

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 734 378**

51 Int. Cl.:

H04S 7/00 (2006.01)

H04S 5/00 (2006.01)

G10L 19/008 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.07.2014** **PCT/EP2014/065037**

87 Fecha y número de publicación internacional: **29.01.2015** **WO15010937**

96 Fecha de presentación y número de la solicitud europea: **14.07.2014** **E 14753020 (8)**

97 Fecha y número de publicación de la concesión europea: **01.05.2019** **EP 3025521**

54 Título: **Mezcla ascendente espacial controlada por renderizador**

30 Prioridad:

22.07.2013 EP 13177368

18.10.2013 EP 13189285

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
05.12.2019

73 Titular/es:

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

72 Inventor/es:

**ERTEL, CHRISTIAN;
HILPERT, JOHANNES;
HÖLZER, ANDREAS;
KUNTZ, ACHIM;
PLOGSTIES, JAN y
KRATSCHMER, MICHAEL**

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 734 378 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

Mezcla ascendente espacial controlada por renderizador

- 5 La presente invención se refiere al procesamiento de señales de audio, y, en particular, a la conversión de formato de señales de audio de multicanal.

La conversión de formato describe el proceso de mapeo de una cierta cantidad de canales de audio en otra representación adecuada para la reproducción por medio de una cantidad diferente de canales de audio.

- 10 Un caso de uso común para la conversión de formato es la mezcla descendente de canales de audio. En la ref. [1], se proporciona un ejemplo, donde la mezcla descendente permite a los usuarios finales la nueva reproducción de una versión del material fuente 5.1 aun cuando no esté disponible un sistema de monitoreo completo 'home-theatre' 5.1. El equipamiento diseñado para aceptar material Dolby Digital, pero que proporciona solo salidas mono o estéreo (por ejemplo, reproductores de DVD portátiles, decodificadores y demás), incorpora instalaciones para la mezcla descendente de los canales 5.1 originales al uno o dos canales de salida, como norma.

- 15 Por otra parte, la conversión de formato además puede describir un proceso de mezcla ascendente, por ejemplo, la mezcla ascendente de material estéreo para formar una versión compatible con 5.1. Además, puede considerarse la renderización binaural como conversión de formato.

- 25 A continuación, se describen implicaciones de conversión de formato para el proceso de decodificación de señales de audio comprimidas. En este caso, la representación comprimida de la señal de audio (archivo mp4) representa una cantidad fija de canales de audio previstos para la reproducción por un contexto de altavoz establecido.

- La interacción entre un decodificador de audio y la posterior conversión de formato en un formato de reproducción deseado puede distinguirse en tres categorías:

- 30 1. El proceso de decodificación es agnóstico con respecto al escenario de reproducción final. Por lo tanto, la representación de audio completa se recupera, y se aplica posteriormente el procesamiento de conversión.

- 35 2. El proceso de decodificación de audio es limitado en sus capacidades, y dará salida a un formato fijo solamente. Ejemplos son radios mono que reciben programas de FM estéreo, o un decodificador HE-AAC mono que recibe una corriente de bits HE-AAC v2.

- 40 3. El proceso de decodificación de audio conoce el contexto de reproducción final y adapta su procesamiento consecuentemente. Un ejemplo es la "Decodificación de canales escalable para reducidas configuraciones de altavoz", tal como se define para MPEG Surround en la ref. [2]. En este caso, el decodificador reduce el número de canales de salida.

- Las desventajas de estos métodos presentan una innecesaria alta complejidad y las posibles fallas por el posterior procesamiento de material decodificado (filtración peine para mezcla descendente, desenmascarado para mezcla ascendente) (1.), y la limitada flexibilidad en relación con el formato de salida final (2. y 3.).

- 45 El objeto de la presente invención es proporcionar conceptos mejorados para el procesamiento de señales de audio. El objeto de la presente invención es resuelto por un decodificador según la reivindicación 1, por un método según la reivindicación 15 y por un programa de ordenador según la reivindicación 16.

- 50 Se proporciona un dispositivo de decodificador de audio para la decodificación de una señal de audio de entrada comprimida que comprende por lo menos un decodificador de núcleo que tiene uno o más procesadores para la generación de una señal de salida del procesador basándose en una señal de entrada del procesador, donde una cantidad de canales de salida de la señal de salida del procesador es mayor que una cantidad de canales de entrada de la señal de entrada del procesador, donde cada uno de los uno o más procesadores comprende un descorrelacionador y una mezcladora, donde una señal de salida de decodificador de núcleo que tiene una pluralidad de canales comprende la señal de salida del procesador, y donde la señal de salida de decodificador de núcleo es adecuada para un contexto de altavoz de referencia;

- 60 Estando por lo menos un conversor de formato configurado para convertir la señal de salida del decodificador de núcleo en una señal de audio de salida, que es adecuada para un contexto de altavoz objetivo; y

un dispositivo de control configurado para el control de por lo menos uno o más procesadores de tal manera que el descorrelacionador del procesador puede ser controlado independientemente de la mezcladora del procesador, donde el dispositivo de control está configurado para el control de por lo menos uno de los descorrelacionadores de los uno o más procesadores de tal manera que dependa del contexto de altavoz objetivo, la mezcladora (40, 40') del

procesador (36, 36') está operativa cuando el descorrelacionador (39, 39') del procesador (36, 36') está apagado.

