

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 995 102**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 19/18 (2013.01)

G10L 19/20 (2013.01)

G10L 19/22 (2013.01)

H04S 3/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **16.07.2014** **E 22159568 (9)**

97 Fecha y número de publicación de la concesión europea: **16.10.2024** **EP 4033485**

54 Título: **Concepto para descodificación de audio para canales de audio y objetos de audio**

30 Prioridad:

22.07.2013 EP 13177378

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

06.02.2025

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.00%)**

Hansastr. 27c

80686 München, DE

72 Inventor/es:

**ADAMI, ALEXANDER;
BORSS, CHRISTIAN;
DICK, SASCHA;
ERTEL, CHRISTIAN;
NEUKAM, SIMONE;
HERRE, JÜRGEN;
HILPERT, JOHANNES;
HÖLZER, ANDREAS;
KRATSCHMER, MICHAEL;
KÜCH, FABIAN;
KUNTZ, ACHIM;
MURTAZA, ADRIAN;
PLOGSTIES, JAN;
SILZLE, ANDREAS y
STENZEL, HANNE**

74 Agente/Representante:

PONTI & PARTNERS, S.L.P.

ES 2 995 102 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

Concepto para descodificación de audio para canales de audio y objetos de audio

5 Especificación

[0001] La presente invención se refiere a codificación/ descodificación de audio y, en particular, a codificación de audio espacial y codificación de un objeto de audio espacial.

- 10 **[0002]** Las herramientas de codificación de audio espacial son bien conocidas en la técnica y están, por ejemplo, estandarizadas en el estándar envolvente MPEG. La codificación de audio espacial comienza con canales de entrada originales, tales como cinco o siete canales que se identifican por su colocación en una configuración de reproducción, es decir, un canal izquierdo, un canal central, un canal derecho, un canal envolvente izquierdo, un canal envolvente derecho y un canal de realce de baja frecuencia. Un codificador de audio espacial normalmente deriva uno o más canales de *downmix* de los canales originales y, de forma adicional, deriva datos paramétricos relacionados con señales espaciales, tales como diferencias de nivel entre canales en los valores de coherencia de canal, diferencias de fase entre canales, diferencias de tiempo entre canales, etc. Los uno o más canales de *downmix* se transmiten junto con la información lateral paramétrica que indica las señales espaciales a un descodificador de audio espacial que descodifica el canal de *downmix* y los datos paramétricos asociados con el fin de obtener finalmente canales de salida que son una versión aproximada de los canales de entrada originales. La colocación de los canales en la configuración de salida es normalmente fija y es, por ejemplo, un formato 5.1, un formato 7.1, etc.

- [0003]** De forma adicional, las herramientas de codificación de un objeto de audio espacial son bien conocidas en la técnica y están estandarizadas en el estándar SAOC de MPEG (SAOC: del inglés «*Spatial Audio Object Coding*», codificación de un objeto de audio espacial, MPEG: del inglés «*Moving Picture Experts Group*», grupo de expertos en imágenes en movimiento). A diferencia de una codificación de audio espacial que se inicia en canales originales, la codificación de un objeto de audio espacial comienza con objetos de audio que no se dedican automáticamente a una determinada configuración de la reproducción de renderizado. En su lugar, la colocación de los objetos de audio en la escena de reproducción es flexible y puede ser determinada por el usuario al introducir determinada información de renderizado en un descodificador de codificación de objetos de audio espacial. De forma alternativa o adicional, la información de renderizado, es decir, la información en cuya posición en la configuración de reproducción deberá colocarse un determinado objeto de audio normalmente con el paso del tiempo puede transmitirse como información lateral adicional o metadatos. Con el fin de obtener una determinada compresión de datos, una cantidad de objetos de audio se codifican por medio de un codificador de SAOC que calcula, a partir de los objetos de entrada, uno o más canales de transporte mediante *downmixing* de los objetos según determinada información del proceso de *downmix*. Además, el codificador de SAOC calcula información lateral paramétrica que representa señales entre objetos, tales como diferencias en el nivel de objetos (OLD, por sus siglas en inglés), valores de coherencia de objetos, etc. Como ocurre en SAC (del inglés «*Spatial Audio Coding*», codificación de audio espacial), los datos paramétricos entre objetos se calculan para bloques individuales de tiempo/frecuencia, es decir, para un determinado cuadro de la señal de audio que comprende, por ejemplo, 1024 o 2048 muestras, se tienen en cuenta 24, 32, o 64, etc., bandas de frecuencia de modo que, al final, existen datos paramétricos para cada cuadro y cada banda de frecuencia. Como ejemplo, cuando una pieza de audio tiene 20 cuadros y cuando cada cuadro se subdivide en 32 bandas de frecuencia, la cantidad de bloques de tiempo/frecuencia es de 640.

- 45 **[0004]** Hasta ahora no existe una tecnología flexible que combine codificación de canales por un lado y codificación de objetos por otro, de modo que se obtengan calidades de audio aceptables a bajas tasas de bits.

- [0005]** El documento WO 201212544 A1 describe una solución integral para crear, codificar, transmitir, descodificar y reproducir bandas sonoras de audio espacial. El formato de codificación de bandas sonoras proporcionado es compatible con los formatos de codificación de sonido envolvente heredados, de modo que las bandas sonoras codificadas en el nuevo formato pueden descodificarse y reproducirse en equipos de reproducción heredados sin pérdida de calidad en comparación con los formatos heredados. Los objetos de audio se incluyen en un *downmix* base en el lado del codificador, y el *downmix* así obtenido y los objetos de audio codificados explícitamente se transmiten a un lado del descodificador. En el lado del descodificador, los objetos se eliminan del *downmix* transmitido y se renderizan y combinan por separado con el *downmix* residual correspondiente al *downmix* base.

- [0006]** El documento US 2010324915 A1 describe un aparato de codificación para un códec de audio multicanal de alta calidad (HQMAC, por sus siglas en inglés) y un aparato de descodificación para dicho HQMAC. Los aparatos de codificación/descodificación del HQMAC pueden realizar una codificación basada en canales del códec de audio multicanal de alta calidad (HQM-CB, por sus siglas en inglés) o una descodificación HQMAC-CB según las características de las señales de audio de entrada para proporcionar compatibilidad con un canal inferior.

- [0007]** Es un objeto de la presente invención proporcionar un concepto mejorado para la descodificación de audio.

65

[0008] Este objeto se logra mediante un descodificador de audio según la reivindicación 1, un procedimiento de descodificación de audio según la reivindicación 16 o un programa informático según la reivindicación 17.

[0009] La presente invención se basa en el hallazgo de que un sistema óptimo que sea flexible por un lado y proporcione una buena eficiencia de compresión con una buena calidad de audio por otro, se logra mediante la combinación de la codificación de audio espacial, es decir, la codificación de audio basada en canales, con la codificación de un objeto de audio espacial, es decir, la codificación basada en objetos. En particular, proporcionar un mezclador para mezclar los objetos y los canales que ya se encuentran en el lado del codificador proporciona una buena flexibilidad, particularmente para aplicaciones de baja tasa de bits, dado que, en ese caso, puede resultar innecesaria cualquier transmisión de objetos, o la cantidad de objetos que deben transmitirse puede reducirse. Por otro lado, se requiere flexibilidad para que el codificador de audio pueda controlarse en dos modos diferentes, es decir, en un modo donde los objetos se mezclan con los canales antes de someterse a codificación principal, mientras que, en el otro modo, los datos de objetos, por un lado, y los datos de canal, por otro lado, se someten directamente a codificación principal sin ninguna mezcla entre ambos.

[0010] Esto garantiza que el usuario pueda separar los objetos y canales procesados en el lado del codificador de modo que se disponga de una flexibilidad total en el lado del descodificador, aunque a cambio de una mayor tasa de bits. Por otro lado, cuando los requisitos de tasa de bits son más rigurosos, la presente invención ya permite realizar una mezcla/prerrenderizado en el lado del codificador, es decir, que algunos o todos los objetos de audio ya se encuentran mezclados con los canales, de modo que el codificador principal únicamente codifica datos de canal y no se requieren bits para transmitir datos de objeto de audio, ni en forma de *downmix* ni en forma de datos paramétricos entre objetos requieren.

[0011] En el lado del descodificador, el usuario una vez más tiene una alta flexibilidad debido a que el mismo descodificador de audio permite el funcionamiento en dos modos diferentes, es decir, en un primer modo donde se realiza la codificación de objetos y canales individuales o separados y el descodificador tiene total flexibilidad para renderizar los objetos y mezclarlos con los datos de canal. Por otro lado, cuando ya se ha realizado una mezcla/renderizado previo en el lado del codificador, el descodificador se configura para realizar un posprocesamiento sin procesamiento intermedio de objetos. Por otro lado, el posprocesamiento también puede aplicarse a los datos en el otro modo, es decir, cuando el renderizado/la mezcla de objetos se realiza en el lado del descodificador. De este modo, la presente invención permite un marco de tareas de procesamiento que permite una gran reutilización de recursos, no solamente en el lado del codificador, sino también en el lado del descodificador. El posprocesamiento puede referirse al *downmixing* y binauralización o cualquier otro procesamiento para obtener un escenario final de canales, tal como una disposición de reproducción pretendida.

[0012] Además, en caso de unos muy bajos requisitos de tasa de bits, la presente invención proporciona al usuario la suficiente flexibilidad para reaccionar a dichos bajos requisitos de tasa de bits, es decir, mediante renderizado previo en el lado del codificador, de modo que, aunque a cambio de cierta flexibilidad, se obtiene, no obstante, una muy buena calidad de audio en el lado del descodificador debido a que los bits que se han ahorrado al dejar de proporcionar datos de objetos del codificador al descodificador pueden usarse para codificar mejor los datos de canal, tal como mediante una cuantificación más precisa de los datos de canal o por otros medios, para mejorar la calidad o para reducir la pérdida por codificación cuando se encuentran disponibles bits suficientes.

[0013] En una realización preferida de la presente invención, el codificador comprende de forma adicional un codificador de SAOC y, además, permite no solo codificar la entrada de objetos en el codificador, sino también codificar datos de canal mediante SAOC con el fin de obtener una buena calidad de audio, incluso con requisitos de tasa de bits más bajos. Otras realizaciones de la presente invención permiten una funcionalidad de posprocesamiento que comprende un renderizador binaural y/o un conversor de formato. Además, se prefiere que todo el procesamiento en el lado del descodificador ya se realice para una determinada cantidad elevada de altavoces, tal como una configuración de altavoces de 22 o 32 canales. Sin embargo, a continuación, el conversor de formato, por ejemplo, determina que únicamente se requiere una salida de 5.1, es decir, una salida para una disposición de reproducción que tenga una cantidad menor que la cantidad máxima de canales. En ese caso, se prefiere que el conversor de formato controle o bien el descodificador de USAC o el descodificador de SAOC, o ambos dispositivos, para restringir la operación de descodificación principal y la operación de descodificación de SAOC de modo que cualquier canal que, al final, no obstante, se someta a *downmixing* en una conversión de formato no se genere en la descodificación. Normalmente, la generación de canales sometidos a *upmixing* requiere procesamiento de decorrelación, y cada procesamiento de decorrelación introduce cierto nivel de artefactos. Por lo tanto, al controlar el descodificador principal y/o el descodificador de SAOC mediante el formato de salida finalmente requerido, se ahorra una gran cantidad de procesamiento de decorrelación adicional en comparación con una situación donde no existe esta interacción. Esto no solo se traduce en una mejora de la calidad de audio, sino que también se traduce en una menor complejidad del descodificador y, al final, en un menor consumo de energía, lo cual resulta particularmente útil para dispositivos móviles que alberguen el codificador o el descodificador inventivo. No obstante, los codificadores/descodificadores inventivos no solo pueden introducirse en dispositivos móviles, tales como teléfonos móviles, teléfonos inteligentes, ordenadores portátiles (*notebooks*) o dispositivos de navegación, sino que también pueden usarse en ordenadores de escritorio sencillos o cualquier otro aparato no móvil.

