decodificador de audio según una de las reivindicaciones 1 a 13, en el que la primera señal residual
 35 y la segunda señal residual se asocian con diferentes posiciones horizontales de una escena de audio o con
 diferentes posiciones azimutales de la escena de audio.
15. El decodificador de audio según una de las reivindicaciones 1 a 14, en el que la primera señal del
 canal de audio y la segunda señal del canal de audio se asocian con posiciones vecinas de forma vertical de una
 40 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 posiciones vecinas
 de forma vertical de la escena de audio.
- 45 16. El decodificador de audio según una de las reivindicaciones 1 a 15, en el que la primera señal del
 canal de audio y la segunda señal del canal de audio se asocian con una primera posición horizontal o posición
 azimutal 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 una segunda
 50 posición horizontal o posición azimutal de la escena de audio, que es diferente de la primera posición horizontal o la
 primera posición azimutal.
17. El decodificador de audio según una de las reivindicaciones 1 a 16, 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
 55 con un lado derecho de una escena de audio.
18. El codificador de audio según la reivindicación 17, en el que la primera señal del canal de audio y la
 segunda señal del canal de audio se asocian con el lado izquierdo 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 el lado derecho de la escena de audio.
- 60 19. El decodificador de audio según la reivindicación 18, en el que la primera señal del canal de audio está

asociada con una posición izquierda inferior de la escena de audio,

en el que la segunda señal del canal de audio está asociada con una posición izquierda superior de la escena de audio,

5 en el que la tercera señal del canal de audio está asociada con una posición derecha inferior de la escena de audio y en el que la cuarta señal del canal de audio está asociada con una posición derecha superior de la escena de audio.

20. El decodificador de audio según una de las reivindicaciones 1 a 19, en el que el decodificador de audio se configura 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 mediante el uso de una decodificación multicanal, en la que 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.

15 21. El decodificador de audio según una de las reivindicaciones 1 a 20, en el que el decodificador de audio se configura 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 de la segunda señal de mezcla descendente mediante el uso de una decodificación multicanal a base de predicciones.

20 22. El decodificador de audio según una de las reivindicaciones 1 a 21, en el que el decodificador de audio se configura 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 de la segunda señal de mezcla descendente mediante el uso de una decodificación multicanal en base a predicciones asistida por señales residuales.

25 23. El decodificador de audio según una de las reivindicaciones 1 a 22, en el que el decodificador de audio se configura para llevar a cabo una primera extensión del ancho de banda multicanal (660; 1390) sobre la base de la primera señal del canal de audio y la tercera señal del canal de audio y

30 en el que el decodificador de audio se configura para llevar a cabo una segunda extensión del ancho de banda multicanal (670; 1394) sobre la base de la segunda señal del canal de audio y la cuarta señal del canal de audio.

24. El decodificador de audio según la reivindicación 23, en el que el decodificador de audio se configura para llevar a cabo la primera extensión del ancho de banda multicanal a fin de obtener dos o más señales de canal de audio de ancho de banda extendido (620, 624;1320, 1324) asociados con un primer plano horizontal común o una primera elevación común de una escena de audio sobre la base de la primera señal del canal de audio y la tercera señal del canal de audio y uno o más parámetros de extensión de ancho de banda (1338), y

40 en el que el decodificador de audio se configura para llevar a cabo la segunda extensión del ancho de banda multicanal a fin de obtener dos o más señales de canal de audio de ancho de banda extendido (622, 626; 1322, 1326) asociada con un segundo plano horizontal común o una segunda elevación común de la escena de audio sobre la base de la segunda señal del canal de audio y la cuarta señal del canal de audio y uno o más parámetros de extensión de ancho de banda (1358).

45 25. El decodificador de audio según una de las reivindicaciones 1 a 24, en el que la representación codificada conjuntamente de la primera señal residual y de la segunda señal residual comprende un elemento del par de canales que comprende una señal de mezcla descendente de la primera y la segunda señal residual y una señal residual común de la primera y la segunda señal residual.

50 26. El decodificador de audio según una de las reivindicaciones 1 a 25, en el que el decodificador de audio se configura 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 mediante el uso de una decodificación multicanal,

55 en el que la representación codificada conjuntamente de la primera señal de mezcla descendente y de la segunda señal de mezcla descendente comprende un elemento del par de canales que comprende una señal de mezcla descendente de la primera y la segunda señal de mezcla descendente y una señal residual común de la primera y la segunda señal de mezcla descendente.

60 27. Un codificador de audio (100; 1100; 1200; 1500; 2100) para proporcionar una representación codificada (130; 1144, 1154; 1220, 1222; 2272, 2282) sobre la base de al menos cuatro señales de canal de audio

(110, 112, 114, 116; 1110, 1112, 1114, 1116; 1210, 1212, 1214, 1216; 2216, 2226, 2218, 2228),

en el que el codificador de audio se configura para codificar conjuntamente al menos una primera señal del canal de audio y una segunda señal del canal de audio mediante el uso de una codificación multicanal asistida por señales residuales (140; 1120; 1230; 2230), para obtener una primera señal de mezcla descendente (120; 1122; 1232; 2234) y una primera señal residual (142; 1124; 1234; 2236); y

en el que el codificador de audio se configura para codificar conjuntamente al menos una tercera señal de canal de audio y una cuarta señal de canal de audio mediante el uso de una codificación multicanal asistida por señales residuales (150; 1130; 1240; 2240), para obtener una segunda señal de mezcla descendente (122; 1132; 1242; 2244) y una segunda señal residual (152; 1134; 1244; 2246); y

en el que el codificador de audio se configura para codificar conjuntamente la primera señal residual y la segunda señal residual mediante el uso de una codificación multicanal (160; 1150; 1260; 2260) que explota las similitudes y/o dependencias entre las señales residuales, para obtener una representación codificada conjuntamente (130; 1154; 1262; 2264) de las señales residuales.

28. El codificador de audio según la reivindicación 27, en el que el codificador de audio se configura para codificar conjuntamente la primera señal de mezcla descendente y la segunda señal de mezcla descendente mediante el uso de una codificación multicanal (1140; 1250; 2250), para obtener una representación codificada conjuntamente (1144; 1252; 2254) de las señales de mezcla descendente.

29. El codificador de audio según la reivindicación 28, en el que el codificador de audio se configura para codificar conjuntamente la primera señal residual y la segunda señal residual mediante el uso de una codificación multicanal en base a predicciones, y

en el que el codificador de audio se configura para codificar conjuntamente la primera señal de mezcla descendente y la segunda señal de mezcla descendente mediante el uso de una codificación multicanal en base a predicciones.

30. El codificador de audio según una de las reivindicaciones 27 a 29, en el que el codificador de audio se configura para codificar conjuntamente al menos la primera señal del canal de audio y la segunda señal del canal de audio mediante el uso de una codificación multicanal asistida por señales residuales en base a parámetros, y

en el que el codificador de audio se configura para codificar conjuntamente al menos la tercera señal del canal de audio y la cuarta señal del canal de audio mediante el uso de una codificación multicanal asistida por señales residuales en base a parámetros.

31. El codificador de audio según una de las reivindicaciones 27 a 30, en el que la primera señal del canal de audio y la segunda señal del canal de audio se asocian con posiciones vecinas de forma vertical 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 posiciones vecinas de forma vertical de la escena de audio.

32. El codificador de audio según una de las reivindicaciones 27 a 31, en el que la primera señal del canal de audio y la segunda señal del canal de audio se asocian con una primera posición horizontal o posición azimutal 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 una segunda posición horizontal o posición azimutal de la escena de audio, que es diferente de la primera posición horizontal o posición azimutal.

33. El codificador de audio según una de las reivindicaciones 27 a 32, 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.

34. El codificador de audio según la reivindicación 33,

en el que la primera señal del canal de audio y la segunda señal del canal de audio se asocian con el lado izquierdo 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 el lado derecho de la escena de audio.

35. El decodificador de audio según la reivindicación 34, en el que la primera señal del canal de audio está asociada con una posición izquierda inferior de la escena de audio,

5 en el que la segunda señal del canal de audio está asociada con una posición izquierda superior de la escena de audio,
en el que la tercera señal del canal de audio está asociada con una posición derecha inferior de la escena de audio y
en el que la cuarta señal del canal de audio está asociada con una posición derecha superior de la escena de audio.

36. El codificador de audio según una de las reivindicaciones 27 a 35, en el que el codificador de audio se
10 configura para codificar conjuntamente la primera señal de mezcla descendente y la segunda señal de mezcla descendente mediante el uso de una codificación multicanal, para obtener una representación codificada conjuntamente de las señales de mezcla descendente, en la que 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.

15 37. Un procedimiento (800) para proporcionar al menos cuatro señales de canal de audio sobre la base de una representación codificada, que comprende:

20 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 mediante el uso de una decodificación multicanal que explota las similitudes y/o dependencias entre las señales residuales;

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 mediante el uso de una decodificación multicanal asistida por señales residuales; y

25 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 mediante el uso de una decodificación multicanal asistida por señales residuales.

38. Un procedimiento (700) para proporcionar una representación codificada sobre la base de al menos
30 cuatro señales de canal de audio, comprendiendo el procedimiento:

la codificación conjuntamente (710) de al menos una primera señal del canal de audio y una segunda señal del canal de audio mediante el uso de una codificación multicanal asistida por señales residuales, para obtener una primera señal de mezcla descendente y una primera señal residual;

35 la codificación conjuntamente (720) de al menos una tercera señal de canal de audio y una cuarta señal de canal de audio mediante el uso de una codificación multicanal asistida por señales residuales, para obtener una segunda señal de mezcla descendente y una segunda señal residual; y

40 la codificación conjuntamente (730) de la primera señal residual y la segunda señal residual mediante el uso de una codificación multicanal que explota las similitudes y/o dependencias entre las señales residuales, para obtener una representación codificada de las señales residuales.

39. Un programa informático para llevar a cabo el procedimiento según la reivindicación 37 ó 38 cuando el programa informático se ejecuta en un ordenador.