El propósito de los procesadores es crear una señal de salida del procesador que tiene una mayor cantidad de canales incoherentes/descorrelacionados que la cantidad de los canales de entrada de la señal de entrada del procesador. Más en particular, cada uno de los procesadores genera una señal de salida del procesador con una pluralidad de canales de salida incoherentes/no correlacionados, por ejemplo, con dos canales de salida, donde las correctas indicaciones espaciales de una señal de entrada del procesador tienen una menor cantidad de canales de entrada, por ejemplo, de una señal de entrada mono.

Dichos procesadores comprenden un descorrelacionador y una mezcladora. El descorrelacionador se usa para crear una señal de descorrelacionador de un canal de la señal de entrada del procesador. Típicamente, un descorrelacionador (filtro de descorrelación) consiste en un prerretardo dependiente de frecuencia seguido de secciones todo paso (IIR).

La señal del descorrelacionador y el canal respectivo de la señal de entrada del procesador entonces se alimentan a la mezcladora. La mezcladora está configurada para establecer una señal de salida del procesador mediante la mezcla de la señal del descorrelacionador y el canal respectivo de la señal de entrada del procesador, donde dicha información se usa con el fin de sintetizar la correcta coherencia/correlación y la correcta relación de intensidad de los canales de salida de la señal de salida del procesador.

Los canales de salida de la señal de salida del procesador son entonces no coherentes/no correlacionados, de manera que los canales de salida del procesador se percibirán como fuentes de sonido independientes si se alimentan a diferentes altavoces en diferentes posiciones.

El conversor de formato puede convertir la señal de salida del decodificador de núcleo de modo de ser adecuada para la reproducción en un contexto de altavoz que puede diferir del contexto de altavoz de referencia. Este contexto se denomina contexto de altavoz objetivo.

En el caso de que los canales de salida de un procesador no sean necesarios para un contexto de altavoz objetivo específico por el posterior conversor de formato en una forma incoherente/no correlacionada, la síntesis de la correcta correlación se torna perceptualmente irrelevante. En consecuencia, para estos procesadores, el descorrelacionador puede omitirse. Sin embargo, en general, la mezcladora permanece completamente operativa cuando el descorrelacionador se apaga. Como resultado, los canales de salida de la señal de salida del procesador son generados incluso si el descorrelacionador se apaga.

Debe observarse que, en este caso, los canales de la señal de salida del procesador son coherentes/correlacionados, pero no idénticos. Esto significa que los canales de la señal de salida del procesador pueden ser adicionalmente procesados de manera independiente entre sí, aguas abajo del procesador, donde, por ejemplo, la relación de intensidad y/u otra información espacial podría usarse por el conversor de formato con el fin de establecer los niveles de los canales de la señal de audio de salida.

Como la filtración de descorrelación requiere una sustancial complejidad computacional, la carga de trabajo de decodificación general puede ser reducida en gran medida por el dispositivo de decodificador propuesto.

Si bien los descorrelacionadores, en particular, sus filtros todo paso, están diseñados de manera que tienen un mínimo impacto sobre la calidad de sonido subjetiva, no siempre puede evitarse la introducción de fallas audibles, por ejemplo, embarrado de transientes debido a distorsiones de fases o "zumbido" de ciertos componentes de frecuencia.

Por lo tanto, puede lograrse una mejora de la calidad de sonido del audio, ya que los efectos secundarios del proceso de descorrelacionador se omiten.

Obsérvese que este procesamiento solo será aplicado para bandas de frecuencia donde se aplica la descorrelación. Las bandas de frecuencia donde se usa la codificación residual no se ven afectadas.

En realizaciones preferidas, el dispositivo de control está configurado para desactivar por lo menos uno o más procesadores de modo tal que los canales de entrada de la señal de entrada del procesador se alimentan a canales de salida de la señal de salida del procesador de una forma no procesada. Mediante esta característica, la cantidad de canales que no son idénticos puede reducirse. Esto podría ser ventajoso, si el contexto de altavoz objetivo comprende varios altavoces, que es muy pequeño en comparación con la cantidad de altavoces del contexto de altavoz de referencia.

En realizaciones ventajosas, el procesador es una herramienta de decodificación de una entrada y dos salidas (OTT, según sus siglas en inglés), donde el descorrelacionador está configurado para crear una señal descorrelacionada

mediante la descorrelación de por lo menos un canal de la señal de entrada del procesador, donde la mezcladora mezcla la señal de audio de entrada del procesador y la señal descorrelacionada basándose en una señal de diferencia de nivel de canal (CLD, según sus siglas en inglés) y/o una señal de coherencia entre canales (ICC, según sus siglas en inglés), de tal manera que una señal de salida del procesador consiste en dos canales de salida incoherentes. Dichas herramientas de decodificación de una entrada a salida permiten la creación de una señal de salida del procesador con un par de canales, que tienen la amplitud y coherencia correctas entre sí de una manera sencilla.