[0014] La implementación anterior, es decir, no generar algunos canales, puede no resultar óptima, dado que puede perderse cierta información (tal como la diferencia de nivel entre los canales que se sometan a *downmixing*). Dicha información de diferencia de nivel puede no ser crítica, pero puede producir una señal de salida de *downmix* diferente, si el *downmixing* aplica diferentes ganancias de *downmix* a los canales sometidos a *downmixing*. Una solución mejorada únicamente desactiva la decorrelación en la *upmix*, pero sigue generando todos los canales de *upmix* con las diferencias de nivel correctas (como señala la SAC paramétrica). La segunda solución produce una mejor calidad de audio, pero la primera solución se traduce en una mayor reducción de la complejidad.

10 **[0015]** A continuación, se describen las realizaciones preferidas haciendo referencia a los dibujos adjuntos, donde:

La FIG. 1 ilustra un primer ejemplo de un codificador;

La FIG. 2 ilustra una primera realización de un descodificador;

15 La FIG. 3 ilustra un segundo ejemplo de un codificador;

La FIG. 4 ilustra una segunda realización de un descodificador;

La FIG. 5 ilustra un tercer ejemplo de un codificador;

La FIG. 6 ilustra una tercera realización de un descodificador;

20 La FIG. 7 ilustra un mapa que indica los distintos modos de funcionamiento de los codificadores/descodificadores según las realizaciones de la presente invención;

La FIG. 8 ilustra una implementación específica del conversor de formato;

La FIG. 9 ilustra una implementación específica del conversor binaural;

La FIG. 10 ilustra una implementación específica del descodificador principal; y

25 La FIG. 11 ilustra una implementación específica de un codificador para procesar un elemento de cuádruple canal (QCE, por sus siglas en inglés) y el correspondiente descodificador QCE.

[0016] La Figura 1 ilustra un codificador según un ejemplo de la presente invención. El codificador se configura para codificar datos de entrada de audio 101 para obtener datos de salida de audio 501. El codificador comprende una interfaz de salida para recibir una pluralidad de canales de audio indicados como CH y una pluralidad de objetos de audio indicados como OBJ. Además, como se ilustra en la Figura 1, la interfaz de entrada 100, de forma adicional, recibe metadatos relacionados con uno o más de la pluralidad de objetos de audio OBJ. Además, el codificador comprende un mezclador 200 para mezclar la pluralidad de objetos y la pluralidad de canales con el fin de obtener una pluralidad de canales previamente mezclados, donde cada canal previamente mezclado comprende datos de audio de un canal y datos de audio de al menos un objeto.

35

[0017] Además, el codificador comprende un codificador principal 300 para realizar la codificación principal de datos de entrada del codificador principal y un compresor de metadatos 400 para comprimir los metadatos relacionados con uno o más de la pluralidad de objetos de audio. Además, el codificador puede comprender un controlador de modos 600 para controlar el mezclador, el codificador principal y/o una interfaz de salida 500 en uno de varios modos de funcionamiento, donde en el primer modo, el codificador principal se configura para codificar la pluralidad de canales de audio y la pluralidad de objetos de audio recibidos por la interfaz de entrada 100 sin ninguna interacción del mezclador, es decir, sin que el mezclador 200 realice ninguna mezcla. Sin embargo, en un segundo modo donde el mezclador 200 está activo, el codificador principal codifica la pluralidad de canales mezclados, es decir, la salida generada por el bloque 200. En este último caso, se prefiere no codificar datos de objetos adicionales. En su lugar, el mezclador 200 ya utiliza los metadatos que indican las posiciones de los objetos de audio para renderizar los objetos en los canales, tal y como indican los metadatos. Dicho de otro modo, el mezclador 200 utiliza los metadatos relacionados con la pluralidad de objetos de audio para prerrenderizar los objetos de audio y, a continuación, los objetos de audio previamente renderizados se mezclan con los canales para obtener canales mezclados en la salida del mezclador. En este ejemplo, es posible que no se transmitan necesariamente objetos, y esto también se aplica a los metadatos comprimidos como salida por el bloque 400. No obstante, si no se mezclan todos los objetos introducidos en la interfaz 100, sino únicamente una determinada cantidad de objetos, solo se transmiten de todas formas al codificador principal 300 o al compresor de metadatos 400, respectivamente, los objetos no mezclados anteriormente y los metadatos asociados.

55 **[0018]** La Figura 3 ilustra un ejemplo adicional de un codificador que, de forma adicional, comprende un codificador de SAOC 800. El codificador de SAOC 800 se configura para generar uno o más canales de transporte y datos paramétricos a partir de los datos de entrada del codificador de objetos de audio espacial. Como se ilustra en la Figura 3, los datos de entrada del codificador de objeto de audio espaciales son objetos que no han sido procesados por el prerrenderizador/ mezclador. Alternativamente, siempre que se haya omitido el prerrenderizador/ mezclador, como en el modo uno, donde está activa una codificación de objetos/canales individuales, todas las entradas de objetos en la interfaz de entrada 100 se codifican por medio del codificador de SAOC 800.

60 **[0019]** Además, como se ilustra en la Figura 3, el codificador principal 300 se implementa preferentemente como un codificador USAC, es decir, como un codificador tal y como se define y estandariza en el estándar MPEG-USAC (del inglés «*Unified Speech and Audio Coding*», codificación unificada de voz y audio). La salida del codificador

completo que se ilustra en la Figura 3 es un flujo de datos MPEG 4 que tiene estructuras similares a un contenedor para distintos tipos de datos. Además, los metadatos se indican como datos «OAM» y el compresor de metadatos 400 de la Figura 1 corresponde al codificador OAM 400 para obtener datos OAM comprimidos que se introducen en el codificador de USAC 300 que, como puede observarse en la Figura 3, comprende de forma adicional la interfaz de salida para obtener el flujo de datos de salida MP4, que no solo contiene los datos de objeto/canal codificados, sino que también contiene los datos OAM comprimidos.

[0020] La Figura 5 ilustra un ejemplo adicional del codificador, donde a diferencia de la Figura 3, el codificador de SAOC puede configurarse de forma indistinta para codificar, con el algoritmo de codificación de SAOC, los canales provistos en el prerrenderizador/mezclador 200 que no están activos en este modo o, alternativamente, para codificar mediante SAOC los canales previamente renderizados más los objetos. De este modo, en la Figura 5, el codificador de SAOC 800 puede funcionar con tres clases diferentes de datos de entrada, es decir, canales sin ningún objeto previamente renderizado, canales y objetos previamente renderizados, o únicamente objetos. Además, se prefiere proporcionar un descodificador OAM adicional 420 en la Figura 5, de modo que el codificador de SAOC 800 use, para su procesamiento, los mismos datos que en el lado del descodificador, es decir, datos obtenidos por una compresión con pérdida en lugar de los datos OAM originales.

[0021] El codificador de la Figura 5 puede funcionar en varios modos distintos.

[0022] Además del primero y segundo modo, como se mencionó en relación con la Figura 1, el codificador de la Figura 5 puede funcionar, de forma adicional, en un tercer modo donde el codificador principal genera los uno o más canales de transporte a partir de los objetos individuales cuando el prerrenderizador/mezclador 200 no estaba activo. De forma alternativa o adicional, en este tercer modo el codificador de SAOC 800 puede generar uno o más canales de transporte alternativos o adicionales a los canales originales, es decir, de nuevo cuando el prerrenderizador/mezclador 200 correspondiente al mezclador 200 de la Figura 1 no estaba activo.

[0023] Por último, el codificador de SAOC 800 puede codificar, cuando el codificador se configura en el cuarto modo, los canales más objetos previamente renderizados según los genera el prerrenderizador /mezclador. De este modo, en el cuarto modo, las aplicaciones con menor tasa de bits ofrecerán buena calidad dado que los canales y objetos se han transformado por completo en canales de transporte de SAOC individuales y la información lateral asociada, que se indica como «SAOC-SI» en las Figuras 3 y 5, y, de forma adicional, no tienen por qué transmitirse metadatos no comprimidos en este cuarto modo.

[0024] La Figura 2 ilustra un descodificador según una realización de la presente invención. El descodificador recibe, como entrada, los datos de audio codificados, es decir, los datos 501 de la Figura 1. El descodificador comprende un descompresor de metadatos 1400, un descodificador principal 1300, un procesador de objetos 1200, un controlador de modos 1600 y un posprocesador 1700.

[0025] Específicamente, el descodificador de audio configura para descodificar datos de audio codificados y la interfaz de entrada se configura para recibir los datos de audio codificados, comprendiendo los datos de audio codificados una pluralidad de canales codificados y la pluralidad de objetos codificados y metadatos comprimidos relacionados con la pluralidad de objetos en un determinado modo.

[0026] Además, el descodificador principal 1300 se configura para descodificar la pluralidad de canales codificados y la pluralidad de objetos codificados y, de forma adicional, el descompresor de metadatos se configura para descomprimir los metadatos comprimidos.

[0027] Además, el procesador de objetos 1200 se configura para procesar la pluralidad de objetos descodificados según los genera el descodificador principal 1300 utilizando los metadatos descomprimidos para obtener una cantidad predeterminada de canales de salida que comprende datos de objeto y los canales descodificados. Estos canales de salida, tal y como se indica en 1205, a continuación, se introducen en un posprocesador 1700. El posprocesador 1700 se configura para convertir la cantidad de canales de salida 1205 en un determinado formato de salida, que puede ser un formato de salida binaural o un formato de salida por altavoz, tal como un formato de salida 5.1, 7.1, etc.

[0028] El descodificador comprende un controlador de modos 1600 que se configura para analizar los datos codificados para detectar una indicación de modo. Por lo tanto, el controlador de modos 1600 está conectado a la interfaz de entrada 1100 en la Figura 2. El descodificador de audio de la Figura 2 y, controlado por el controlador de modos 1600, se configura para omitir el procesador de objetos y para introducir la pluralidad de canales descodificados al posprocesador 1700. Este es el funcionamiento en el modo 2, es decir, donde solo se reciben canales previamente renderizados, es decir, cuando se ha aplicado el modo 2 en el codificador de la Figura 1. Alternativamente, cuando se ha aplicado el modo 1 en el codificador, es decir, cuando el codificador ha realizado la codificación de canales/objetos individuales, no se omite el procesador de objetos 1200, sino que la pluralidad de canales descodificados y la pluralidad de objetos descodificados se introducen en el procesador de objetos 1200 junto con los metadatos descomprimidos generados por el descompresor de metadatos 1400.

[0029] Según la invención, la indicación de si debe aplicarse el modo 1 o el modo 2 se incluye los datos de audio codificados y luego el controlador de modos 1600 analiza los datos codificados para detectar una indicación de modo. Se usa el modo 1 cuando la indicación de modo indica que los datos de audio codificados comprenden canales codificados y objetos codificados y se aplica el modo 2 cuando la indicación de modo indica que los datos de audio codificados no contienen ningún objeto de audio, es decir, únicamente contienen canales previamente renderizados obtenidos mediante el modo 2 del codificador de la Figura 1.

[0030] La Figura 4 ilustra una realización preferida comparada con la del descodificador de la Figura 2, y la realización de la Figura 4 corresponde al codificador de la Figura 3. Además de la implementación del descodificador de la Figura 2, el descodificador de la Figura 4 comprende un descodificador de SAOC 1800. Además, el procesador de objetos 1200 de la Figura 2 se implementa como un renderizador de objetos por separado 1210 y el mezclador 1220 mientras que, dependiendo del modo, la funcionalidad del renderizador de objetos 1210 también puede implementarse mediante el descodificador de SAOC 1800.