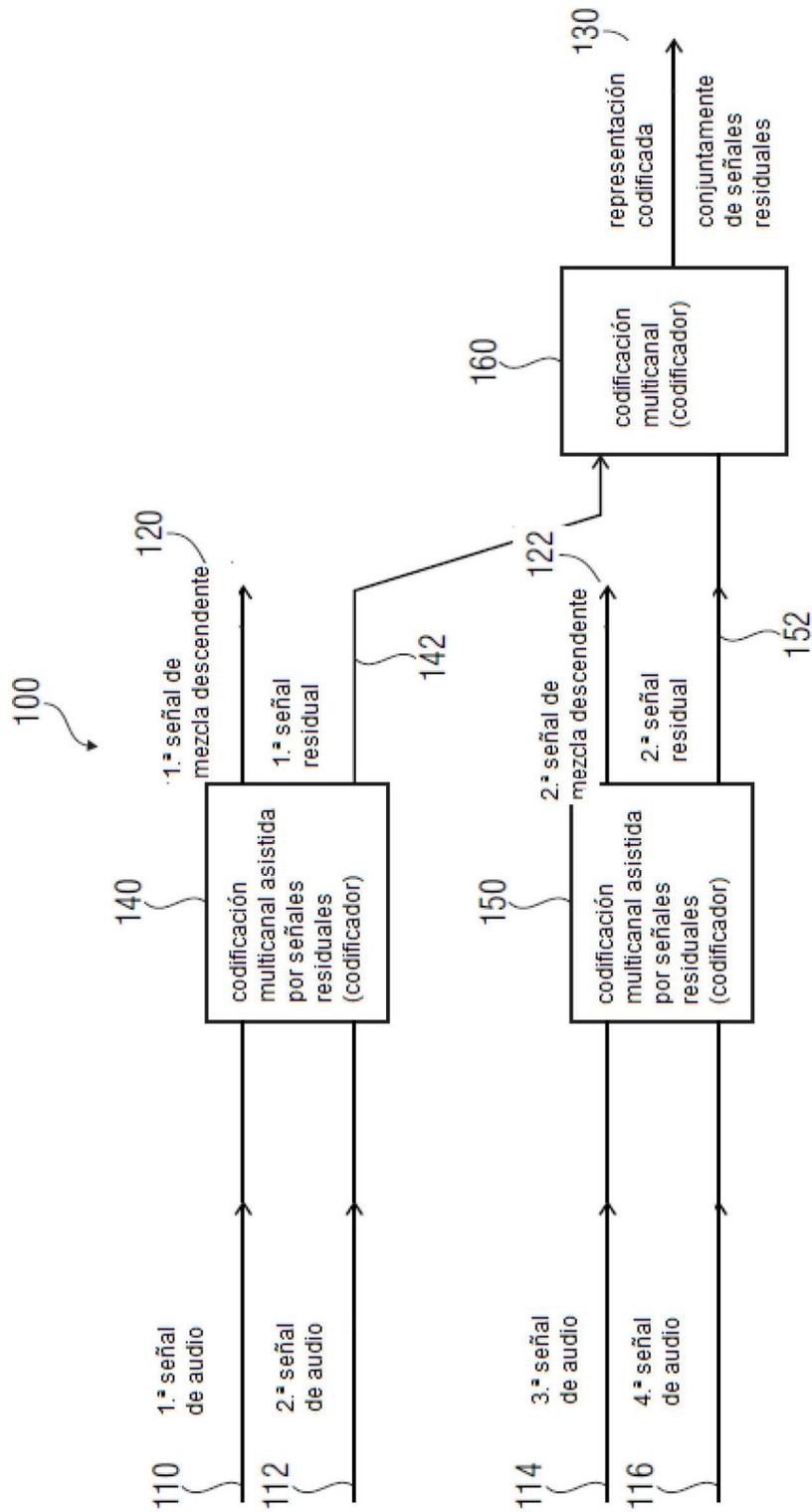


FIG. 1

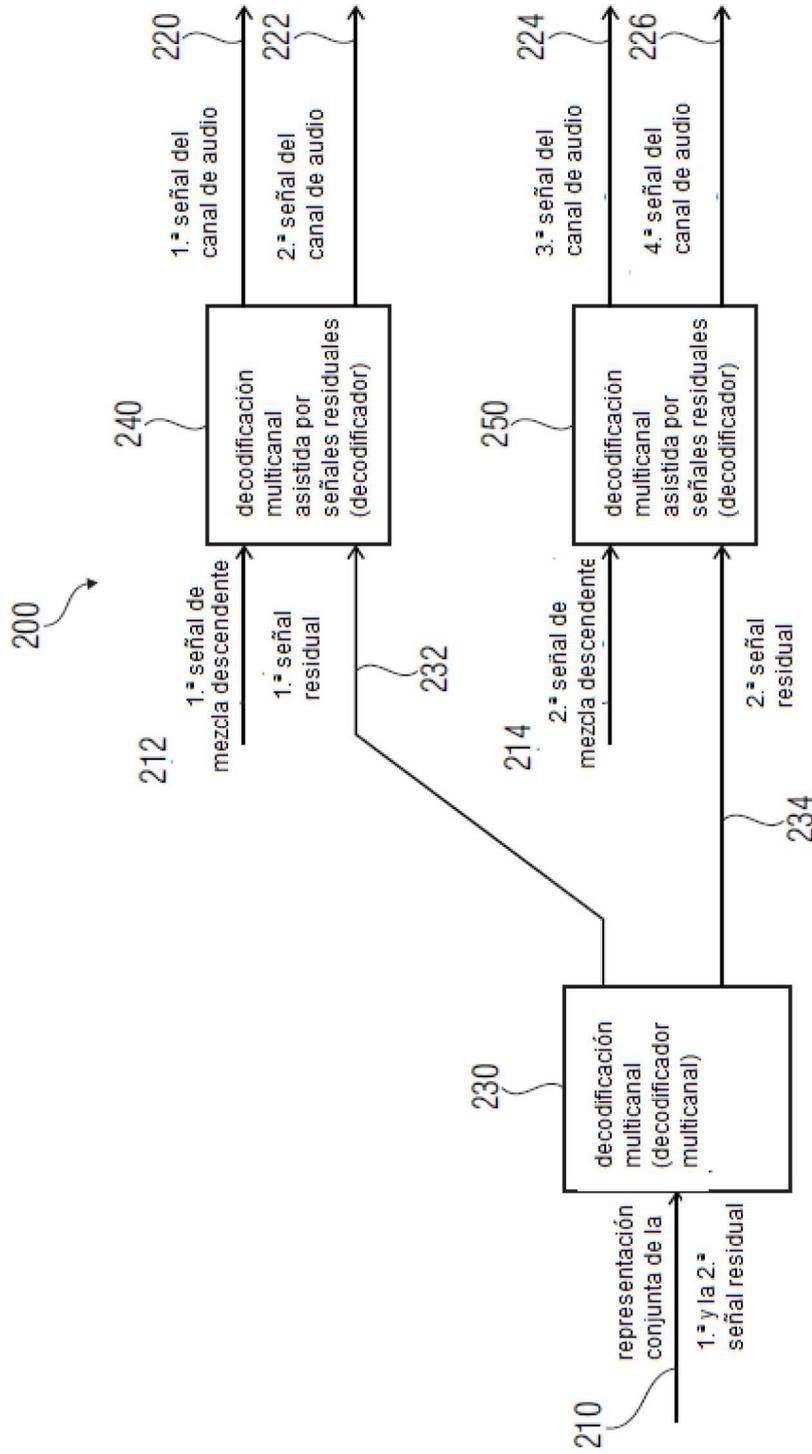


FIG. 2

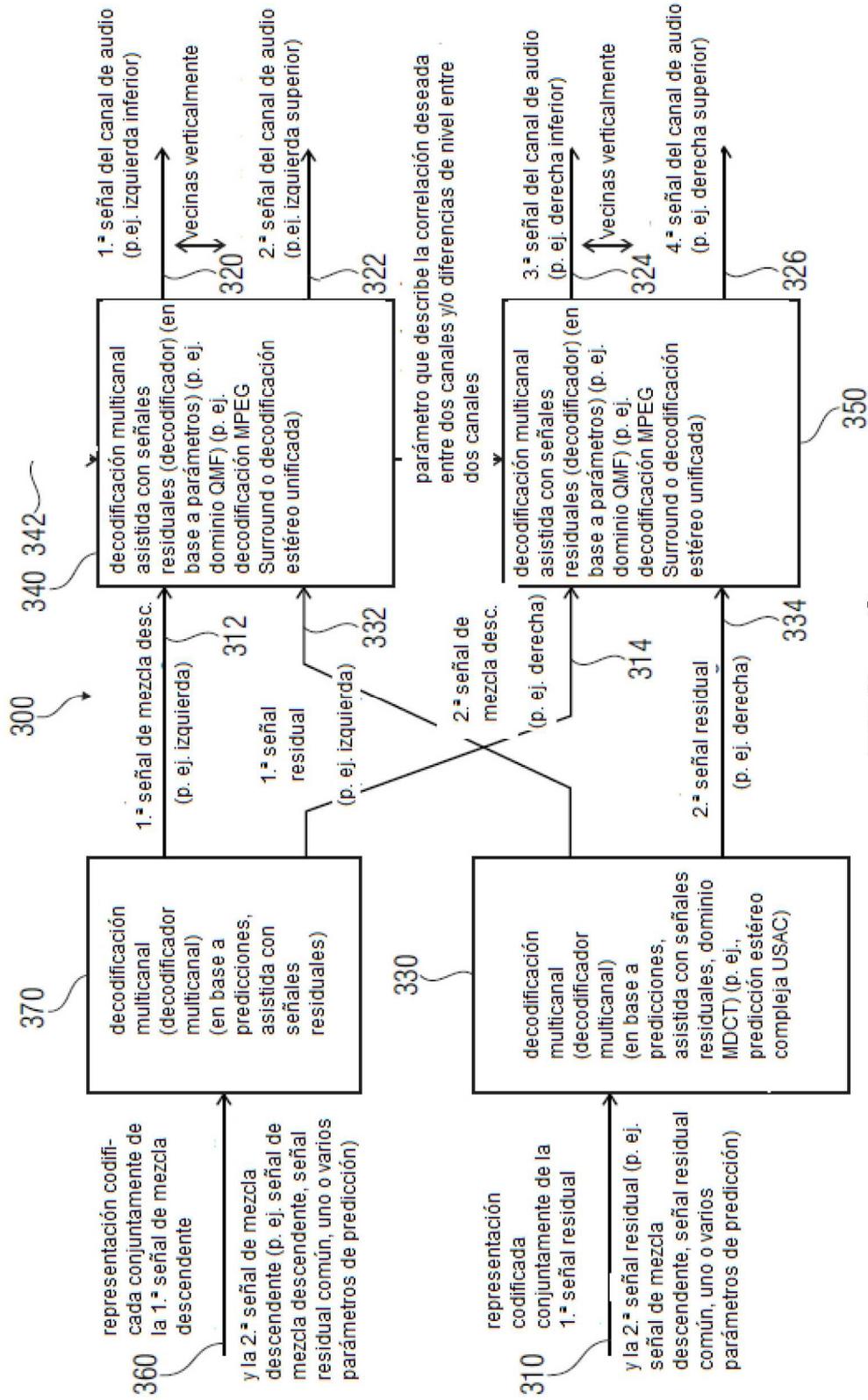


FIG. 3

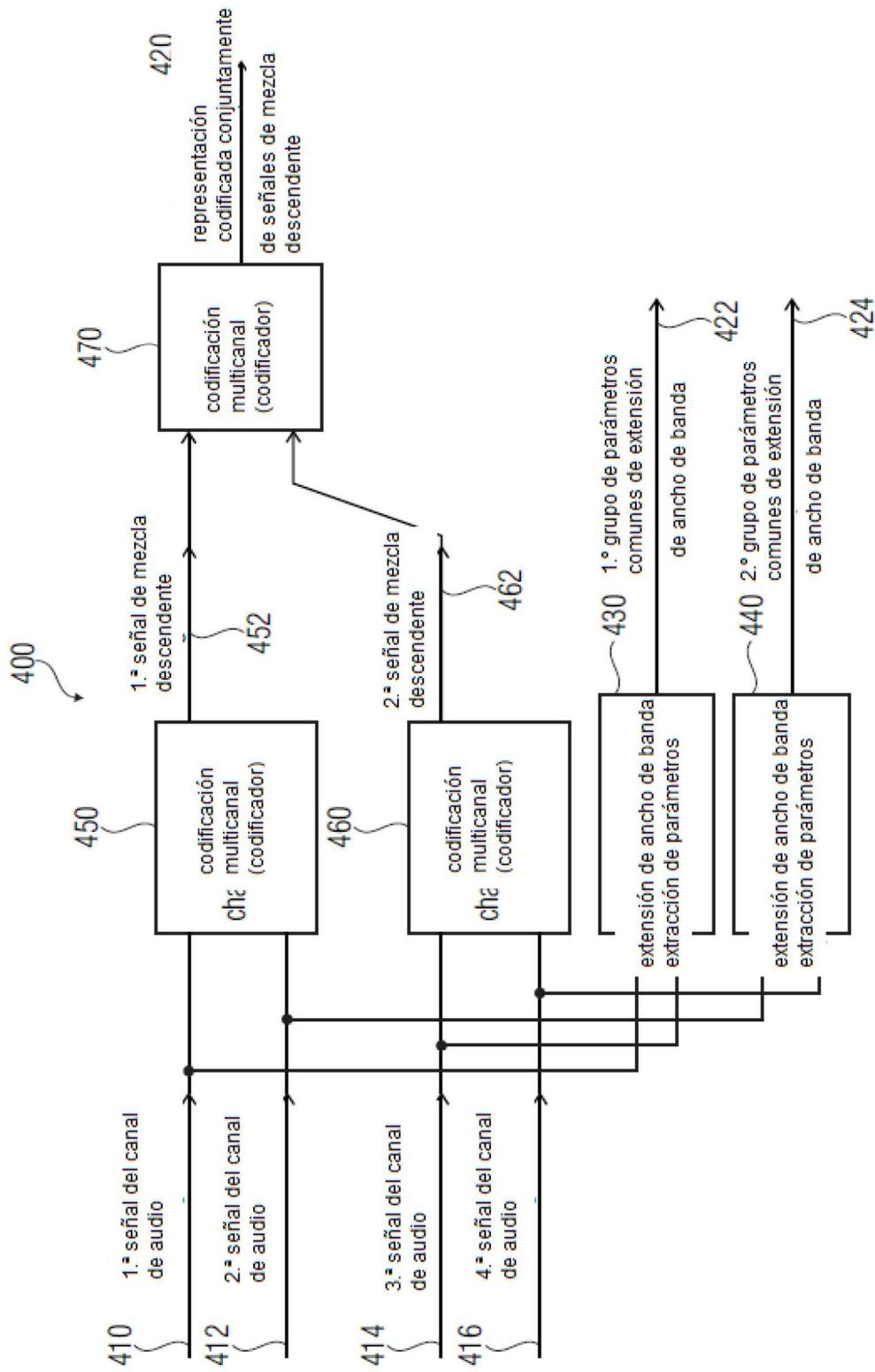


FIG. 4

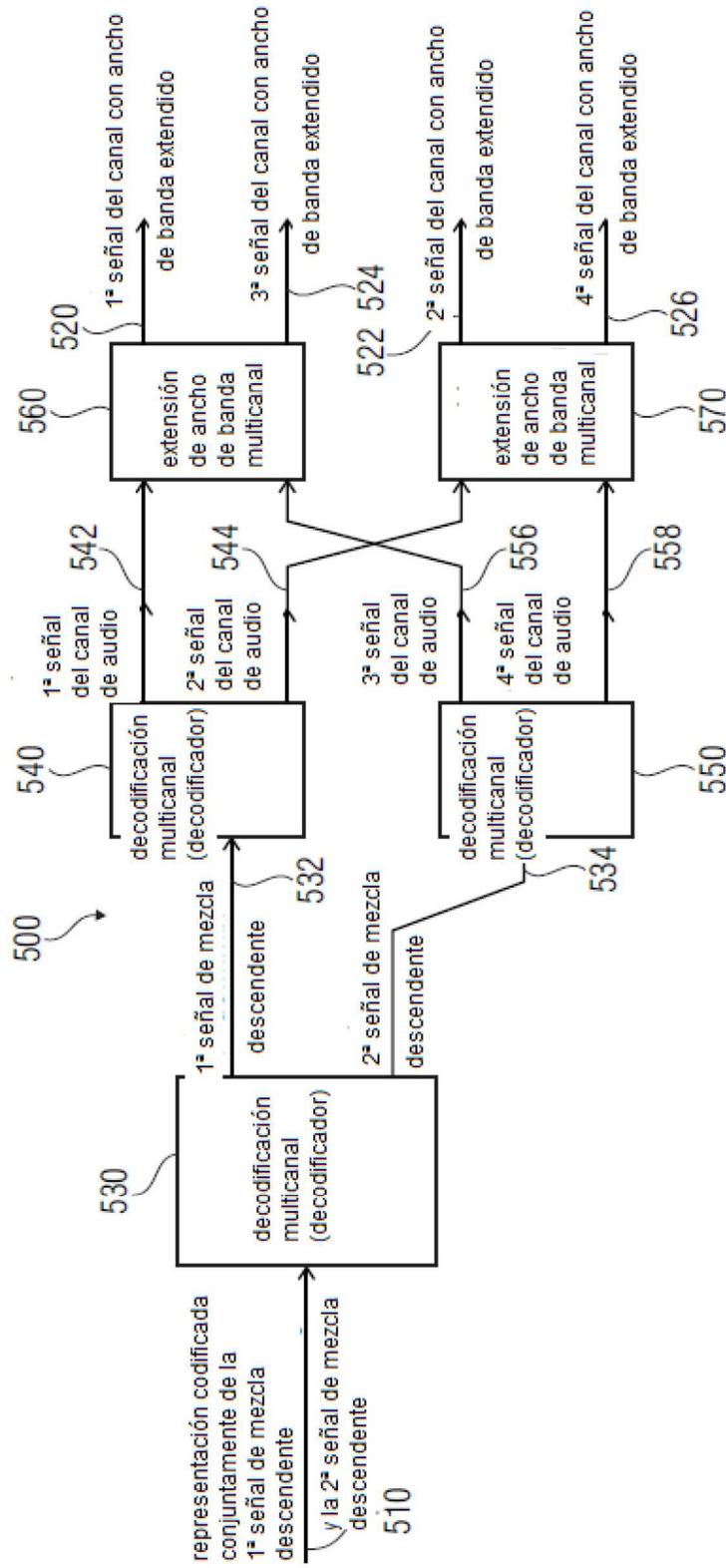