En algunas realizaciones, el dispositivo de control está configurado para apagar el descorrelacionador de uno de los procesadores, estableciendo la señal de audio descorrelacionada en cero, o evitando que la mezcladora mezcle la señal descorrelacionada en una señal de salida del procesador del procesador respectivo. Ambos métodos permiten el apagado del descorrelacionador de una manera fácil.

En realizaciones preferidas, el decodificador de núcleo es un decodificador tanto para música como para voz, tal como un decodificador USAC, donde una señal de entrada del procesador de por lo menos uno de los procesadores contiene elementos de pares de canales, por ejemplo, elementos de pares de canales USAC. En este caso, es posible omitir la decodificación de los elementos de pares de canales, si esto no es necesario para el contexto de altavoz objetivo actual. De esta manera, la complejidad computacional y las fallas que se originan a partir del proceso de descorrelación, al igual que del proceso de mezcla descendente, pueden reducirse significativamente.

En algunas realizaciones, el decodificador de núcleo es un codificador de objeto paramétrico, tal como un decodificador SAOC. De este modo, la complejidad computacional y las fallas que se originan a partir del proceso de descorrelación, al igual que del proceso de mezcla descendente, pueden reducirse adicionalmente.

En algunas realizaciones, la cantidad de altavoces de un contexto de altavoz de referencia es mayor que una cantidad de altavoces del contexto de altavoz objetivo. En este caso, el conversor de formato puede mezclar de forma descendente la señal de salida del decodificador de núcleo a un audio con la señal de audio de salida, donde la cantidad de los canales de salida es menor que la cantidad de canales de salida de la señal de salida del decodificador de núcleo.

En este caso, la mezcla descendente describe el caso donde se presenta una cantidad mayor de altavoces en el contexto de altavoz de referencia, que la utilizada en el contexto de altavoz objetivo. En dichos casos, los canales de salida de uno o más procesadores a menudo no son necesarios en la forma de señales incoherentes. Si los descorrelacionadores de dichos procesadores se apagan, la complejidad computacional y las fallas que se originan a partir del proceso de descorrelación, al igual que del proceso de mezcla descendente, pueden reducirse significativamente.

En algunas realizaciones, el dispositivo de control está configurado para apagar los descorrelacionadores para por lo menos uno primero de dichos canales de salida de la señal de salida del procesador, y uno segundo de dichos canales de salida de la señal de salida del procesador, si el primero de dichos canales de salida y el segundo de dichos canales de salida son, según el contexto de altavoz objetivo, mezclados en un canal común de la señal de audio de salida, siempre y cuando un primer factor de escala para la mezcla del primero de dichos canales de salida de la señal de salida del procesador en el canal común exceda un primer umbral, y/o un segundo factor de escala para la mezcla del segundo de dichos canales de salida de la señal de salida del procesador en el canal común exceda un segundo umbral.

En el caso de que el primero de dichos canales de salida y el segundo de dichos canales de salida estén mezclados en un canal común de la señal de audio de salida, la descorrelación en el decodificador de núcleo puede omitirse para los canales de salida primero y segundo. De esta manera, la complejidad computacional y las fallas que se originan a partir del proceso de descorrelación, al igual que del proceso de mezcla descendente, pueden reducirse significativamente. De este modo, puede evitarse una descorrelación innecesaria.

En una realización más avanzada, puede preverse el primer factor de escala para la mezcla del primero de dichos canales de salida de la señal de salida del procesador. De la misma manera, puede usarse un segundo factor de escala para la mezcla del segundo de dichos canales de salida de señal de salida del procesador. En este caso, un factor de escala es un valor numérico, habitualmente, entre cero y uno, que describe la relación entre la intensidad de señal en el canal original (canal de salida de la señal de salida del procesador) y la intensidad de señal de la señal resultante en el canal mezclado (canal común de la señal de audio de salida). Los factores de escala pueden estar contenidos en una matriz de mezcla descendente. Mediante el uso de un primer umbral para el primer factor de escala, y/o el uso de un segundo umbral para el segundo factor de escala, puede garantizarse que la descorrelación para el primer canal de salida y el segundo canal de salida solo se apague, si por lo menos se mezclan una parte determinada del primer canal de salida y/o por lo menos una parte determinada del segundo canal de salida en el canal común. A modo de ejemplo, el umbral puede establecerse en cero.

En realizaciones preferidas, el dispositivo de control está configurado para recibir un conjunto de pautas desde el conversor de formato según el cual el conversor de formato mezcla los canales de la señal de salida del procesador con los canales de la señal de audio de salida según el contexto de altavoz objetivo, donde el dispositivo de control está configurado para controlar los procesadores según el conjunto recibido de pautas. En este caso, el control de los procesadores puede incluir el control de los descorrelacionadores y/o de las mezcladoras. Mediante esta característica, puede garantizarse que el dispositivo de control controle los procesadores de manera precisa.