[0031] Además, el posprocesador 1700 puede implementarse como un renderizador binaural 1710 o un conversor de formato 1720. Alternativamente, una salida directa de datos 1205 de la Figura 2 también puede implementarse como se ilustra por medio de 1730. Por lo tanto, se prefiere realizar el procesamiento en el descodificador sobre la cantidad más elevada de canales, tales como 22,2 o 32, con el fin de tener flexibilidad y posprocesar a continuación si se requiere un formato más pequeño. Sin embargo, cuando resulte obvio desde el principio que únicamente se requiere un formato pequeño, tal como un formato 5.1, se prefiere, como se indica en la Figura 2 o 6 mediante el «atajo» 1727, que pueda aplicarse un determinado control sobre el descodificador de SAOC y/o el descodificador de USAC aplicarse con el fin de evitar operaciones de *upmixing* innecesarias y las posteriores operaciones de *downmixing*.

[0032] En una realización preferida de la presente invención, el procesador de objetos 1200 comprende el descodificador de SAOC 1800 y el descodificador de SAOC se configura para descodificar uno o más canales de transporte emitidos por el descodificador principal y los datos paramétricos asociados y, utilizando los metadatos descomprimidos, para obtener la pluralidad de objetos de audio renderizados. Para este fin, la salida de OAM se conecta al módulo 1800.

[0033] Además, el procesador de objetos 1200 se configura para renderizar objetos descodificados emitidos por el descodificador principal que no se codifican en los canales de transporte de SAOC, sino que se codifican individualmente, normalmente en elementos de un solo canal, como indica el renderizador de objetos 1210. Además, el descodificador comprende una interfaz de salida que corresponde a la salida 1730 para emitir una salida del mezclador a los altavoces.

[0034] En una realización adicional, el procesador de objetos 1200 comprende un descodificador de codificación de objetos de audio espaciales 1800 para descodificar uno o más canales de transporte y la información lateral paramétrica asociada que representa objetos de audio codificados o canales de audio codificados, donde el descodificador de codificación de objetos de audio espaciales se configura para transcódecificar la información paramétrica asociada y los metadatos descomprimidos en información lateral paramétrica transcódecificada susceptible de usarse para renderizar directamente el formato de salida, como se define, por ejemplo, en una versión anterior de SAOC. El posprocesador 1700 se configura para calcular canales de audio del formato de salida utilizando los canales de transporte descodificados y la información lateral paramétrica transcódecificada. El procesamiento realizado por el posprocesador puede ser similar al procesamiento MPEG Surround o puede tratarse de cualquier otro procesamiento, tal como procesamiento BCC y similares.

[0035] En una realización adicional, el procesador de objetos 1200 comprende un descodificador de codificación de objetos de audio espaciales 1800 configurado para someter directamente a *upmixing* y renderizar señales de canal para el formato de salida utilizando los canales de transporte descodificados (por el descodificador principal) y la información lateral paramétrica.

De forma adicional e importante, el procesador de objetos 1200 de la Figura 2 comprende, de forma adicional, el mezclador 1220 que recibe, como entrada, datos generados por el descodificador de USAC 1300 directamente cuando existen objetos previamente renderizados mezclados con canales, es decir, cuando el mezclador 200 de la Figura 1 estaba activo. De forma adicional, el mezclador 1220 recibe datos del renderizador de objetos que realiza el renderizado de objetos sin descodificación de SAOC. Además, el mezclador recibe datos de salida del descodificador de SAOC, es decir, objetos renderizados mediante SAOC.

[0036] El mezclador 1220 se conecta a la interfaz de salida 1730, el renderizador binaural 1710 y el conversor de formato 1720. El renderizador binaural 1710 se configura para renderizar los canales de salida en dos canales binaurales utilizando funciones de transferencia relacionadas con la cabeza o respuestas binaurales a los impulsos en el espacio (BRIR, por sus siglas en inglés). El conversor de formato 1720 se configura para convertir los canales de salida en un formato de salida que tenga una cantidad menor de canales que los canales de salida 1205 del mezclador, y el conversor de formato 1720 requiere información sobre la disposición de reproducción, tal como altavoces 5.1 y

similares.

[0037] El descodificador de la Figura 6 se diferencia del descodificador de la Figura 4 en el hecho de que el descodificador de SAOC no solo puede generar objetos renderizados, sino también canales renderizados, como sucede cuando se ha utilizado el codificador de la Figura 5 y la conexión 900 entre los canales/objetos previamente renderizados y la interfaz de entrada del codificador de SAOC 800 está activa.

[0038] Además, se configura una etapa de paneo de amplitud basado en vectores (VBAP, por sus siglas en inglés) 1810 que recibe, del descodificador de SAOC, información sobre la disposición de reproducción y que emite una matriz de renderizado al descodificador de SAOC, de modo que el descodificador de SAOC puede, al final, proporcionar canales renderizados sin un funcionamiento adicional del mezclador en el formato de canal alto de 1205, es decir, 32 altavoces.

[0039] El bloque de VBAP recibe preferentemente los datos OAM descodificados para derivar las matrices de renderizado. Más en general, requiere preferentemente información geométrica, no solo de la disposición de reproducción, sino también de las posiciones donde deben renderizarse las señales de entrada en la disposición de reproducción. Estos datos de entrada geométricos pueden ser datos OAM para los objetos, o información de posición de canal para canales que se hayan transmitido mediante SAOC.

[0040] Sin embargo, si solo se requiere una interfaz de salida específica, el estado de VBAP 1810 ya puede proporcionar la matriz de renderizado requerida para la salida, por ejemplo, 5.1. A continuación, el descodificador de SAOC 1800 realiza un renderizado directo de los canales de transporte de SAOC, los datos paramétricos asociados y los metadatos descomprimidos, un renderizado directo en el formato de salida requerido sin ninguna interacción del mezclador 1220. Sin embargo, cuando se aplica una determinada mezcla entre modos, es decir, cuando varios canales se codifican con SAOC, pero no todos los canales se codifican con SAOC, o cuando varios objetos se codifican con SAOC, pero no todos los objetos se codifican con SAOC, o cuando solo una determinada cantidad de objetos previamente renderizados con canales se descodifican con SAOC y los canales restantes no se procesan con SAOC, el mezclador unificará los datos procedentes de las distintas porciones de entrada, es decir, directamente del descodificador principal 1300, del renderizador de objetos 1210 y del descodificador de SAOC 1800.

[0041] A continuación, se describe la Figura 7 para indicar determinados modos de codificador/descodificador que puede aplicar el concepto de codificador/descodificador de audio sumamente flexible y de alta calidad de la invención.

[0042] Según el primer modo de codificación, se omite el mezclador 200 del codificador de la Figura 1 y, por lo tanto, no se omite el procesador de objetos del descodificador de la Figura 2.

[0043] En el segundo modo, el mezclador 200 de la Figura 1 está activo y se omite el procesador de objetos de la Figura 2.

[0044] A continuación, en el tercer modo de codificación, el codificador de SAOC de la Figura 3 está activo, pero SAOC codifica únicamente los objetos, en lugar de los canales o los canales emitidos por el mezclador. Por lo tanto, el modo 3 requiere que, en el lado del descodificador que se ilustra en la Figura 4, el descodificador de SAOC solo esté activo para los objetos y genere objetos renderizados.

[0045] En un cuarto modo de codificación, como se ilustra en la Figura 5, el codificador de SAOC se configura para codificar mediante SAOC canales previamente renderizados; es decir, el mezclador está activo como en el segundo modo. En el lado del descodificador, se realiza la descodificación mediante SAOC para objetos previamente renderizados. de modo que se omite el procesador de objetos como en el segundo modo de codificación.

[0046] Además, existe un quinto modo de codificación, que puede realizarse mezclando cualquiera de los modos 1 a 4. En particular, existirá un modo de codificación por mezclado cuando el mezclador 1220 de la Figura 6 reciba canales directamente del descodificador de USAC y, de forma adicional, reciba canales con objetos previamente renderizados del descodificador de USAC. Además, en este modo de codificación por mezclado, los objetos se codifican directamente utilizando, preferentemente, un elemento de un solo canal del descodificador de USAC. En este contexto, el renderizador de objetos 1210 a continuación renderizará estos objetos descodificados y los enviará al mezclador 1220. Además, se codifican adicionalmente varios objetos mediante un codificador de SAOC, de modo que el descodificador de SAOC generará objetos renderizados al mezclador y/o canales renderizados cuando existan varios canales codificados mediante tecnología SAOC.

[0047] A continuación, cada porción de entrada del mezclador 1220 puede, ejemplarmente, tener al menos potencial para recibir la cantidad de canales, tal como 32, que se indica en 1205. De este modo, básicamente, el mezclador podría recibir 32 canales del descodificador de USAC y, de forma adicional, 32 canales previamente renderizados/mezclados del descodificador de USAC y, de forma adicional, 32 «canales» del renderizador de objetos y, de forma adicional, 32 «canales» del descodificador de SAOC, donde cada «canal» entre los bloques 1210 y 1218,

por un lado, y el bloque 1220, por otro, tiene una contribución de los correspondientes objetos en un canal de altavoz correspondiente y, a continuación, el mezclador 1220 mezcla, es decir, agrega las distintas contribuciones para cada canal de altavoz.

5 **[0048]** En una realización preferida de la presente invención, el sistema de codificación/descodificación se basa en un códec de USAC MPEG-D para codificar las señales de canal y objeto. Para aumentar la eficiencia para codificar una gran cantidad de objetos, se ha adaptado tecnología SAOC MPEG. Tres tipos de renderizadores realizan la tarea de renderizar objetos a los canales, renderizar canales a auriculares, o renderizar canales a una configuración de altavoces distinta. Cuando las señales de objeto se transmiten explícitamente o se codifican paramétricamente
10 mediante SAOC, la correspondiente información de metadatos de objetos se comprime y multiplexa en los datos de salida codificados.

[0049] En una realización, el prrenderizador/mezclador 200 se utiliza para convertir una escena de entrada de objeto más canal en una escena de canal antes de la codificación. Funcionalmente, es idéntica a la combinación
15 de renderizador /mezclador de objetos en el lado del descodificador que se ilustra en la Figura 4 o la Figura 6 y que se indica mediante el procesador de objetos 1200 de la Figura 2. El renderizado previo de objetos garantiza una entropía de señal determinista en la entrada del codificador que es básicamente independiente de la cantidad de señales de objeto simultáneamente activas. Con el renderizado previo de objetos, no es necesario transmitir metadatos de objetos. Las señales discretas de objeto se renderizan en la disposición de canales que el codificador esté configurado
20 para usar. El peso de los objetos para cada canal se obtiene de los metadatos de objeto OAM asociados, como indica la flecha 402.

[0050] Como codificador/descodificador/principal para señales de canal de altavoz, señales discretas de objeto, señales de *downmix* de objetos y señales previamente renderizadas, se prefiere una tecnología USAC. Realiza la
25 codificación de la multitud de señales creando información de mapeo de canal y objeto (la información geométrica y semántica de la asignación de objeto y canal de entrada). Dicha información de mapeo describe cómo los canales y objetos de entrada se asignan a elementos de canal de USAC como se ilustra en la Figura 10, es decir, elementos de par de canales (CPE, por sus siglas en inglés), elementos de un solo canal (SCE), elementos de cuádruple canal (QCE), y la correspondiente información se transmite al descodificador principal desde el codificador principal. Todas
30 las cargas útiles adicionales, como datos SAOC o metadatos de objetos, se han pasado a través de elementos de extensión y se han tenido en cuenta en el control de tasa del codificador.