FIG. 5

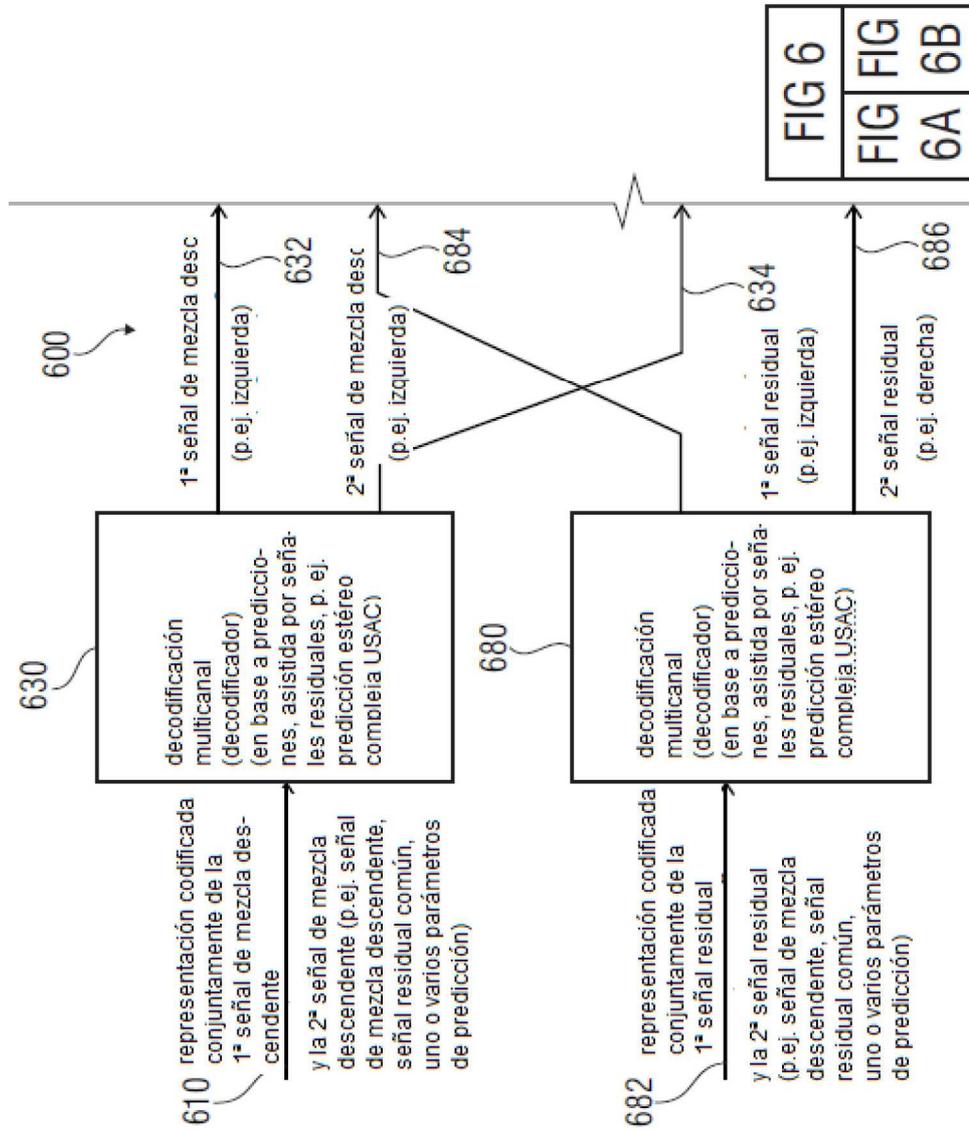


FIG 6
FIG 6A
6B

FIG. 6A

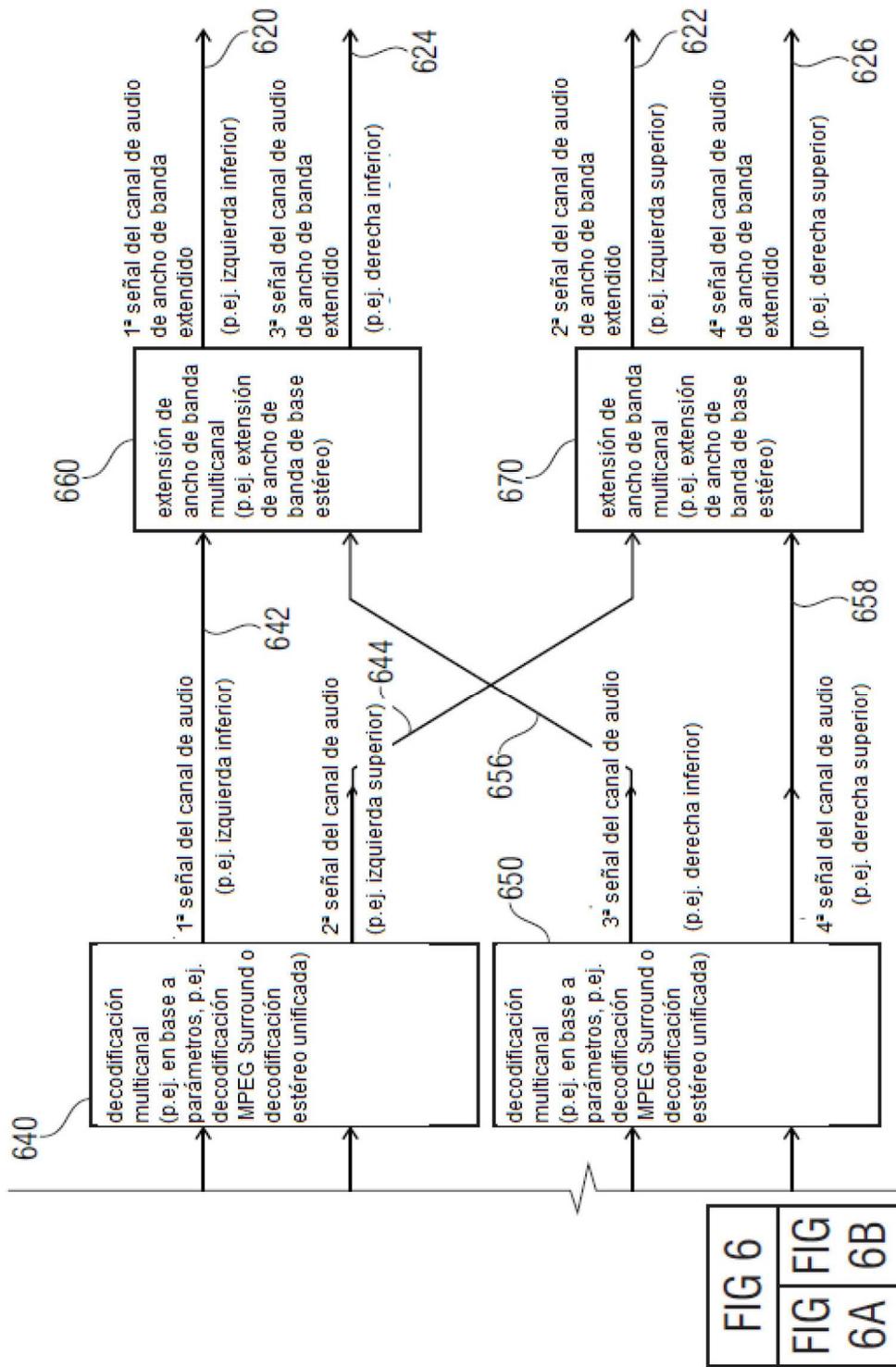


FIG. 6B

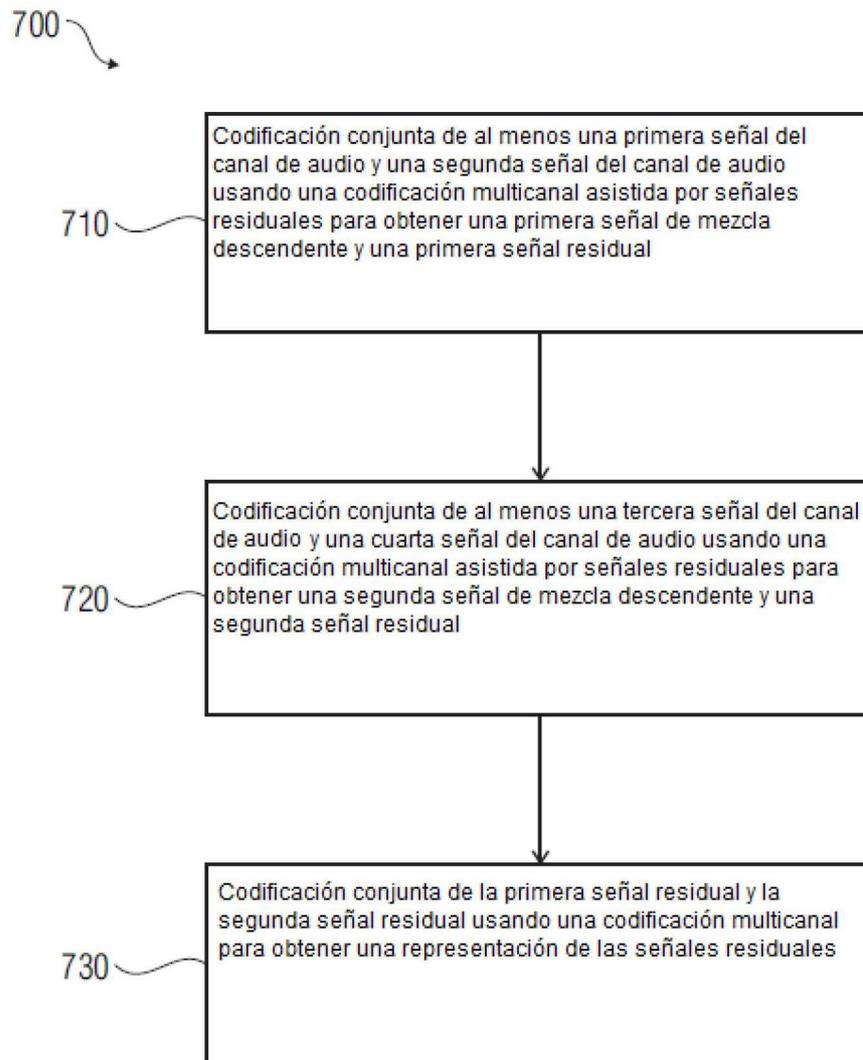


FIG. 7

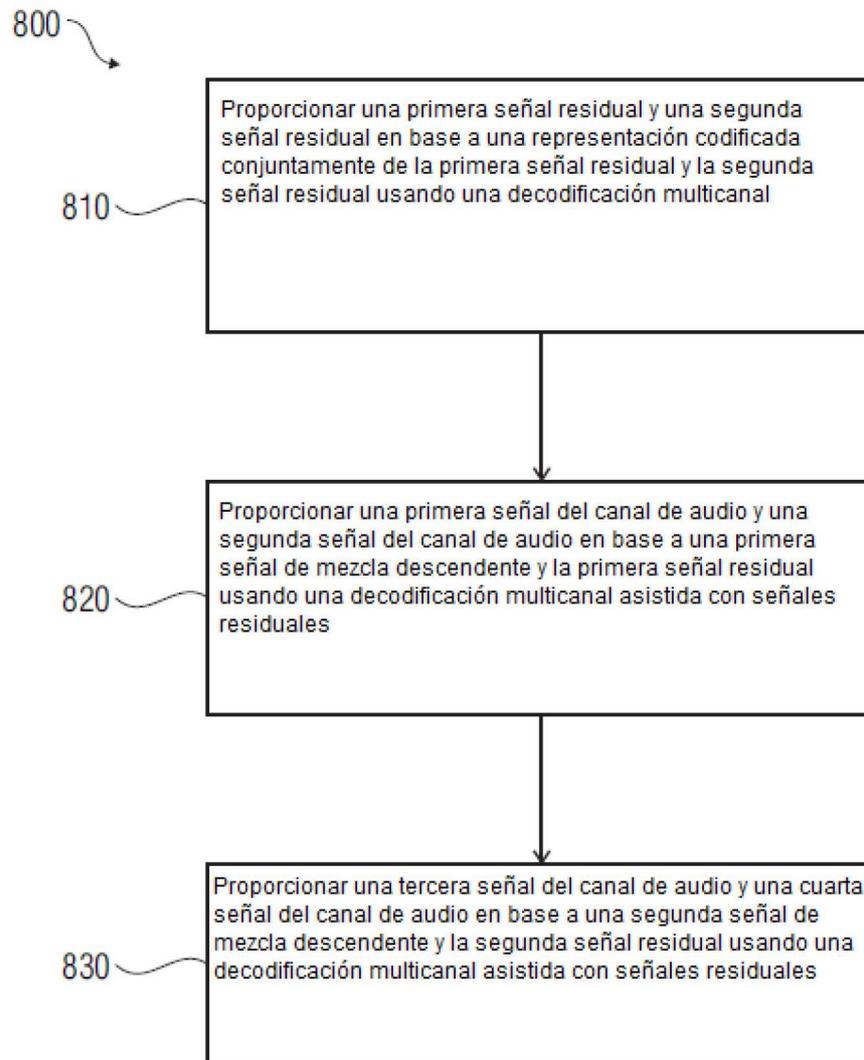