Por medio del conjunto de pautas, puede proporcionarse información en términos de si los canales de salida de un procesador son combinados por una etapa posterior de conversión de formato al dispositivo de control. Las pautas recibidas por el dispositivo de control típicamente se presentan en la forma de una matriz de mezcla descendente que define factores de escala para cada canal de salida del decodificador, a cada canal de salida de audio utilizado por el conversor de formato. En una siguiente etapa, pueden calcularse las pautas de control para el control de los descorrelacionadores por el dispositivo de control a partir de las pautas de mezcla descendente. Estas pautas de control pueden estar contenidas en una llamada matriz de mezcla, que puede ser generada por el dispositivo de control según el contexto de altavoz objetivo. Estas pautas de control pueden usarse entonces para el control de los descorrelacionadores y/o las mezcladoras. Como resultado, el dispositivo de control puede adaptarse a diferentes contextos de altavoz objetivos sin intervención manual.

En realizaciones preferidas, el dispositivo de control está configurado para controlar los descorrelacionadores del decodificador de núcleo de modo tal que una cantidad de canales incoherentes de la señal de salida del decodificador de núcleo es igual a la cantidad de altavoces del contexto de altavoz objetivo. En este caso, la complejidad computacional y las fallas que se originan a partir del proceso de descorrelación, al igual que del proceso de mezcla descendente, pueden reducirse significativamente.

En realizaciones, el conversor de formato comprende un mezclador descendente para la mezcla descendente de la señal de salida del decodificador de núcleo. El mezclador descendente puede producir directamente la señal de audio de salida. Sin embargo, en algunas realizaciones, el mezclador descendente puede conectarse a otro elemento del conversor de formato, que entonces produce la señal de audio de salida.

En algunas realizaciones, el conversor de formato comprende un renderizador binaural. Los renderizadores binaurales se usan generalmente para convertir una señal de multicanal en una señal estéreo adaptada para el uso con auriculares estéreo. El renderizador binaural produce una mezcla descendente binaural de la señal alimentada, de modo tal que cada canal de esta señal es representado por una fuente de sonido virtual. El procesamiento puede llevarse a cabo a modo de trama en un dominio de filtro espejo en cuadratura (QMF, según sus siglas en inglés). La binauralización se realiza basándose en respuestas de impulsos ambientales binaurales medidos, y causa una complejidad computacional extremadamente alta, que se correlaciona con la cantidad de canales incoherentes/descorrelacionados de la señal alimentada al renderizador binaural.

En realizaciones preferidas, la señal de salida del decodificador de núcleo se alimenta al renderizador binaural como una señal de entrada de renderizador binaural. En este caso, el dispositivo de control habitualmente está configurado para controlar los procesadores del decodificador de núcleo de modo tal que una cantidad de los canales de la señal de salida del decodificador de núcleo es mayor que la cantidad de altavoces de los auriculares. Esto puede ser deseable, ya que, por ejemplo, el renderizador binaural puede utilizar la información de sonido espacial contenida en los canales para el ajuste de las características de frecuencia de la señal estéreo alimentada a los auriculares con el fin de generar una impresión de audio tridimensional.

En algunas realizaciones, la señal de salida del mezclador descendente se alimenta al renderizador binaural como una señal de entrada de renderizador binaural. En el caso de que la señal de audio de salida del mezclador descendente sea alimentada al renderizador binaural, la cantidad de canales de esta señal de entrada es significativamente menor que en los casos en los cuales la señal de salida del decodificador de núcleo se alimenta al renderizador binaural, de modo que se reduce la complejidad computacional.

Adicionalmente, se proporciona un método para la decodificación de una señal de audio de entrada comprimida, donde el método comprende las etapas de: proporcionar por lo menos un decodificador de núcleo que tiene uno o más procesadores para la generación de una señal de salida del procesador basándose en una señal de entrada del procesador, donde una cantidad de canales de salida de la señal de salida del procesador es mayor que una cantidad de canales de entrada de la señal de entrada del procesador, donde cada uno de los uno o más procesadores comprende un descorrelacionador y una mezcladora, donde una señal de salida del decodificador de núcleo que tiene una pluralidad de canales comprende una señal de salida del procesador, y donde la señal de salida del decodificador de núcleo es adecuada para un contexto de altavoz de referencia; proporcionar por lo menos un conversor de formato configurado para convertir la señal de salida del decodificador de núcleo en una señal de audio de salida, que es adecuada para un contexto de altavoz objetivo; y proporcionar un dispositivo de control configurado para el control de por lo menos uno o más procesadores de tal manera que el descorrelacionador del procesador puede ser controlado independientemente de la mezcladora del procesador,

donde el dispositivo de control está configurado para el control de por lo menos uno de los descorrelacionadores de los uno o más procesadores de tal manera que dependiendo de la configuración de altavoz objetivo, la mezcladora (40, 40') del procesador (36, 36') esté operativa cuando el descorrelacionador (39, 39') del procesador (36, 36') esté apagado.

5 Aún más, se proporciona un programa de ordenador para la implementación del método mencionado con anterioridad, cuando se ejecuta en un ordenador o en un procesador de señales.