[0051] Es posible realizar la codificación de objetos de distintas maneras, dependiendo de los requisitos de tasa/distorsión y los requisitos de interactividad del renderizador. Son posibles las siguientes variantes de codificación
35 de objetos:

- Objetos previamente renderizados: las señales de objeto se renderizan previamente y se mezclan con las señales del canal 22.2 antes de codificar. La posterior cadena de codificación ve señales del canal 22.2.
- Formas de onda de objetos discretos: los objetos se suministran al codificador como formas de onda monofónicas.
- 40 El codificador usa elementos de un solo canal SCE para transmitir los objetos además de las señales de canal. Los objetos descodificados se renderizan y se mezclan en el lado del receptor. La información de metadatos comprimidos de objetos se transmite al receptor/renderizador junto con estos.
- Formas de onda de objetos paramétricos: las propiedades de los objetos y su relación mutua se describen por medio de parámetros de SAOC. La *downmix* de las señales de objeto se codifica con USAC. La información paramétrica se
45 transmite junto con esta. La cantidad de canales de *downmix* se elige dependiendo de la cantidad de objetos y la tasa de datos total. La información de metadatos comprimidos de objetos se transmite al renderizador de SAOC.

[0052] El codificador y descodificador de SAOC para las señales de objeto se basan en tecnología MPEG SAOC. El sistema es capaz de recrear, modificar y renderizar una cantidad de objetos de audio basándose en una
50 menor cantidad de canales transmitidos y datos paramétricos adicionales (OLD, IOC (del inglés «*Inter Object Coherence*», coherencia entre objetos), DMG (del inglés «*Down Mix Gains*», ganancias de *downmix*)). Los datos paramétricos adicionales muestran una tasa de datos considerablemente más baja que la requerida para transmitir todos los objetos de forma individual, lo que hace que la codificación sea muy eficiente.

55 **[0053]** El codificador de SAOC toma como entrada las señales de objeto/canal como formas de onda monofónicas y genera la información paramétrica (que se empaqueta en el flujo de bits de Audio 3D) y los canales de transporte de SAOC (que se codifican mediante elementos de un solo canal y se transmiten).

[0054] El descodificador de SAOC reconstruye las señales de objeto/canal a partir de los canales de transporte
60 de SAOC descodificados y la información paramétrica, y genera la escena de audio de salida según la disposición de reproducción, la información de metadatos descomprimidos de objetos y, opcionalmente, la información de interacción con el usuario.

[0055] Para cada objeto, los metadatos asociados que especifican la posición geométrica y el volumen del
65 objeto en el espacio 3D se codifican eficazmente por cuantificación de las propiedades del objeto en el tiempo y el

espacio. Los metadatos comprimidos de objetos cOAM se transmiten al receptor en forma de información lateral. El volumen del objeto puede comprender información sobre un grado espacial y/o información sobre el nivel de señal de la señal de audio de dicho objeto de audio.

5 **[0056]** El renderizador de objetos utiliza los metadatos comprimidos de objetos para generar formas de onda de objeto según el formato de reproducción dado. Cada objeto se renderiza en determinados canales de salida según sus metadatos. La salida de este bloque es el resultado de la suma de los resultados parciales.

10 **[0057]** Si se descodifican tanto contenidos basados en canales como objetos discretos/paramétricos, las formas de onda basadas en canales y las formas de onda del objeto renderizado se mezclan antes de generar las formas de onda resultantes (o antes de introducirlas en un módulo posprocesador, como el renderizador binaural o el módulo renderizador de altavoces).

15 **[0058]** El módulo del renderizador binaural produce un *downmix* 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 cuadro a cuadro en el dominio de QMF (del inglés «*Quadrature Mirror Filterbank*», banco de filtros espejo en cuadratura).

20 **[0059]** La binauralización se basa en repuestas binaurales medidas a los impulsos en el espacio. La Figura 8 ilustra una realización preferida del conversor de formato 1720. El renderizador de altavoces o conversor de formato convierte entre la configuración de canal transmisor y el formato de reproducción deseado. Dicho conversor de formato realiza conversiones hacia una cantidad menor de canales de salida, es decir, crea *downmixes*. Para este fin, un *downmixer* 1722 que funciona preferentemente en el dominio de QMF recibe señales de salida del mezclador 1205 y genera señales de altavoz. Preferentemente, se proporciona un controlador 1724 para configurar el *downmixer* 1722 que recibe, como entrada de control, una disposición de salida del mezclador, es decir, la disposición para lo que se determinan los datos 1205, y normalmente se ha introducido una disposición de reproducción deseada en el bloque de conversión de formato 1720 que se ilustra en la Figura 6. Basándose esta información, el controlador 1724 genera, preferentemente de forma automática, matrices de *downmix* optimizadas para la combinación dada de formatos de entrada y salida y aplica dichas matrices en el bloque del *downmixer* 1722 en el proceso de *downmix*. El conversor de formato permite tanto las configuraciones de altavoces estándar como configuraciones aleatorias con posiciones de altavoz no estándar.

35 **[0060]** Como se ilustra en relación con la Figura 6, el descodificador de SAOC está diseñado para renderizar en la disposición de canales predefinida, como 22.2, con una conversión posterior de formato a la disposición de reproducción buscada. No obstante, y de forma alternativa, el descodificador de SAOC se implementa para admitir el modo de «baja energía», donde el descodificador de SAOC se configura para descodificar directamente a la disposición de reproducción, sin la posterior conversión de formato. En esta implementación, el descodificador de SAOC 1800 genera directamente la señal de altavoz, como las señales de altavoz 5.1, y el descodificador de SAOC 1800 requiere la información de la disposición de reproducción y la matriz de renderizado para que pueda funcionar el paneo de amplitud basado en vectores o cualquier otra clase de procesador para generar información de *downmix*.

45 **[0061]** La Figura 9 ilustra una realización adicional del renderizador binaural 1710 de la Figura 6. Específicamente para dispositivos móviles, se requiere el renderizado binaural para auriculares conectados a dichos dispositivos móviles o para altavoces directamente conectados a dispositivos móviles normalmente pequeños. Para dichos dispositivos móviles, pueden existir restricciones que limiten la complejidad del descodificador y el renderizado. Además de omitir la decorrelación en este tipo de escenario de procesamiento, se prefiere realizar primero *downmixing* mediante el *downmixer* 1712 a un *downmix* intermedio, es decir, a una cantidad menor de canales de salida, que a su vez se traduce en una menor cantidad de canales de entrada para el conversor binaural 1714. Ejemplarmente, el *downmixer* 1712 somete a *downmixing* el material del canal 22.2 para crear un *downmix* intermedio 5.1 o, 50 alternativamente, el *downmix* intermedio se calcula directamente mediante el descodificador de SAOC 1800 de la Figura 6 en una especie de «atajo». A continuación, el renderizado binaural únicamente tiene que aplicar diez funciones HRTF (funciones de transferencia relacionadas con la cabeza) o BRIR para renderizar los cinco canales individuales en diferentes posiciones, en comparación con aplicar 44 HRTF para funciones BRIR si los canales de entrada 22.2 ya se hubiesen renderizado directamente. Específicamente, las operaciones de convolución necesarias 55 para un renderizado binaural requieren una gran capacidad de procesamiento y, por lo tanto, la reducción de dicha capacidad de procesamiento mientras se sigue obteniendo una calidad de audio aceptable resulta particularmente útil para dispositivos móviles.

60 **[0062]** Preferentemente, el «atajo» que se ilustra mediante la línea de control 1727 comprende controlar el descodificador 1300 para descodificar a una cantidad menor de canales, es decir, omitir el bloque de procesamiento OTT completo en el descodificador o una conversión de formato a una cantidad menor de canales y, como se ilustra en la Figura 9, el renderizado binaural se realiza para la menor cantidad de canales. El mismo procesamiento puede aplicarse no solamente al procesamiento binaural, sino también a una conversión de formato, como se ilustra mediante la línea 1727 en la Figura 6.

65

[0063] En una realización adicional, se requiere una interfaz eficiente entre bloques de procesamiento. En particular en la Figura 6, se representa la ruta de la señal de audio entre los diferentes bloques de procesamiento. Tanto el renderizador binaural 1710 como el conversor de formato 1720, el descodificador de SAOC 1800 y el descodificador de USAC 1300, en caso de que se aplique SBR (del inglés «*Spectral Band Replication*», replicación de la banda espectral), funcionan en un dominio QMF o QMF híbrido. Según una realización, todos estos bloques de procesamiento proporcionan una interfaz QMF o QMF híbrida para permitir el pase de señales de audio entre sí en el dominio QMF de forma eficiente. De forma adicional, se prefiere implementar el módulo del mezclador y el módulo del renderizador de objetos para que también funcionen en el QMF o dominio QMF híbrido. Como consecuencia, pueden evitarse etapas de síntesis y análisis separadas de QMF o QMF híbrido, lo que se traduce en un ahorro significativo de complejidad y, a continuación, únicamente se requiere una etapa de síntesis de QMF final para generar los altavoces que se indican en 1730, o para generar los datos binaurales en la salida del bloque 1710, o para generar las señales de altavoz de disposición de reproducción en la salida del bloque 1720.

[0064] Posteriormente, se hace referencia a la Figura 11 con el fin de explicar los elementos de cuádruple canal (QCE). A diferencia de un elemento de par de canales, tal y como se define en el estándar US AC-MPEG, un elemento de cuádruple canal requiere cuatro canales de entrada 90 y genera un elemento QCE codificado 91. En una realización, se proporciona una jerarquía de dos módulos MPEG Surround en Modo 2-1-2 o dos módulos TTO (del inglés «*Two To One*», dos a uno) y herramientas adicionales de codificación de Joint Stereo (por ejemplo, MS-Stereo), como se define en MPEG USAC o MPEG Surround, y el elemento QCE no solamente comprende dos canales de *downmix* codificados en Joint Stereo y opcionalmente, dos canales residuales codificados en Joint Stereo y, de forma adicional, datos paramétricos derivados de los, por ejemplo, dos módulos TTO. En el lado del descodificador, se aplica una estructura donde se aplica la descodificación de Joint Stereo de los dos canales de *downmix* y, opcionalmente, de los dos canales residuales y, en una segunda etapa con dos módulos OTT, los canales de *downmix* y canales residuales opcionales se someten a *upmixing* hasta los cuatro canales de salida. Sin embargo, pueden aplicarse operaciones de procesamiento alternativas para un codificador QCE en lugar del funcionamiento jerárquico. De este modo, además de la codificación de canal conjunta de un grupo de dos canales, el codificador/descodificador principal utiliza adicionalmente una codificación de canal conjunta de un grupo de cuatro canales.

[0065] Además, se prefiere realizar un procedimiento de relleno de ruido optimizado para permitir la codificación de banda completa sin pérdidas (18 kHz) a 1200 kbps.

[0066] El codificador se ha utilizado en un modo de «tasa constante con reserva de bits», utilizando un máximo de 6144 bits por canal como buffer de tasa para los datos dinámicos.

[0067] Todas las cargas útiles adicionales, como datos de SAOC o metadatos de objetos, se han pasado a través de elementos de extensión y se han tenido en cuenta en el control de tasa del codificador.

[0068] Con el fin de aprovechar las funcionalidades de SAOC también para el contenido de audio 3D, se han implementado las siguientes extensiones de MPEG SAOC:

- *Downmixing* a una cantidad arbitraria de canales de transporte de SAOC.
- Renderizado optimizado para configuraciones de salida con un número elevado de altavoces (hasta 22.2).

[0069] El módulo del renderizador binaural produce un *downmix* binaural del material de audio multicanal, de modo que cada canal de entrada (sin incluir los canales LFE) está representado por una fuente de sonido virtual. El procesamiento se realiza cuadro a cuadro en el dominio QMF.

[0070] La binauralización se basa en repuestas binaurales medidas a los impulsos en el espacio. El sonido directo y las reflexiones tempranas quedan grabados en el material de audio mediante un enfoque de convolución en un dominio pseudo FFT utilizando una convolución rápida sobre el dominio QMF.

[0071] Aunque algunos aspectos se han descrito en relación con un aparato, resulta evidente que dichos aspectos también representan una descripción del procedimiento correspondiente, donde un bloque o dispositivo corresponde a una etapa del procedimiento o una característica de una etapa del procedimiento. De forma similar, los aspectos descritos en relación con una etapa del procedimiento también representan una descripción de un bloque o elemento o característica correspondiente de un aparato correspondiente. Algunas o todas las etapas del procedimiento pueden ejecutarse mediante (o utilizando) un aparato de *hardware*, como, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, alguna o más de las etapas más importantes del procedimiento pueden ser ejecutadas por dicho aparato.