FIG. 8

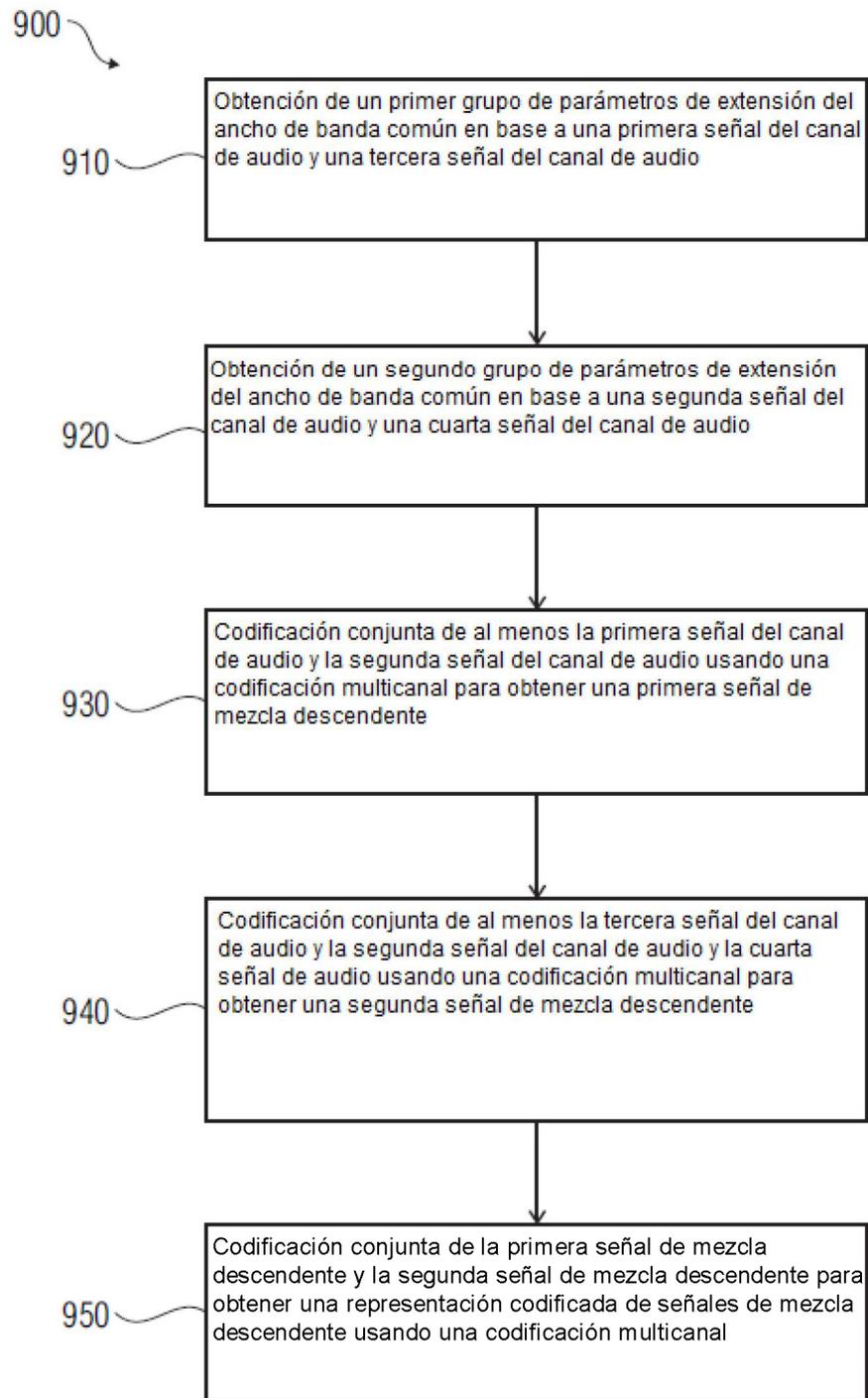


FIG. 9

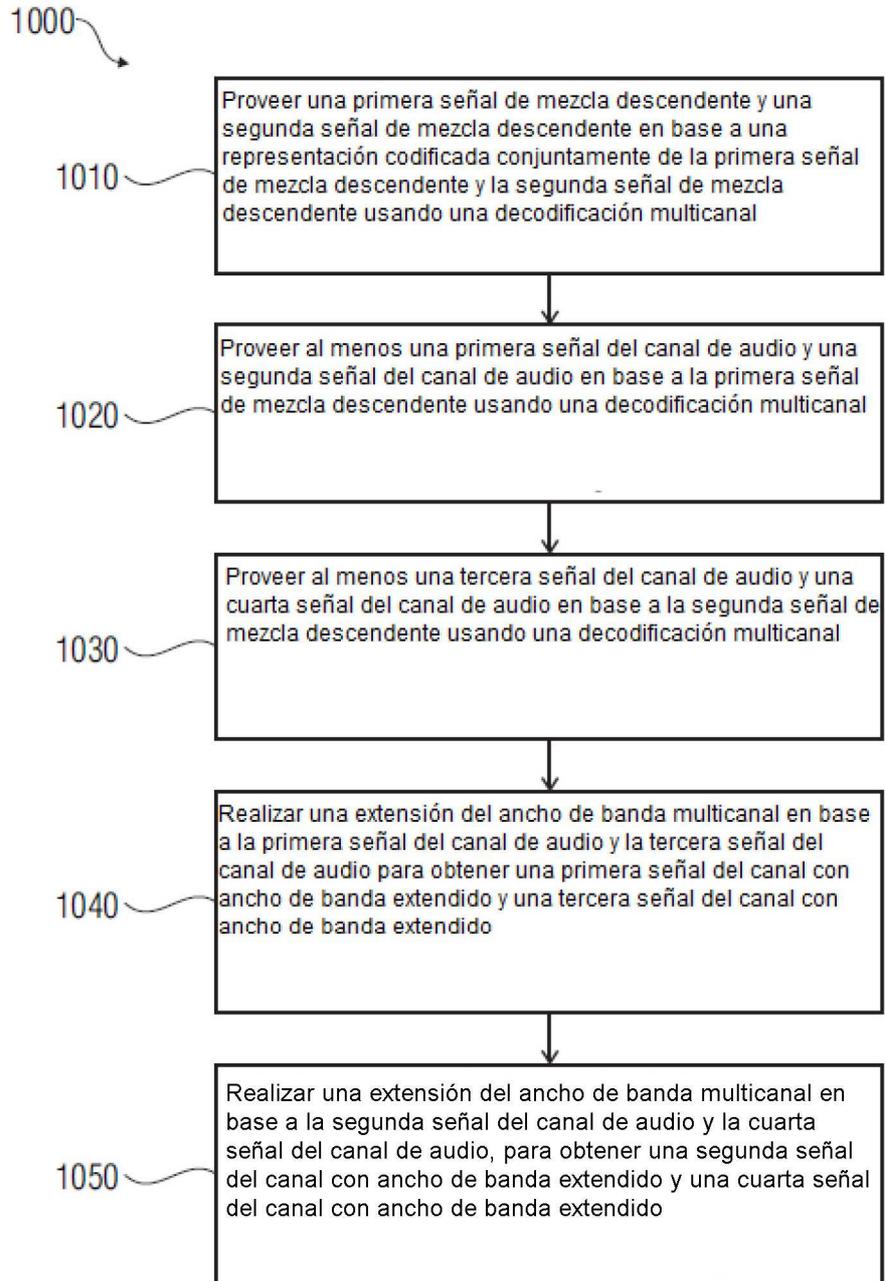


FIG. 10

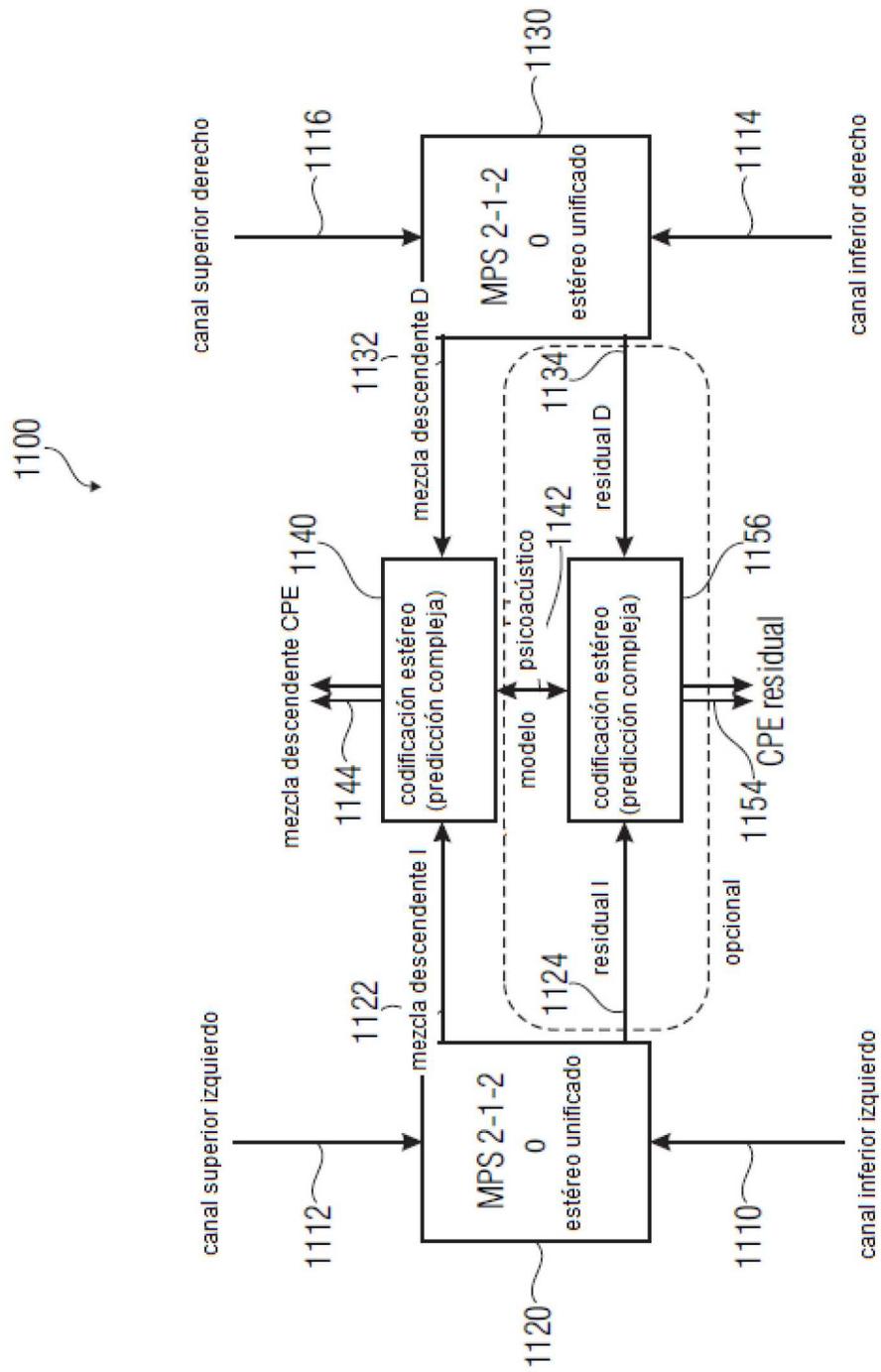


FIG. 11

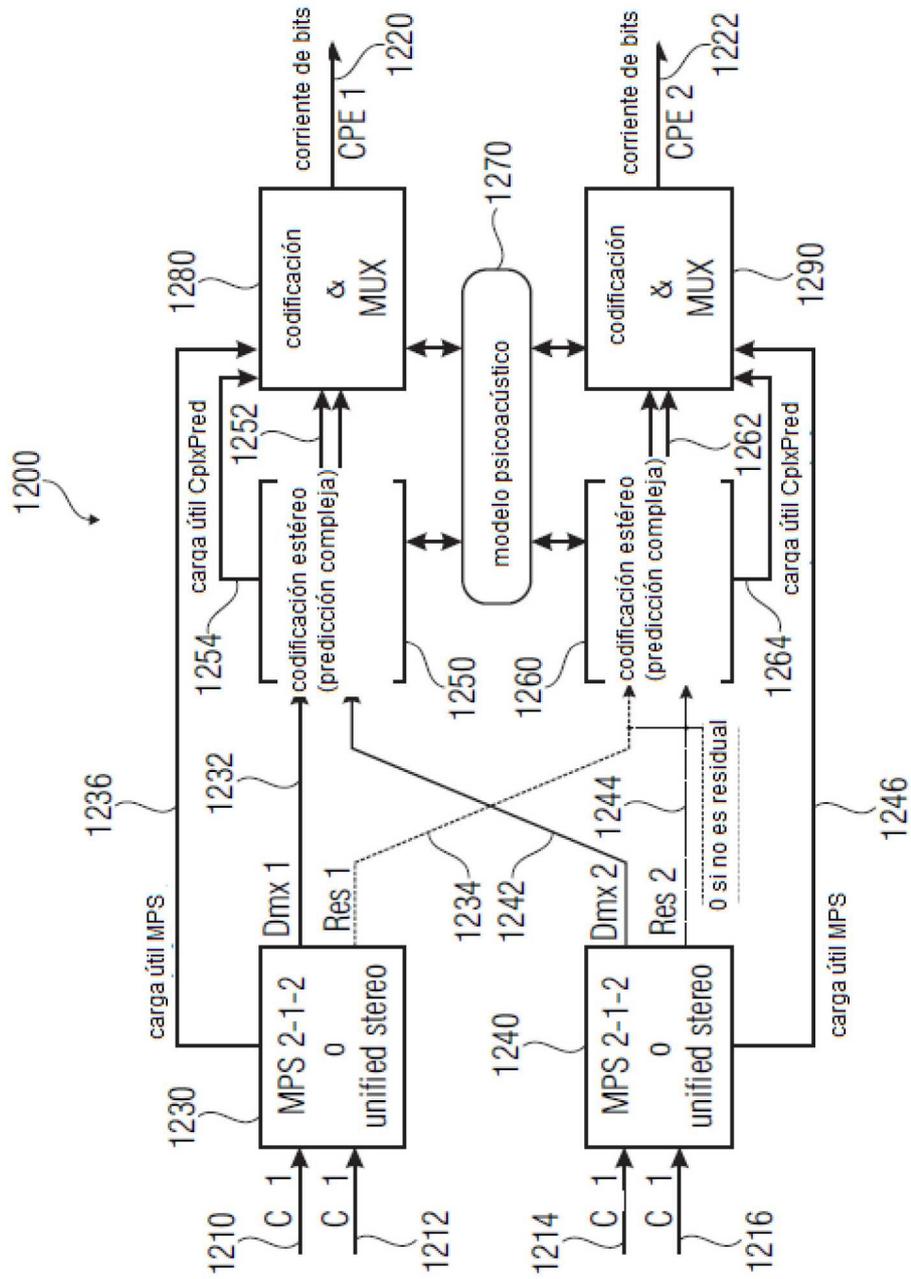


FIG. 12