10 A continuación, se describen las realizaciones de la presente invención en más detalle, con referencia a las figuras, en las que:

La figura 1 muestra un diagrama de bloques de una realización preferida de un decodificador según la invención;

15 la figura 2 muestra un diagrama de bloques de una segunda realización de un decodificador según la invención;

la figura 3 muestra un modelo de un procesador conceptual, donde el descorrelacionador está encendido;

la figura 4 muestra un modelo de un procesador conceptual, donde el descorrelacionador está apagado;

20 la figura 5 ilustra una interacción entre la conversión de formato y la decodificación;

la figura 6 muestra un diagrama de bloques de un detalle de una realización de un decodificador según la invención, donde se genera una señal de 5.1 canales;

25 la figura 7 muestra un diagrama de bloques de un detalle de la realización de la figura 6 de un decodificador según la invención, donde el canal 5.1 es mezclado descendentemente hasta una señal de canal 2.0;

la figura 8 muestra un diagrama de bloques de un detalle de la realización de la figura 6 de un decodificador según la invención, donde la señal de canal 5.1 es mezclada descendentemente hasta una señal de canal 4.0;

30 la figura 9 muestra un diagrama de bloques de un detalle de una realización de un decodificador según la invención, donde se genera una señal de canal 9.1;

35 la figura 10 muestra un diagrama de bloques de un detalle de la realización de la figura 9 de un decodificador según la invención, donde la señal de canal 9.1 es mezclada descendentemente hasta una señal de canal 4.0;

la figura 11 muestra un diagrama de bloques esquemático de una reseña conceptual de un codificador de audio 3D;

40 la figura 12 muestra un diagrama de bloques esquemático de una reseña conceptual de un decodificador de audio 3D; y

la figura 13 muestra un diagrama de bloques esquemático de una reseña conceptual de un conversor de formato.

45 Antes de describir las realizaciones de la presente invención, se proporcionan más antecedentes sobre sistemas de codificadores-decodificadores de la técnica anterior.

La figura 11 muestra un diagrama de bloques esquemático de una reseña conceptual de un codificador de audio 3D 1, mientras que la figura 12 muestra un diagrama de bloques esquemático de una reseña conceptual de un decodificador de audio 3D 2.

50 El sistema de códec de audio 3D 1, 2, puede basarse en un codificador MPEG-D de codificación unificada de voz y audio (USAC, según sus siglas en inglés) 3, para la codificación de señales de canales 4 y señales de objetos 5, al igual que basándose en un decodificador MPEG-D de codificación unificada de voz y audio (USAC) 6, para la decodificación de la señal de audio de salida 7 del codificador 3. Con el fin de incrementar la eficiencia para la

55 codificación de una gran cantidad de objetos 5, se ha adaptado la tecnología de codificación de objeto de audio espacial (SAOC, según sus siglas en inglés). Tres tipos de renderizadores 8, 9, 10 realizan las tareas de renderización de los objetos 11, 12 a los canales 13, renderizar los canales 13 a auriculares, o renderizar los canales a un contexto de altavoz diferente.

60 Cuando las señales de objeto son explícitamente transmitidas o paramétricamente codificadas usando SAOC, la correspondiente información de metadatos de objetos (OAM) 14 es comprimida y multiplexada en la corriente de bits de audio 3D 7.

El prerrrenderizador/mezcladora 15 puede usarse opcionalmente para convertir una escena de entrada de canal y

objeto 4, 5, en una escena de canal 4, 16, antes de la codificación. Funcionalmente, es idéntica al renderizador de objeto/mezcladora 15 que se describe a continuación.

La prerrenderización de objetos 5 garantiza la entropía de señal determinística en la entrada del codificador 3 que es básicamente independiente de la cantidad de señales de objeto activas simultáneamente 5. Con la prerrenderización de objetos 5, no se requiere la transmisión de metadatos de objeto 14.

Las señales de objetos discretas 5 son renderizadas a la disposición de canal que el codificador 3 está configurado para usar. Los pesos de los objetos 5 para cada canal 16 se obtienen a partir de los metadatos de objetos asociados 14.

El códec de núcleo para señales de canal de altavoz 4, señales de objetos discretas 5, señales de mezcla descendente de objetos 14 y señales prerrenderizadas 16 pueden basarse en la tecnología MPEG-D USAC. Esta manipula la codificación de la multitud de señales 4, 5, 14 mediante la creación de la información de mapeo de canal y objeto basándose en la información geométrica y semántica del canal de entrada y asignación de objeto. Esta información de mapeo describe la manera en que los canales de entrada 4 y los objetos 5 son mapeados hasta elementos de canales USAC, concretamente, hasta elementos de pares de canales (CPE, según sus siglas en inglés), elementos de canales únicos (SCE, según sus siglas en inglés), mejoras de baja frecuencia (LFE, según sus siglas en inglés), y la información correspondiente es transmitida al decodificador 6.

Todas las cargas de transmisión de datos adicionales, como datos SAOC 17 o metadatos de objetos 14 pueden pasarse a través de elementos de extensión y pueden considerarse en el control de tasa del codificador 3.

La codificación de objetos 5 es posible de diferentes maneras, según los requerimientos de tasa/distorsión y los requerimientos de interactividad para el renderizador. Son posibles las siguientes variantes de codificación de objetos:

- Objetos prerrenderizados 16: las señales de objetos 5 son prerrenderizadas y mezcladas con las señales de canales 4, por ejemplo con señales de canales 22.2 4, antes de la codificación. La posterior cadena de codificación observa señales de canales 22.2 4.