[0072] Dependiendo de determinados requisitos de implementación, las realizaciones de la invención pueden implementarse en *hardware* o en *software*. La implementación puede realizarse utilizando un medio de almacenamiento no transitorio, tal como un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un Blu-Ray, un CD, una ROM, una PROM, y EPROM, una memoria EEPROM o flash, que tenga almacenadas en su interior señales de control legibles de manera electrónica, que cooperan (o son capaces de cooperar) con un sistema

informático programable, de modo que se realice el respectivo procedimiento. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

5 **[0073]** Algunas realizaciones según la invención comprenden un soporte de datos que tiene señales de control legibles de manera electrónica, que son capaces de cooperar con un sistema informático programable, de modo que se realice uno de los procedimientos descritos en esta invención.

10 **[0074]** En general, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa, estando operativo el código de programa para realizar uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede, por ejemplo, almacenarse en un soporte legible por máquina.

15 **[0075]** Otras realizaciones comprenden el programa informático para realizar uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por máquina.

[0076] Dicho de otro modo, una realización del procedimiento inventivo es, por lo tanto, un programa informático que tiene un código de programa para realizar uno de los procedimientos descritos en esta invención, cuando el programa informático se ejecuta en un ordenador.

20 **[0077]** Una realización adicional del procedimiento inventivo es, por lo tanto, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en este, el programa informático para realizar uno de los procedimientos descritos en esta invención. El soporte de datos, el medio de almacenamiento digital o el medio grabado normalmente son tangibles y/o no transitorios.

25 **[0078]** Una realización adicional del procedimiento inventivo es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa informático para realizar uno de los procedimientos descritos en esta invención. El flujo de datos o la secuencia de señales pueden, por ejemplo, configurarse para transferirse a través de una conexión de comunicación de datos, por ejemplo, a través de la Internet.

30 **[0079]** Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador o un dispositivo lógico programable, configurado para, o adaptado para, realizar uno de los procedimientos descritos en esta invención.

35 **[0080]** Una realización adicional comprende un ordenador que tiene instalado en este el programa informático para realizar uno de los procedimientos descritos en esta invención.

40 **[0081]** Una realización adicional según la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente u ópticamente) un programa informático para realizar uno de los procedimientos descritos en esta invención a un receptor. El receptor puede, por ejemplo, ser un ordenador, un dispositivo móvil, un dispositivo de memoria o similares. El aparato o sistema puede, por ejemplo, comprender un servidor de archivos para transferir el programa informático al receptor.

45 **[0082]** En algunas realizaciones, puede usarse un dispositivo lógico programable (por ejemplo, una matriz de puertas programables en campo) para realizar todas o algunas de las funcionalidades de los procedimientos descritos en esta invención. En algunas realizaciones, una matriz de puertas programables en campo puede cooperar con un microprocesador para realizar uno de los procedimientos descritos en esta invención. En general, los procedimientos se realizan preferentemente mediante cualquier aparato de *hardware*.

50 **[0083]** Las realizaciones anteriormente descritas son meramente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en esta invención resultarán evidentes a otros expertos en la materia. Por lo tanto, se pretende quedar limitado únicamente por el alcance de las reivindicaciones de patente en curso y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones de esta invención.

REIVINDICACIONES

1. Codificador de audio para descodificar datos de audio codificados, que comprende:
 - 5 una interfaz de entrada (1100) configurada para recibir los datos de audio codificados, comprendiendo los datos de audio codificados una pluralidad de canales de audio codificados o una pluralidad de objetos de audio codificados y metadatos comprimidos relacionados con la pluralidad de objetos de audio, y una indicación de modo; un descodificador principal (1300) configurado para descodificar la pluralidad de canales de audio codificados y la pluralidad de objetos de audio codificados;
 - 10 un descompresor de metadatos (1400) configurado para descomprimir los metadatos comprimidos; un procesador de objetos (1200) configurado para procesar la pluralidad de objetos de audio descodificados utilizando los metadatos descomprimidos para obtener una cantidad de canales de audio de salida (1205) que comprende datos de audio a partir de los objetos de audio y los canales de audio descodificados;
 - 15 un controlador de modos (1600) conectado a la interfaz de entrada (1100) y configurado para analizar los datos de audio codificados para detectar la indicación de modo que indique un primer modo o un segundo modo, donde, en el primer modo, los datos de audio codificados comprenden canales de audio codificados y objetos de audio codificados y donde, en el segundo modo, los datos de audio codificados comprenden únicamente la pluralidad de canales de audio codificados; y
 - 20 un posprocesador (1700) configurado para convertir la cantidad de canales de audio de salida (1205) en un formato de salida, donde el descodificador de audio, controlado por el controlador de modo (1600), está configurado para omitir el procesador de objetos (1200) y para introducir una pluralidad de canales descodificados en el posprocesador (1700) cuando el controlador de modo (1600) ha detectado el segundo modo, y para introducir la pluralidad de objetos de audio descodificados y la pluralidad de canales de audio descodificados en el procesador de objetos
 - 25 (1200) cuando el controlador de modo (1600) ha detectado el primer modo.
2. El descodificador de audio según la reivindicación 1, donde el posprocesador (1700) configurado para convertir la cantidad de canales de audio de salida (1205) en una representación binaural o en un formato de reproducción que tiene una menor cantidad de canales de audio que la cantidad de canales de audio de salida (1205),
 - 30 y donde el descodificador de audio está configurado para controlar el posprocesador (1700) según una entrada de control derivada de una interfaz del usuario o extraída de los datos de audio codificados.
3. El descodificador de audio según la reivindicación 1 o 2, donde el procesador de objetos (1200)
 - 35 comprende:
 - un renderizador de objetos (1210) para renderizar los objetos de audio descodificados con el fin de obtener objetos de audio renderizados utilizando los metadatos descomprimidos; y
 - 40 un mezclador (1220) para mezclar los objetos de audio renderizados y los canales de audio descodificados para obtener la cantidad de canales de audio de salida (1205).
4. El descodificador de audio según cualquiera de las reivindicaciones 1 a 3, donde el procesador de objetos (1200) comprende:
 - 45 un descodificador de codificación de objetos de audio espaciales (1800) para descodificar uno o más canales de transporte y la información lateral paramétrica asociada que representa objetos de audio codificados, donde el descodificador de codificación de objetos de audio espaciales (1800) está configurado para renderizar los objetos de audio codificados según la información de renderizado relacionada con una colocación de los objetos de audio para obtener los objetos de audio renderizados y para controlar el procesador de objetos (1200) con el fin de mezclar los objetos de audio renderizados y los canales de audio descodificados para obtener la cantidad de canales de audio de
 - 50 salida (1205).
5. El descodificador de audio según cualquiera de las reivindicaciones 1 a 3, donde el procesador de objetos (1200) comprende un descodificador de codificación de objetos de audio espaciales (1800) para descodificar uno o más canales de transporte y la información lateral paramétrica asociada que representa objetos de audio
 - 55 codificados y canales de audio codificados, donde el descodificador de codificación de objetos de audio espaciales (1800) está configurado para descodificar los objetos de audio codificados y canales de audio codificados utilizando los uno o más canales de transporte y la información lateral paramétrica, y
 - 60 donde el procesador de objetos (1200) está configurado para renderizar la pluralidad de objetos de audio utilizando los metadatos descomprimidos para obtener objetos de audio renderizados y para descodificar los canales de audio y para mezclar los canales de audio con los objetos de audio renderizados para obtener la cantidad de canales de audio de salida (1205).
6. El descodificador de audio según cualquiera de las reivindicaciones 1 a 3, donde el procesador de

objetos (1200) comprende un descodificador de codificación de objetos de audio espaciales (1800) para descodificar uno o más canales de transporte y la información lateral paramétrica asociada que representa objetos de audio codificados o canales de audio codificados,

- 5 donde el descodificador de codificación de objetos de audio espaciales (1800) está configurado para transcodificar la información paramétrica asociada y los metadatos descomprimidos en información lateral paramétrica transcodificada susceptible de usarse para renderizar directamente el formato de salida, y donde el posprocesador (1700) está configurado para calcular canales de audio del formato de salida utilizando los canales de transporte descodificados y la información lateral paramétrica transcodificada, o
- 10 donde el descodificador de codificación de objetos de audio espaciales (1800) está configurado para someter directamente a *upmixing* y renderizar señales de canal para el formato de salida utilizando los canales de transporte descodificados y la información paramétrica lateral.

7. El descodificador de audio según cualquiera de las reivindicaciones anteriores,

- 15 donde el procesador de objetos (1200) comprende un descodificador de codificación de objetos de audio espaciales (1800) para descodificar uno o más canales de transporte generados por el descodificador principal (1300) y la información lateral paramétrica asociada y los metadatos descomprimidos para obtener una pluralidad de objetos de audio renderizados,
- 20 donde el procesador de objetos (1200) comprende un renderizador de objetos (1210) que está configurado para renderizar los objetos de audio descodificados generados por el descodificador principal (1300) para obtener objetos de audio descodificados renderizados;
- donde el procesador de objetos (1200) comprende un mezclador (1220) que está configurado para mezclar los objetos de audio renderizados descodificados y la pluralidad de objetos de audio renderizados con los canales de
- 25 audio descodificados,
- donde el descodificador de audio comprende, además, una interfaz de salida (1730) para generar una salida del mezclador (1220) a los altavoces;
- donde el posprocesador (1700) comprende, además:
- 30 un renderizador binaural (1700) para renderizar los canales de audio de salida (1205) en dos canales binaurales utilizando funciones de transferencia relacionadas con la cabeza o respuestas binaurales a los impulsos, representando los dos canales binaurales la representación binaural, y
- un conversor de formato (1720) para convertir los canales de audio de salida (1205) en el formato de salida que tiene una cantidad menor de canales de audio que los canales de audio de salida (1205) del mezclador
- 35 utilizando información sobre una disposición de reproducción.

8. El descodificador de audio según cualquiera de las reivindicaciones 1 a 7,

- 40 donde la pluralidad de canales de audio codificados o la pluralidad de objetos de audio codificados se codifican como elementos de par de canales, elementos de un solo canal, elementos de baja frecuencia o elementos de cuádruple canal, donde un elemento de cuádruple canal comprende cuatro canales de audio u objetos de audio originales, y
- donde el descodificador principal (1300) está configurado para descodificar los elementos de par de canales, los elementos de un solo canal, los elementos de baja frecuencia, o los elementos de cuádruple canal según la
- 45 información lateral incluida en los datos de audio codificados que indica un elemento de par de canales, un elemento de un solo canal, un elemento de baja frecuencia, o un elemento de cuádruple canal.

9. El descodificador de audio según cualquiera de las reivindicaciones 1 a 8,

- 50 donde el descodificador principal (1300) está configurado para aplicar una operación de descodificación de banda completa utilizando una operación de relleno de ruido.

10. El descodificador de audio según la reivindicación 7, donde los elementos que comprenden el renderizador binaural (1710), el conversor de formato (1720), el mezclador (1220), el descodificador de SAOC (1800) y el descodificador principal (1300) y el renderizador de objetos (1210) funcionan en un dominio de banco de filtros espejo en cuadratura (QMF) y donde los datos del dominio de filtros espejo en cuadratura se transmiten de uno de los

55 elementos a otro de los elementos sin ningún banco de filtros de síntesis ni procesamiento posterior de análisis mediante banco de filtros.

11. El descodificador de audio según cualquiera de las reivindicaciones 1 a 10,

- 60 donde el posprocesador (1700) está configurado para someter a *downmixing* la cantidad de canales de audio de salida (1205) generados por el procesador de objetos (1200) hasta un formato que tiene tres o más canales de audio y que tiene menos canales de audio que la cantidad de canales de audio de salida (1205) generados por el procesador de objetos (1200) para obtener canales de un *downmix* intermedio, y para renderizar binauralmente (1710) los canales del *downmix* intermedio en la representación binaural que tiene una señal de salida binaural de dos canales.