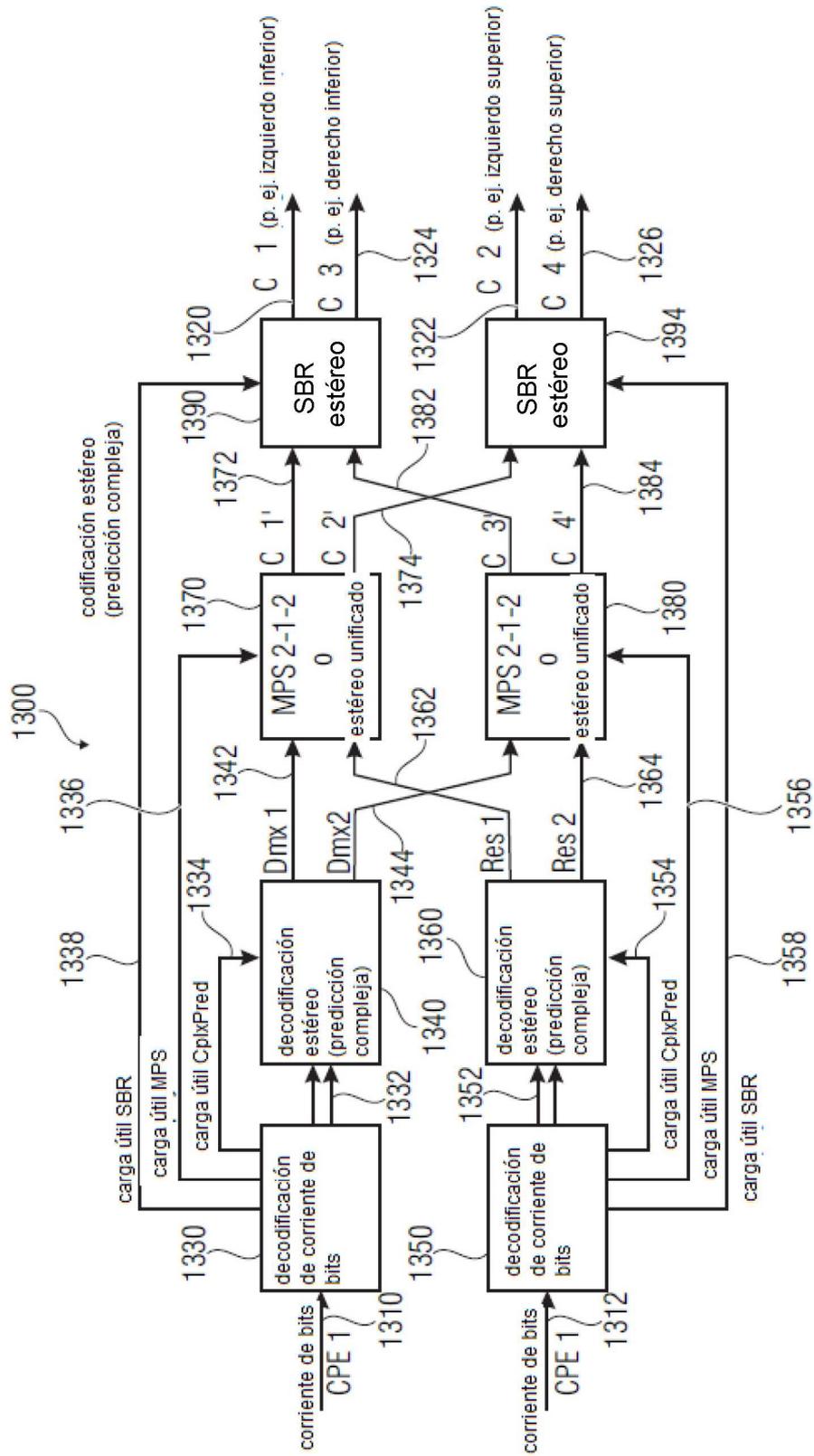


FIG. 13

```

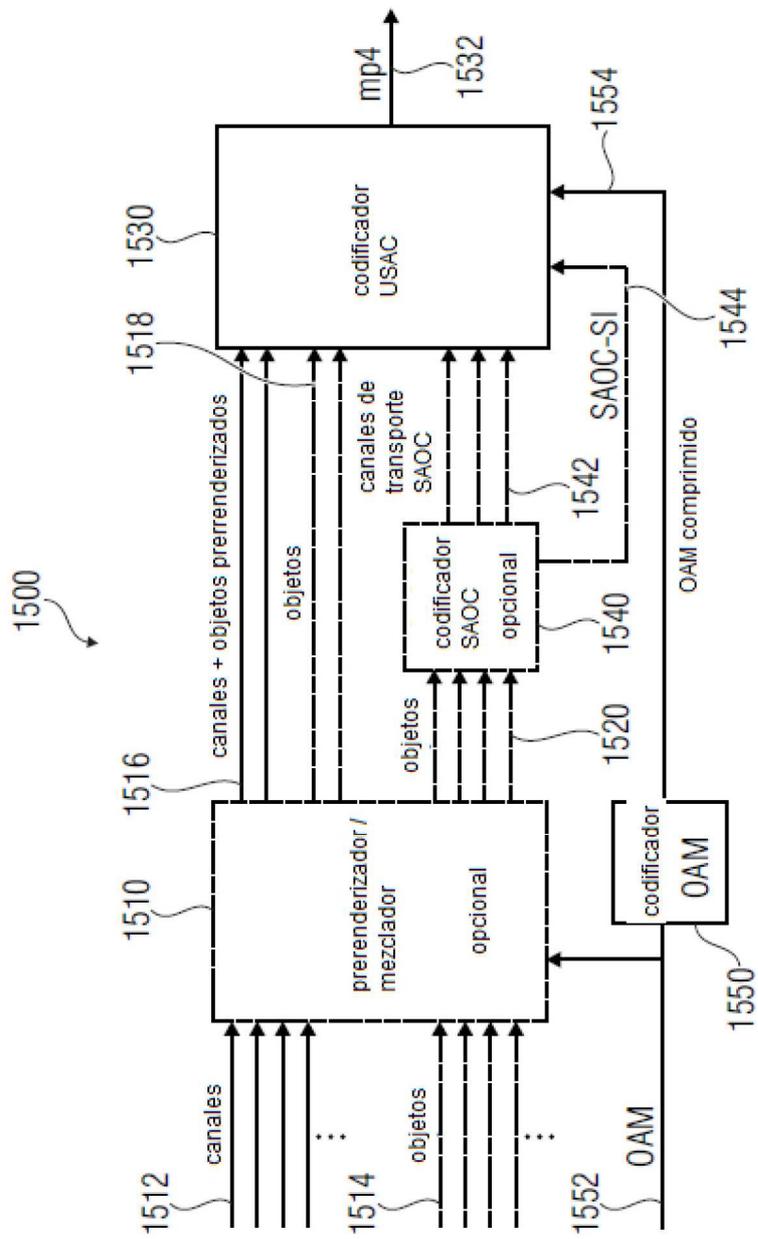
UsacChannelPairElementConfig (sbrRatiolIndex)
{
    UsacCoreConfig ();
    if (sbrRatiolIndex > 0) {
        SbrConfig ();