- Formas de onda de objetos discretas: los objetos 5 son suministrados como formas de onda monofónicas al codificador 3. El codificador 3 utiliza elementos de canales (SCE) independientes para transmitir los objetos 5 además de las señales de canal 4. Los objetos decodificados 18 son renderizados y mezclados del lado del receptor. Se transmite información de metadatos de objetos comprimidos 19, 20 al receptor/renderizador 21 junto con lo anterior.

- Formas de onda de objetos paramétricas 17: las propiedades de objeto y su relación entre sí se describen por medio de parámetros SAOC 22, 23. La mezcla descendente de las señales de objetos 17 es codificada con USAC. La información paramétrica 22 es transmitida junto con lo anterior. La cantidad de canales de mezcla descendente 17 se selecciona según la cantidad de objetos 5 y la tasa de datos general. La información de metadatos de objetos comprimidos 23 es transmitida al renderizador SAOC 24.

El codificador SAOC 25 y el decodificador 24 para señales de objetos 5 se basan en la tecnología MPEG SAOC. El sistema es capaz de recrear, modificar y renderizar una cantidad de objetos de audio 5 basándose en una cantidad menor de canales transmitidos 7 y datos paramétricos 22, 23 adicionales, tales como las diferencias de niveles de objetos (OLD, según sus siglas en inglés), las correlaciones entre objetos (IOC, según sus siglas en inglés) y los valores de ganancia de mezcla descendente (DMG, según sus siglas en inglés). Los datos paramétricos 22, 23 adicionales exhiben una tasa de datos significativamente menor que la requerida para la transmisión de todos los objetos 5 individualmente, lo que hace que la codificación sea muy eficaz.

El codificador SAOC 25 toma como entrada las señales de objeto/canal 5 como formas de onda monofónicas y emite la información paramétrica 22 (que está empaquetada en la corriente de bits de audio 3D 7) y los canales de transporte SAOC 17 (que están codificados usando elementos de canales independientes y transmitidos). El decodificador SAOC 24 reconstruye las señales de objeto/canal 5 a partir de los canales de transporte SAOC decodificados 26 y la información paramétrica 23, y genera la escena de audio de salida 27 basándose en la disposición de reproducción, la información de metadatos de objetos descomprimidos 20 y, opcionalmente, sobre la información de interacción de usuario.

Para cada objeto 5, los metadatos de objetos asociados 14 que especifican la posición geométrica y el volumen del objeto en espacio 3D son eficazmente codificados por un codificador de metadatos de objetos 28 mediante la cuantificación de las propiedades de objetos en tiempo y espacio. Los metadatos de objetos comprimidos (cOAM) 19 son transmitidos al receptor como información secundaria 20, que puede ser decodificada por un decodificador OAM 29.