65

12. El descodificador de audio según cualquiera de las reivindicaciones 1 a 8, donde el posprocesador (1700) comprende:

- 5 un *downmixer* controlado (1722) para aplicar una matriz de *downmix*; y
- un controlador (1724) para determinar una matriz de *downmixing* específica utilizando información sobre una configuración de canales de una salida del procesador de objetos (1200) e información sobre una disposición de reproducción pretendida.

13. El descodificador de audio según cualquiera de las reivindicaciones 1 a 12,

- 10 donde el descodificador principal (1300) o el procesador de objetos (1200) son controlables, y
- donde el posprocesador (1700) está configurado para controlar el descodificador principal (1300) o el procesador de objetos (1200) según información sobre el formato de salida, de modo que se reduce o elimina un renderizado que conlleve el procesamiento de decorrelación de objetos de audio o canales de audio que no aparecen como
- 15 canales de audio separados en el formato de salida, o de modo que, para los objetos de audio o canales de audio que no aparecen como los canales de audio separados en el formato de salida, se realizan operaciones de *upmixing* o descodificación como si los objetos de audio o canales de audio aparecieran como los canales separados en el formato de salida, excepto que se desactiva todo procesamiento de decorrelación de los objetos de audio o los canales de audio que no aparecen como los canales de audio separados en el formato de salida.

14. El descodificador de audio según la reivindicación 8,

- donde el descodificador principal (1300) está configurado para realizar descodificación por transformación y una descodificación con replicación de la banda espectral para el elemento de un solo canal, y para realizar descodificación por transformación, descodificación estéreo paramétrica y descodificación con reproducción de la banda espectral
- 25 para los elementos de par de canales y los elementos de cuádruple canal.

15. Procedimiento de descodificación de datos de audio codificados, que comprende:

- la recepción (1100) de los datos de audio codificados, comprendiendo los datos de audio codificados o bien una
- 30 pluralidad de canales de audio codificados y una pluralidad de objetos de audio codificados y metadatos comprimidos relacionados con la pluralidad de objetos de audio, o una pluralidad de canales de audio codificados sin objetos de audio codificados, y una indicación de modo;
- la descodificación principal (1300) de la pluralidad de canales de audio codificados y la pluralidad de objetos de audio codificados;
- 35 la descompresión (1400) de los metadatos comprimidos;
- el procesamiento (1200) de la pluralidad de objetos de audio descodificados utilizando los metadatos comprimidos para obtener una cantidad de canales de audio de salida (1205) que comprende datos de audio de los objetos de audio y los canales de audio descodificados;
- el análisis, mediante un controlador de modos (1600), de los datos de audio codificados para detectar una
- 40 indicación de modo que indique un primer modo o un segundo modo, donde, en el primer modo, los datos de audio codificados comprenden canales de audio codificados y objetos de audio codificados y donde, en el segundo modo, los datos de audio codificados comprenden únicamente la pluralidad de canales de audio codificados; y
- la conversión (1700) de la cantidad de canales de audio de salida (1205) en un formato de salida,
- 45 donde, en el procedimiento de descodificación de audio y controlado por el controlador de modos (1600), se omite el procesamiento (1200) de la pluralidad de objetos descodificados y se introduce una pluralidad de canales descodificados en la conversión (1700) cuando el controlador de modo (1600) ha detectado el segundo modo, y
- donde la pluralidad de objetos de audio descodificados y la pluralidad de canales de audio descodificados se introducen en el procesamiento (1200) la pluralidad de objetos descodificados, cuando el controlador de modo (1600) ha detectado el primer modo.

16. Un programa informático que comprende instrucciones que, cuando el programa se ejecuta mediante un ordenador o un procesador, hacen que el ordenador o procesador ejecute el procedimiento según la reivindicación 15.

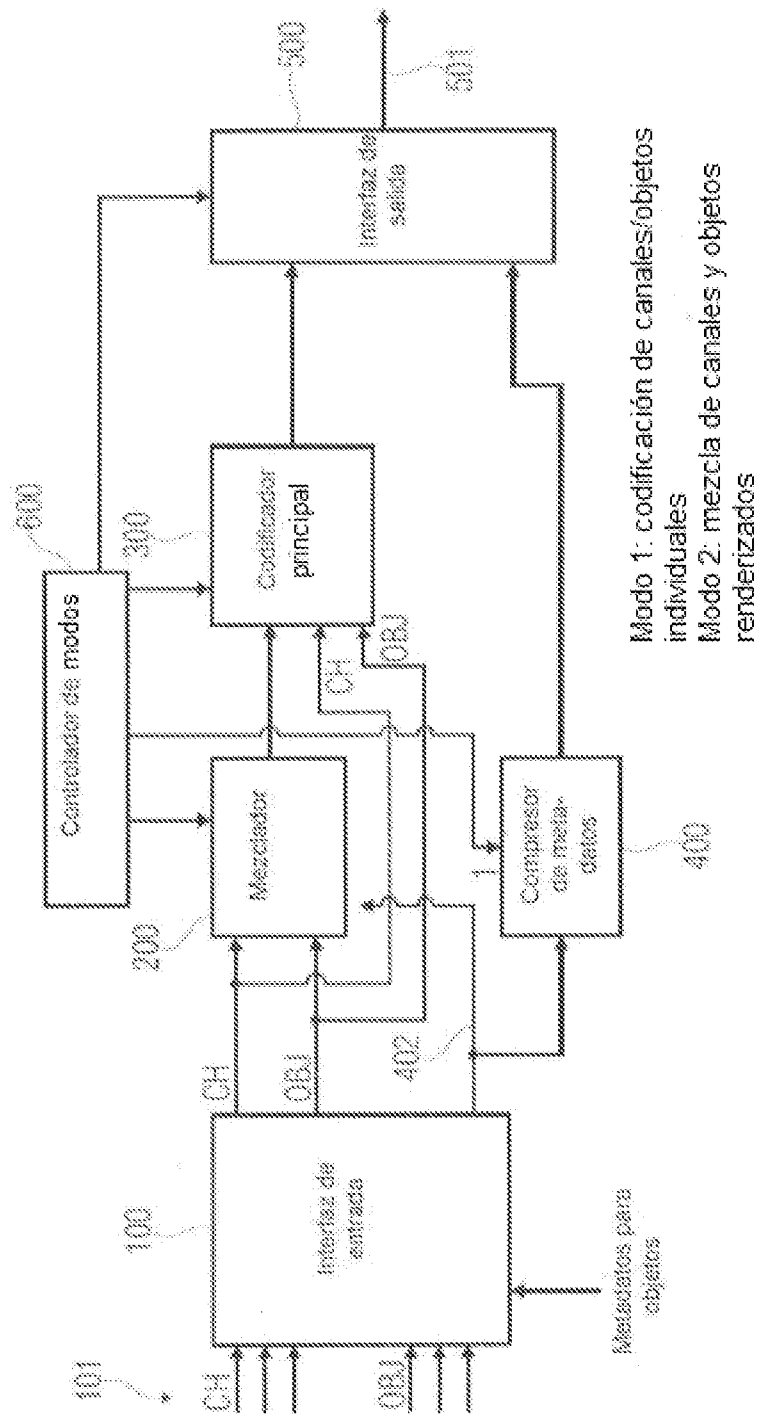


FIG. 1
(CODIFICADOR)

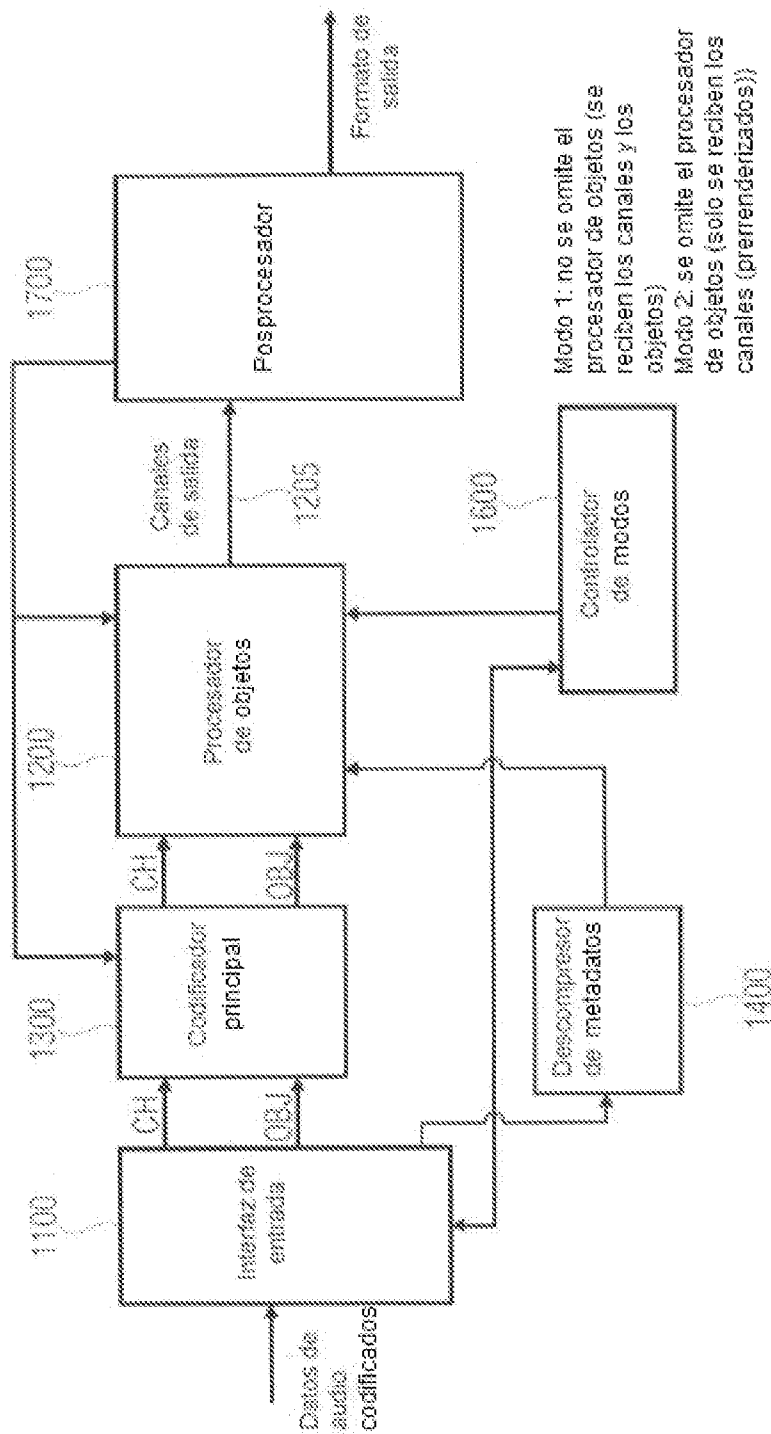


FIG. 2
(DESCODIFICADOR)