        stereoConfigIndex;                2      uimsbf
    } else {
        stereoConfigIndex = 0;
    }
    if (stereoConfigIndex > 0) {
        Mps212Config(stereoConfigIndex);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

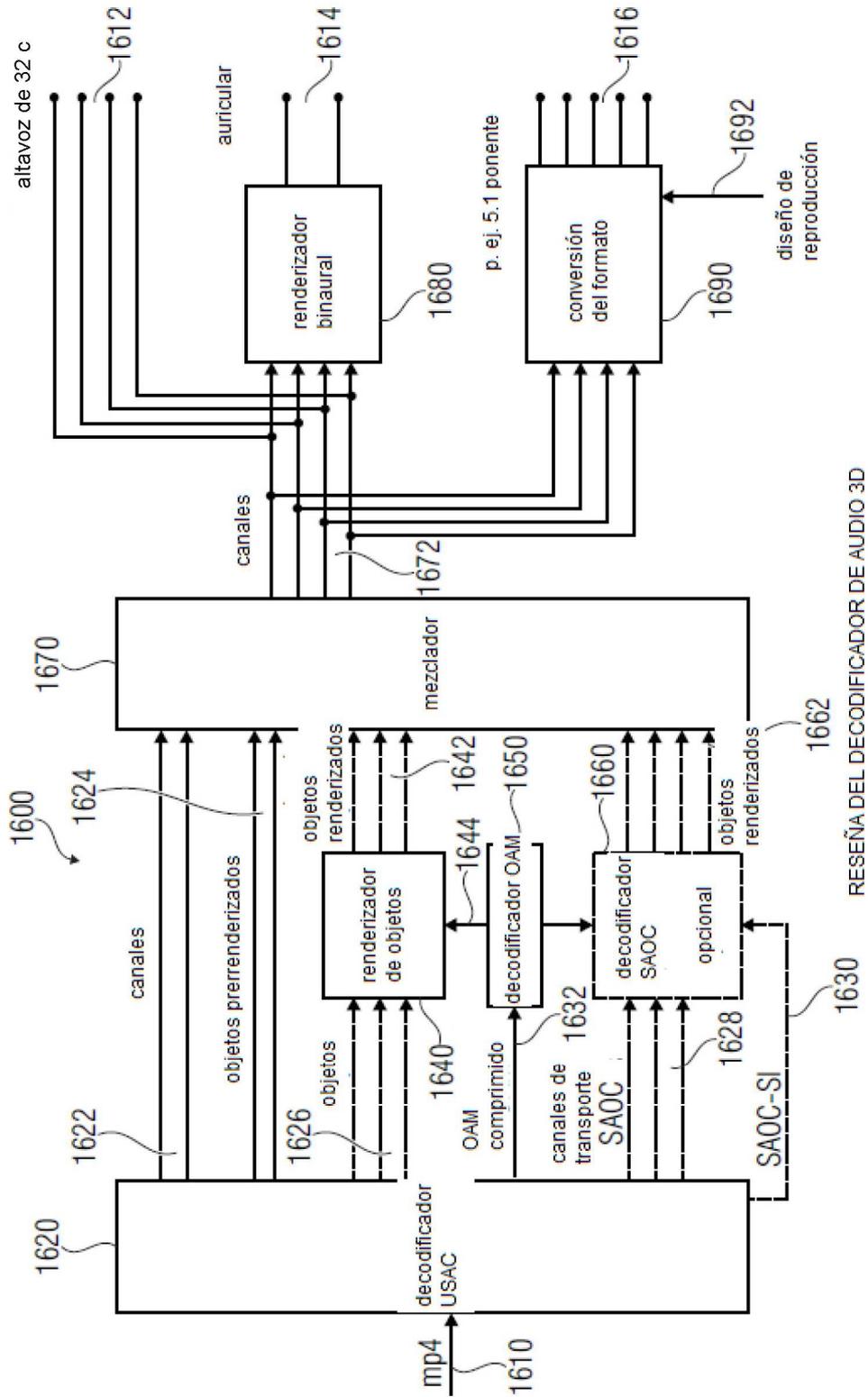
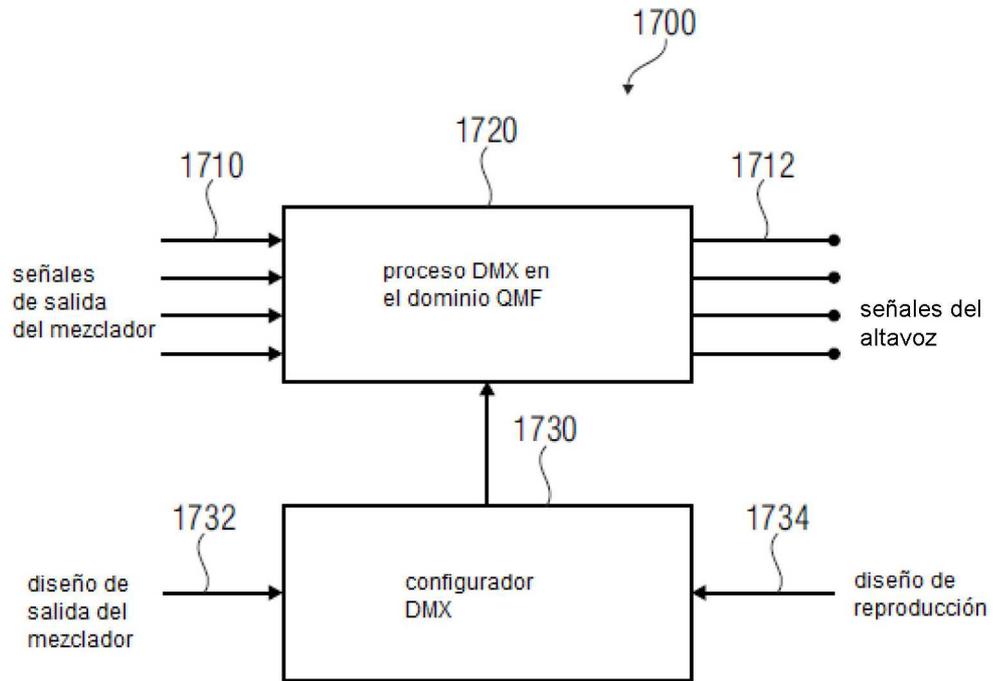


FIG. 16

RESEÑA DEL DECODIFICADOR DE AUDIO 3D



ESTRUCTURA DEL CONVERTIDOR DE FORMATO

FIG. 17

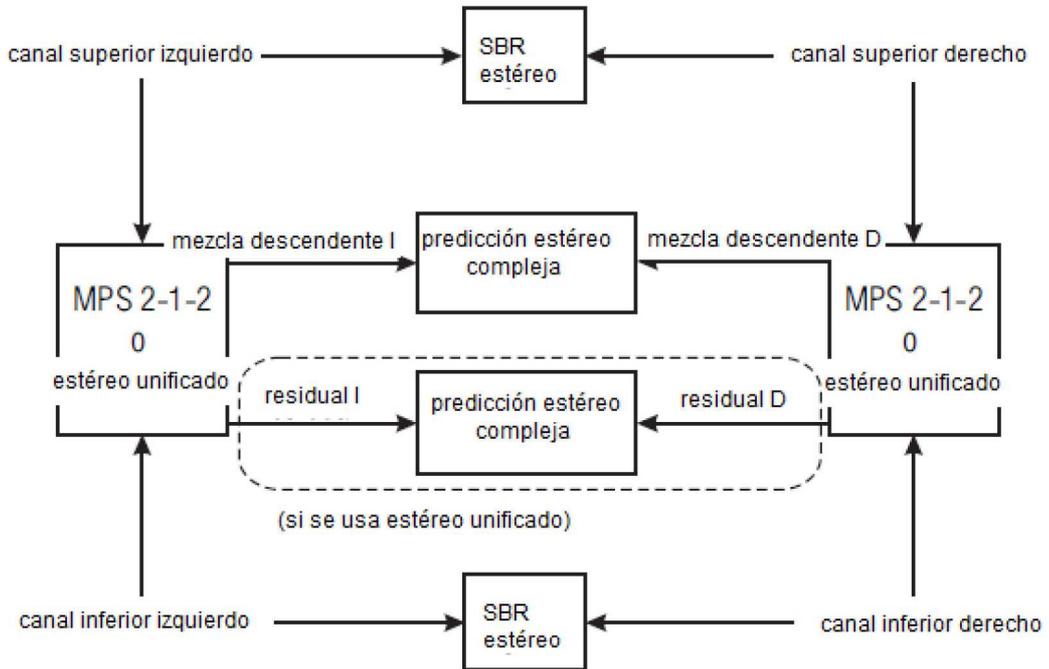


FIG. 18