- El renderizador de objeto 21 utiliza los metadatos de objetos comprimidos 20 para generar formas de onda de objetos 12 según el formado de reproducción determinado. Cada objeto 5 es renderizado a ciertos canales de salida 12 según sus metadatos 19, 20. La salida de este bloque 21 resulta de la suma de los resultados parciales. Si tanto el contenido basado en el canal 11, 30, como los objetos discretos/paramétricos 12, 27 son decodificados, las formas de onda basadas en el canal 11, 30, y las formas de ondas de objetos renderizados 12, 27 se mezclan antes de emitir las formas de onda 13 resultantes (o antes de alimentarlas a un módulo posprocesador 9, 10 como el renderizador binaural 9 o el módulo renderizador de altavoz 10) por una mezcladora 8.
- El módulo de renderizador binaural 9 produce una mezcla descendente binaural del material de audio de multicanal 13, de modo tal que cada canal de entrada 13 sea representado por una fuente de sonido virtual. El procesamiento se lleva a cabo a modo de trama en un dominio de filtro espejo en cuadratura (QMF). La binauralización se basa en las respuestas de impulsos ambientales binaurales medidos.
- El renderizador de altavoz 10 mostrado en la figura 13 en más detalle convierte entre la configuración de canal transmitida 13 y el formato de reproducción deseado 31. Se denomina entonces 'convertor de formato' 10 a continuación. El convertor de formato 10 realiza conversiones a números menores de canales de salida 31, es decir, crea mezclas descendentes por medio de un mezclador descendente 32. El configurador de DMX (mezcla descendente) 33 genera automáticamente matrices optimizadas de mezcla descendente para la combinación dada de formatos de entrada 13 y formatos de salida 31, y aplica estas matrices en un proceso de mezcla descendente 32, donde se usan una disposición de salida de mezcladora 34 y una disposición de reproducción 35. El convertor de formato 10 permite las configuraciones de altavoz habituales al igual que las configuraciones aleatorias con posiciones de altavoz no habituales.
- La figura 1 muestra un diagrama de bloques de una realización preferida de un decodificador 2 según la invención.
- El dispositivo decodificador de audio 2 para la decodificación de una señal de audio de entrada comprimida 38, 38' comprende por lo menos un decodificador de núcleo 6 que tiene uno o más procesadores 36, 36' para la generación de una señal de salida del procesador 37, 37' basándose en una señal de entrada del procesador 38, 38', donde una cantidad de canales de salida 37.1, 37.2, 37.1', 37.2' de la señal de salida del procesador 37, 37' es mayor que una cantidad de canales de entrada 38.1, 38.1' de la señal de entrada del procesador 38, 38', donde cada uno de los uno o más procesadores 36, 36' comprende un descorrelacionador 39, 39' y una mezcladora 40, 40', donde una señal de salida de decodificador de núcleo 13 que tiene una pluralidad de canales 13.1, 13.2, 13.3, 13.4 comprende una señal de salida del procesador 37, 37', y donde la señal de salida del decodificador de núcleo 13 es adecuada para un contexto de altavoz de referencia 42.
- Adicionalmente, el dispositivo decodificador de audio 2 comprende por lo menos un dispositivo convertor de formato 9, 10 configurado para convertir la señal de salida del decodificador de núcleo 13 en una señal de audio de salida 31, adecuada para un contexto de altavoz objetivo 45.
- Además, el dispositivo decodificador de audio 2 comprende un dispositivo de control 46 configurado para el control de por lo menos uno o más procesadores 36, 36' de modo tal que el descorrelacionador 39, 39' del procesador 36, 36' puede ser controlado independientemente de la mezcladora 40, 40' del procesador 36, 36', donde el dispositivo de control 46 está configurado para el control de por lo menos uno de los descorrelacionadores 39, 39' de los uno o más procesadores 36, 36' de tal manera que, dependiendo del contexto de altavoz objetivo 45, el mezclador 40, 40' del procesador 36, 36' esté operativo cuando el descorrelacionador 39, 39' del procesador 36, 36' esté apagado.
- El propósito de los procesadores 36, 36' es crear una señal de salida del procesador 37, 37' que tiene una mayor cantidad de canales incoherentes/descorrelacionados 37.1, 37.2, 37.1', 37.2' que la cantidad de los canales de entrada 38.1, 38.1' de la señal de entrada del procesador 38. Más en particular, cada uno de los procesadores 36, 36' puede generar una señal de salida del procesador 37 con una pluralidad de canales de salida incoherentes/no correlacionados 37.1, 37.2, 37.1', 37.2' con las correctas indicaciones espaciales de una señal de entrada del procesador 38, 38' que tiene una menor cantidad de canales de entrada 38.1, 38.1'.
- En la realización que se muestra en la figura 1, un primer procesador 36 tiene dos canales de salida 37.1, 37.2, que son generados a partir de una señal de entrada mono 38, y un segundo procesador 36' tiene dos canales de salida 37.1', 37.2', que son generados a partir de una señal de entrada mono 38'.
- El dispositivo convertor de formato 9, 10 puede convertir la señal de salida del decodificador de núcleo 13 de modo que sea adecuada para la reproducción en un contexto de altavoz 45 que puede diferir del contexto de altavoz de referencia 42. Este contexto se denomina contexto de altavoz objetivo 45.
- En la realización de la figura 1, el contexto de altavoz de referencia 42 comprende un altavoz frontal izquierdo (L), un altavoz frontal derecho (R), un altavoz envolvente izquierdo (LS) y un altavoz envolvente derecho (RS).

Adicionalmente, el contexto de altavoz objetivo 42 comprende un altavoz frontal izquierdo (L), un altavoz frontal derecho (R) y un altavoz envolvente central (CS).

En el caso de que los canales de salida 37.1, 37.2, 37.1', 37.2' de un procesador 36, 36' no sean necesarios para un contexto de altavoz objetivo específico 45 por el dispositivo de conversor de formato 9, 10 posterior en una forma incoherente/no correlacionada, la síntesis de la correcta correlación se torna perceptualmente irrelevante. En consecuencia, para estos procesadores 36, 36', el descorrelacionador 39, 39' puede omitirse. Sin embargo, en general, la mezcladora 40, 40' permanece completamente operativa cuando el descorrelacionador se apaga. Como consecuencia, los canales de salida 37.1, 37.2, 37.1', 37.2' de la señal de salida del procesador se generan incluso si el descorrelacionador 39, 39' está apagado.

Debe observarse que, en este caso, los canales 37.1, 37.2, 37.1', 37.2' de la señal de salida del procesador 37, 37' son coherentes/correlacionados, pero no idénticos. Esto significa que los canales 37.1, 37.2, 37.1', 37.2' de la señal de salida del procesador 37, 37' pueden procesarse adicionalmente de manera independiente entre sí, aguas abajo del procesador 36, 36', donde, por ejemplo, la relación de intensidad y/u otra información espacial podría usarse por el dispositivo conversor de formato 9, 10 con el fin de establecer los niveles de los canales 31.1, 31.2, 31.3 de la señal de audio de salida 31.

Como la filtración de descorrelación requiere una complejidad computacional sustancial, la carga de trabajo de decodificación general puede reducirse en gran medida por el dispositivo de decodificador 2 propuesto.