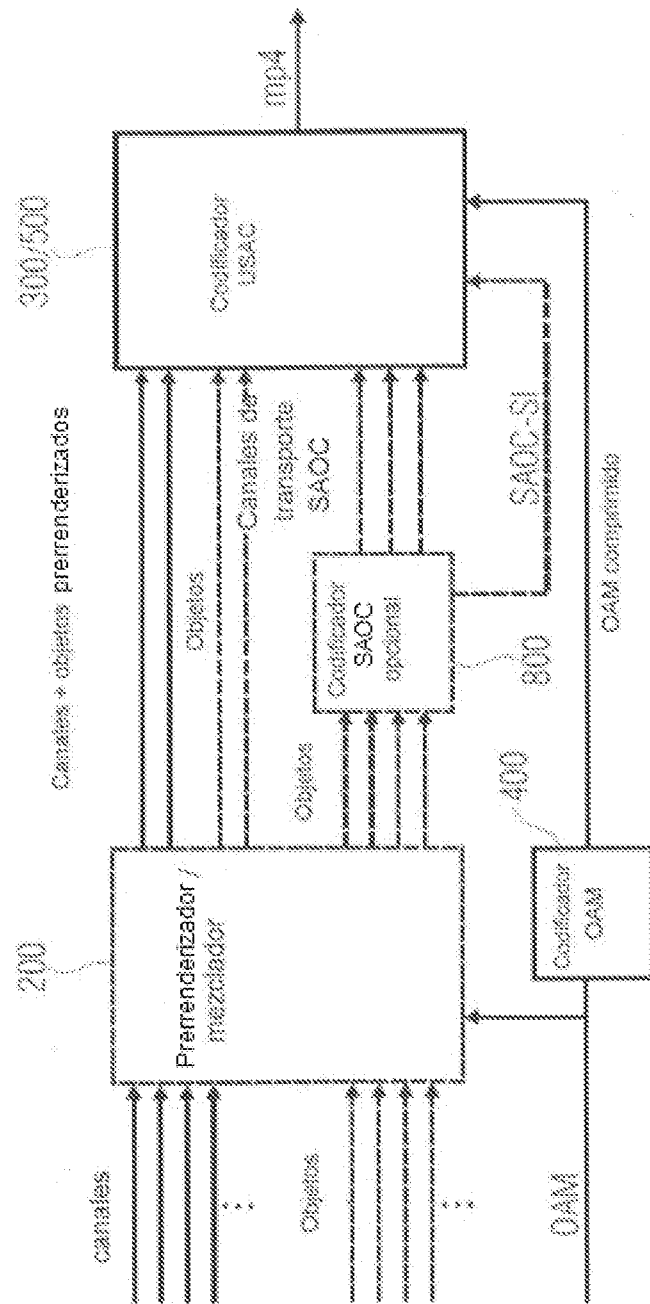


FIG. 3
(CODIFICADOR)

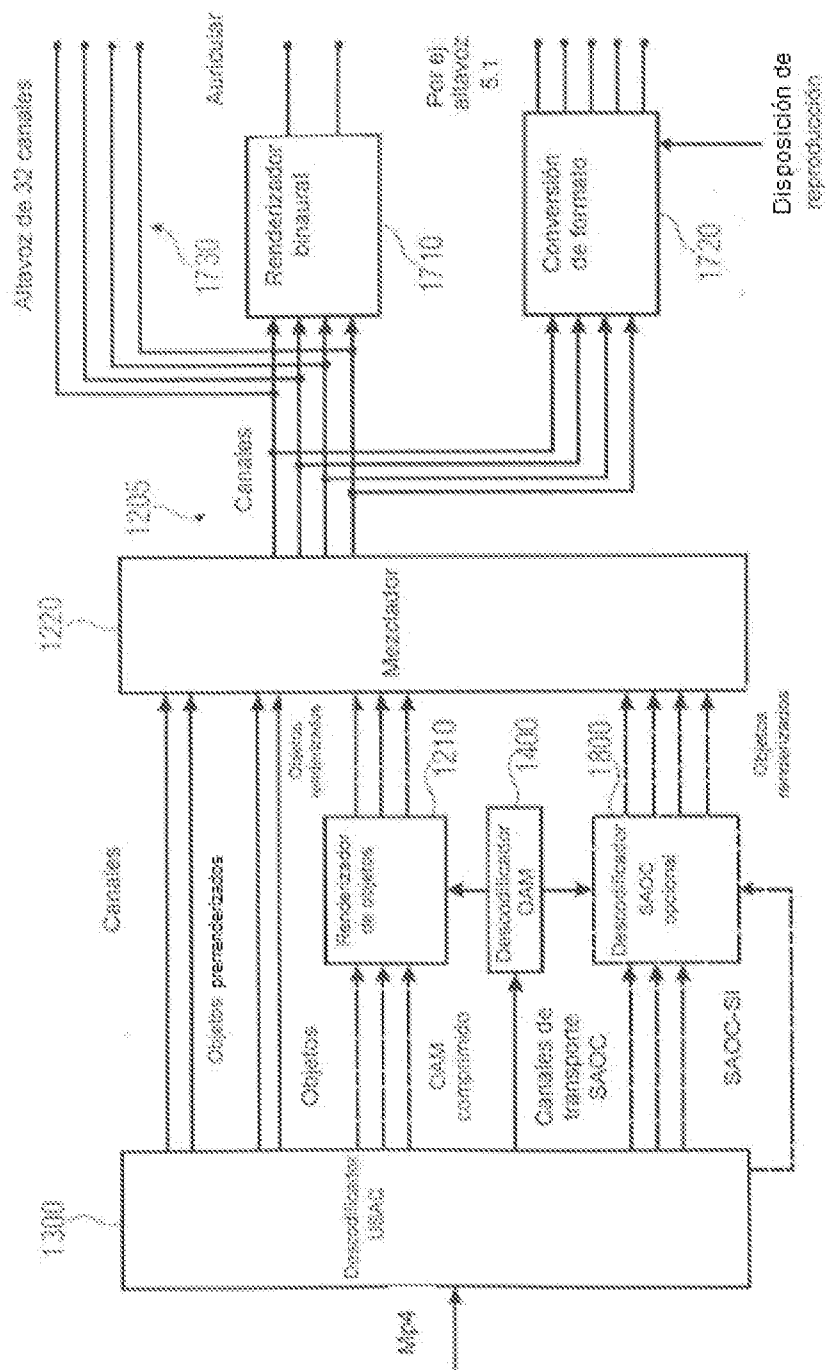


FIG. 4
(DESCODIFICADOR)

Dominio QMF directo.
 Procesamiento en renderizador binaural,
 conversor de formato, descodificador
 STOC, descodificador USAC en modo
 SBR

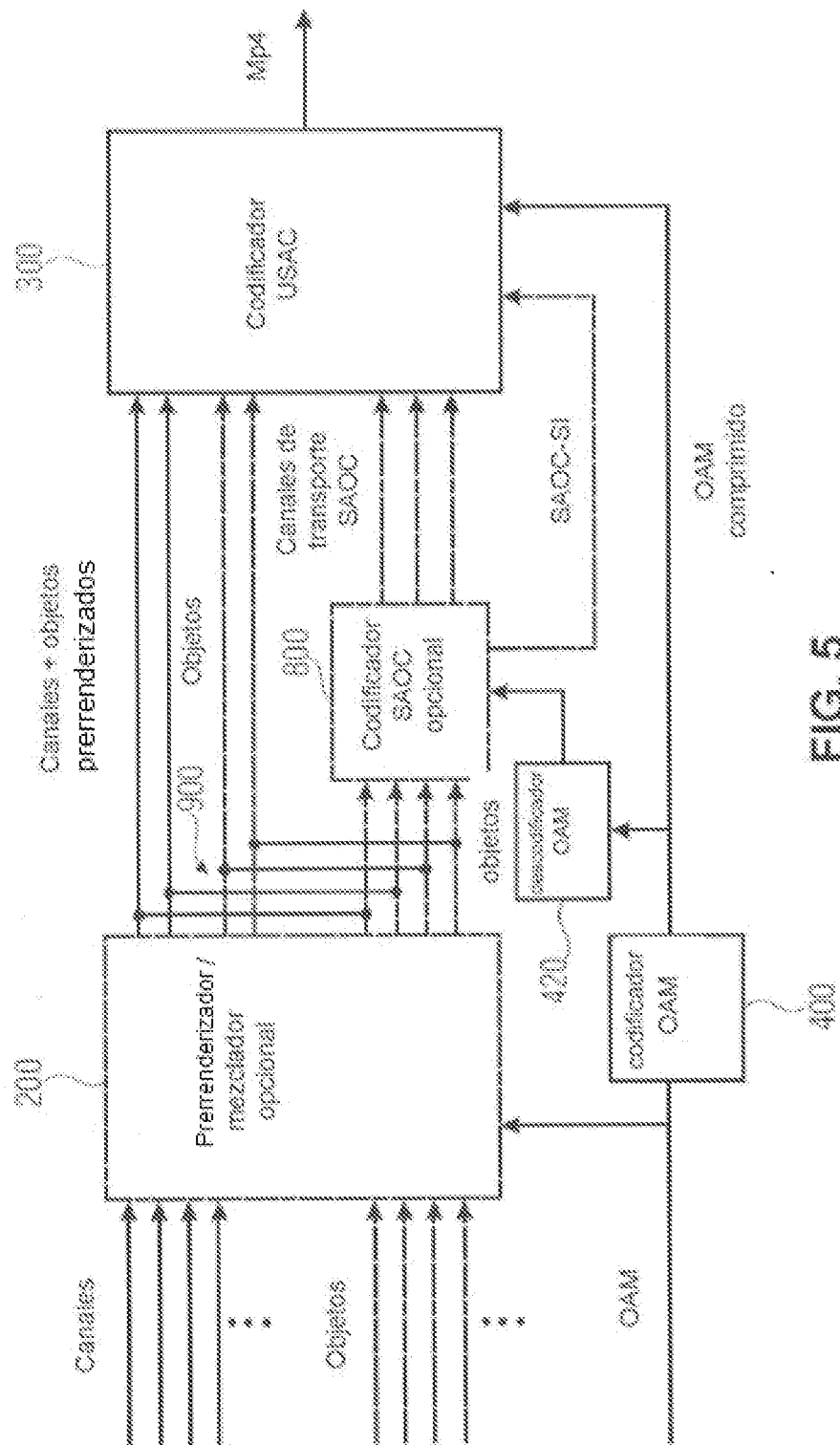


FIG. 5
(CODIFICADOR)

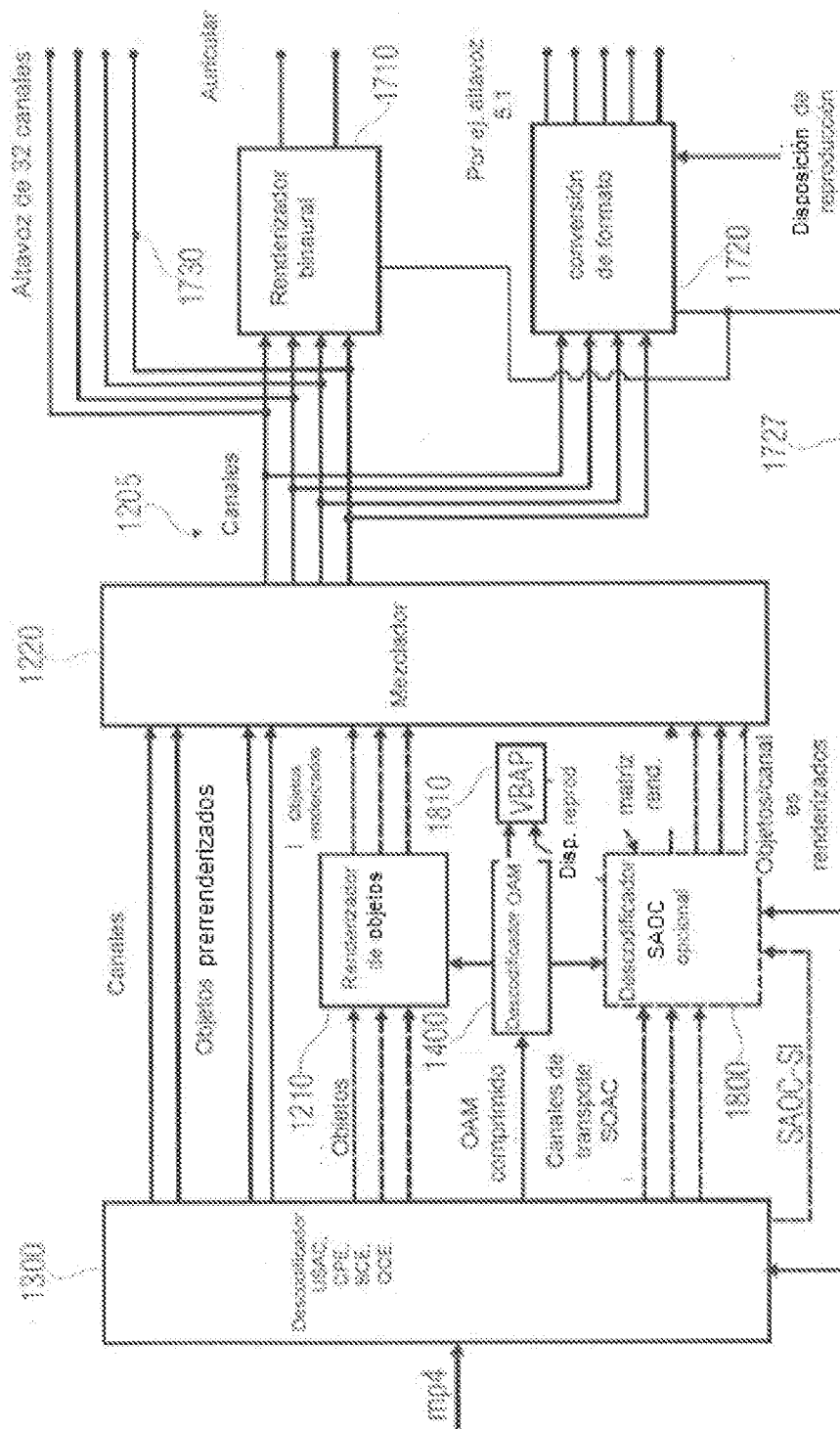


FIG. 6
(DESCODIFICADOR)