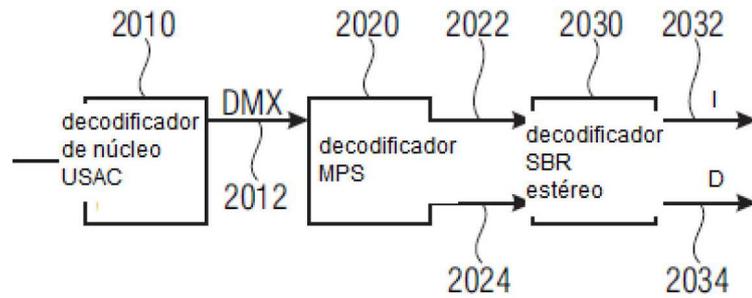
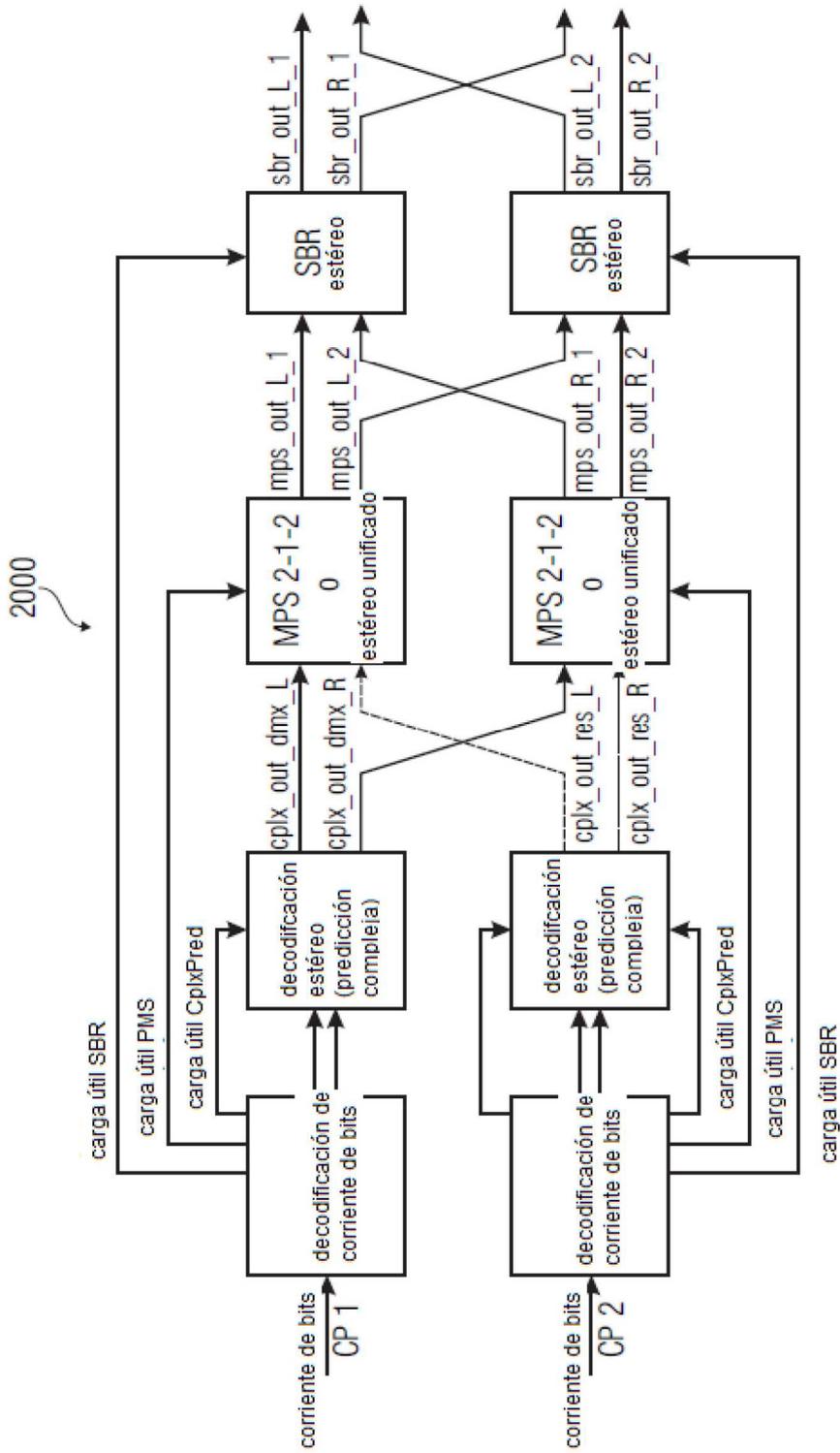
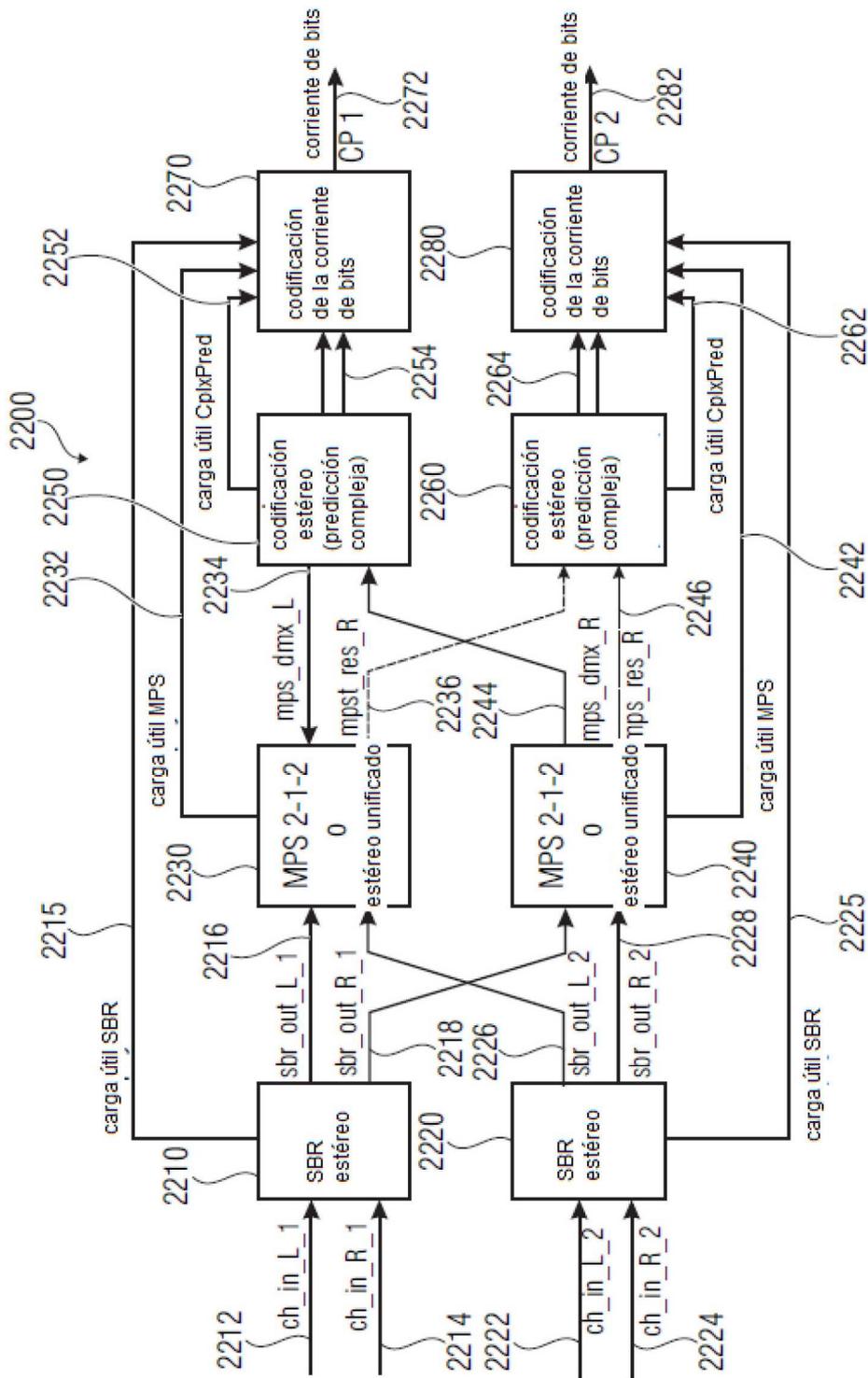


FIG. 19



ESQUEMAS DEL DECODIFICADOR QCE

FIG. 20



ESQUEMAS DEL CODIFICADOR DEL CANAL QUAD

FIG. 21