modo	codificador	descodificador
1	se omite el mezclador	no se omite el procesador de objetos
2	mezclador activo	se omite el procesador de objetos
3	Codificación SAOC solo para objetos	descodificación SAOC solo para objetos
4	Codificación SAOC para canales prrenderizados/ mezclador activo	Descodificación SAOC para objetos prrenderizados (se omite el proc. de obj.)
5	Cualquier mezcla de modos 1 a 4	Cualquier mezcla de modos 1 a 4

FIG. 7

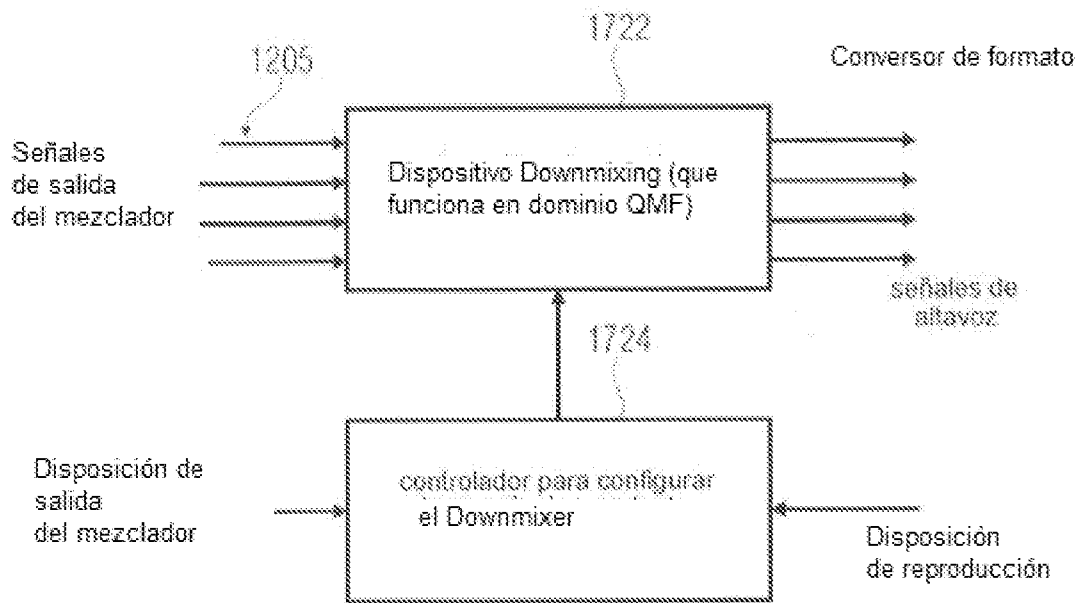


FIG. 8

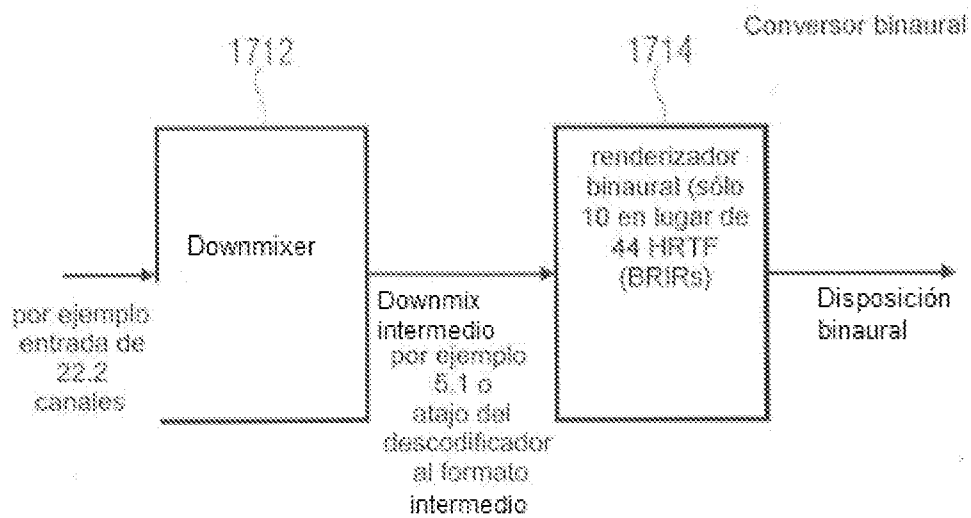


FIG. 9

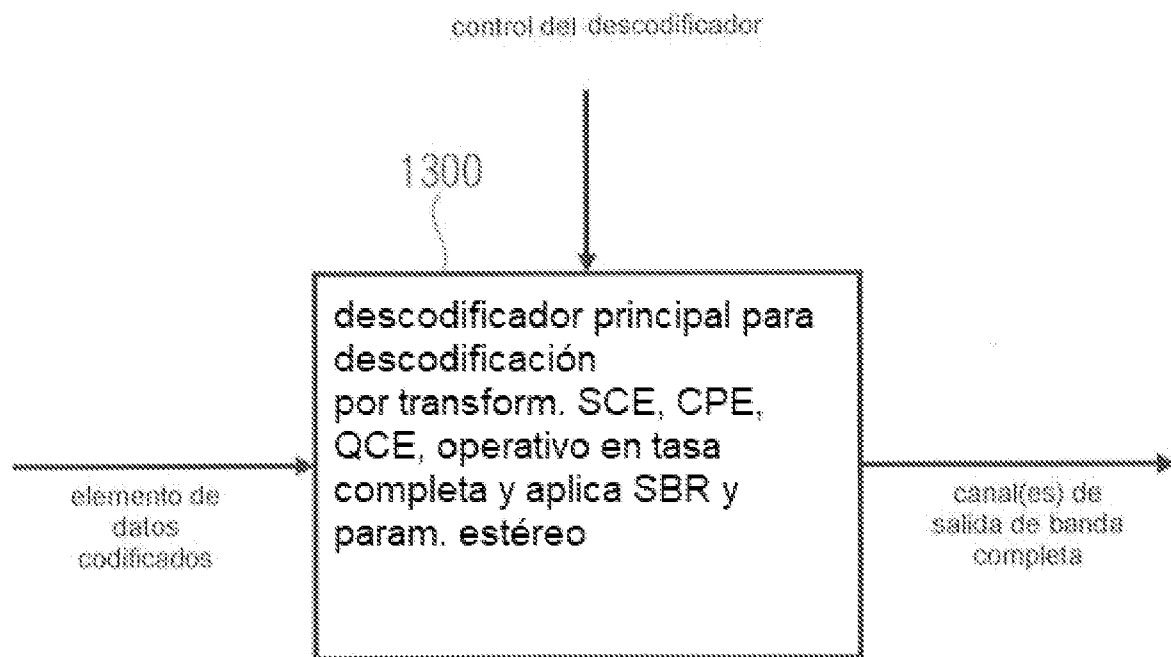


FIG. 10

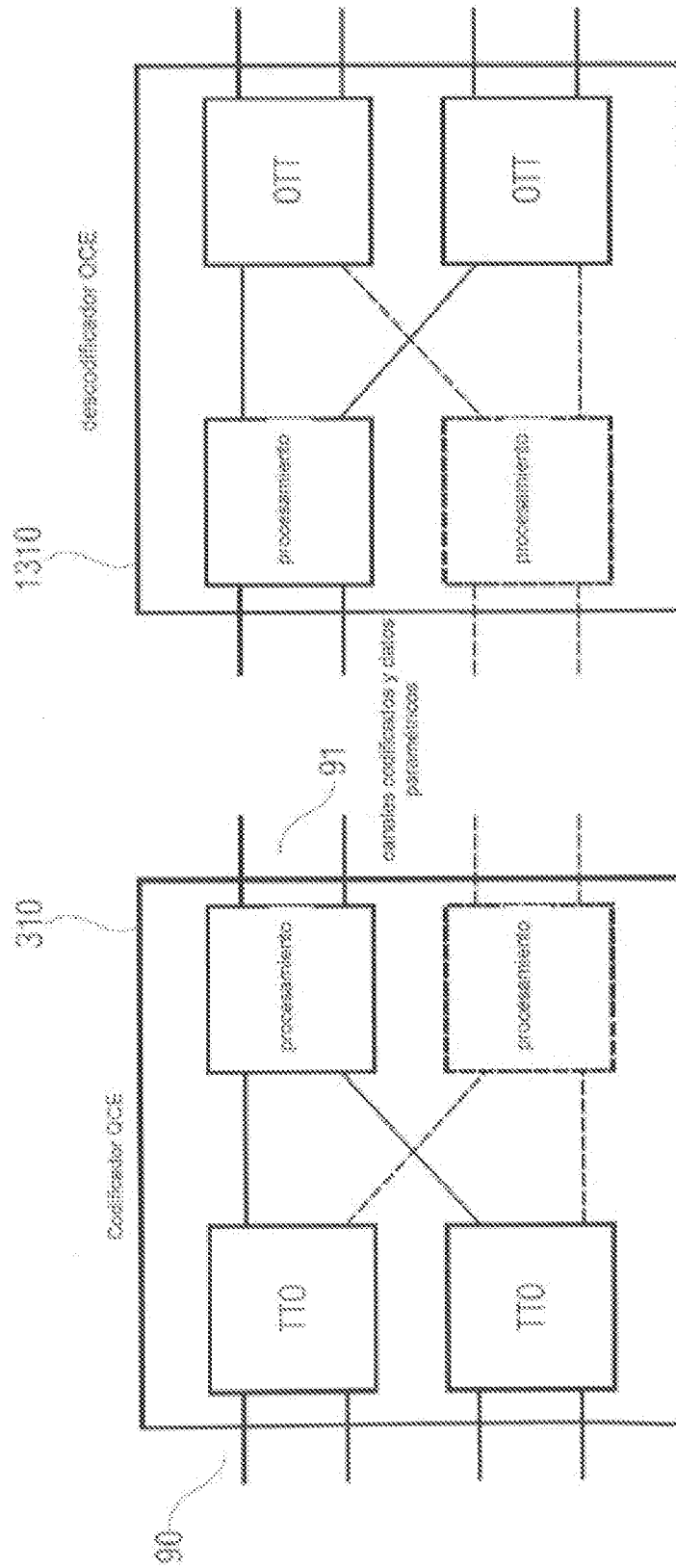


FIG. 11