Si bien los descorrelacionadores 39, 39', en particular, sus filtros paso todo, están diseñados de manera que tienen un impacto mínimo sobre la calidad de sonido subjetiva, no siempre puede evitarse la introducción de fallas audibles, por ejemplo, el embarrado de transientes debido a distorsiones de fases o "zumbido" de ciertas componentes de frecuencia. Por lo tanto, puede lograrse una mejora de la calidad de sonido del audio, ya que los efectos secundarios del proceso de descorrelacionador se omiten.

Obsérvese que este procesamiento solo se aplicará a bandas de frecuencia donde se aplica la descorrelación. Las bandas de frecuencia donde se usa la codificación residual no se ven afectadas.

En realizaciones preferidas, el dispositivo de control 46 está configurado para desactivar por lo menos uno o más procesadores 36, 36', de modo tal que los canales de entrada 38.1, 38.1' de la señal de entrada del procesador se alimentan a canales de salida 37.1, 37.2, 37.1', 37.2' de la señal de salida del procesador 37, 37' de una forma no procesada. Mediante esta característica, puede reducirse la cantidad de canales que no son idénticos. Esto podría ser ventajoso, si el contexto de altavoz objetivo 45 comprende una cantidad de altavoces, que es muy pequeña en comparación con la cantidad de altavoces del contexto de altavoz de referencia 42.

En realizaciones preferidas, el decodificador de núcleo 6 es un decodificador 6 tanto para música como para voz, tal como un decodificador USAC 6, donde una señal de entrada del procesador 38, 38' de por lo menos uno de los procesadores contiene elementos de pares de canales, tales como elementos USAC de pares de canales. En este caso, es posible omitir la decodificación de los elementos de pares de canales, si esto no es necesario para el contexto actual de altavoz objetivo 45. De esta manera, la complejidad computacional y las fallas que se originan a partir del proceso de descorrelación al igual que del proceso de mezcla descendente pueden reducirse significativamente.

En algunas realizaciones, el decodificador de núcleo es un codificador de objeto paramétrico 24, tal como un decodificador SAOC 24. De esta manera, la complejidad computacional y las fallas que se originan a partir del proceso de descorrelación al igual que del proceso de mezcla descendente pueden reducirse adicionalmente.

En algunas realizaciones, la cantidad de altavoces de un contexto de altavoz de referencia 42 es mayor que una cantidad de altavoces del contexto de altavoz objetivo 45. En este caso, el dispositivo conversor de formato 9, 10 puede mezclar de manera descendente la señal de salida del decodificador de núcleo 13 con la señal de audio de salida 31, donde la cantidad de los canales de salida 31.1, 31.2, 31.3 es menor que la cantidad de canales de salida 13.1, 13.2, 13.3, 13.4 de la señal de salida del decodificador de núcleo 13.

En este caso, la mezcla descendente describe el caso en que se presenta una cantidad mayor de altavoces en el contexto de altavoz de referencia 42, que la utilizada en el contexto de altavoz objetivo 45. En dichos casos, los canales de salida 37.1, 37.2, 37.1', 37.2' de uno o más procesadores 36, 36' con frecuencia no son necesarios en la forma de señales incoherentes. En la figura 1, se presentan cuatro canales de salida del decodificador 13.1, 13.2, 13.3, 13.4 de la señal de salida del decodificador de núcleo 13, si bien solo tres canales de salida 31.1, 31.2, 31.3 de la señal de salida de audio 31. Si los descorrelacionadores 39, 39' de dichos procesadores 36, 36' están apagados, la complejidad computacional y las fallas que se originan a partir del proceso de descorrelación al igual que del proceso de mezcla descendente pueden reducirse significativamente.

Por las razones que se explican a continuación, los canales de salida del decodificador 13.3 y 13.4 en la figura 1 no son necesarios en la forma de señales incoherentes. Por lo tanto, el descorrelacionador 39' se apaga por el dispositivo de control 46, mientras que el descorrelacionador 39 y la mezcladoras 40, 40' se encienden.

- 5 En algunas realizaciones, el dispositivo de control 46 está configurado para apagar los descorrelacionadores 39' para por lo menos uno primero de dichos canales de salida 37.1' de la señal de salida del procesador 37, 37' y uno segundo de dichos canales de salida 37.2, 37.2' de la señal de salida del procesador 37, 37', si el primero de dichos canales de salida 37.1' y el segundo de dichos canales de salida 37.2', según el contexto de altavoz objetivo 45, se mezclan en un canal común 31.3 de la señal de audio de salida 31, siempre y cuando un primer factor de escala para la mezcla del primero de dichos canales de salida 37.1' de la señal de salida del procesador 37' en el canal común 31.3 exceda un primer umbral, y/o un segundo factor de escala para la mezcla del segundo de dichos canales de salida 37.2' de la señal de salida del procesador 37' en el canal común 31.3 exceda un segundo umbral.

- 15 En la figura 1, los canales de salida del decodificador 13.3 y 13.4 están mezclados en un canal común 31.3 de la señal de audio de salida 31. Los factores de escala primero y segundo pueden ser 0.7071. Como los umbrales primero y segundo en esta realización están establecidos en cero, su descorrelacionador 39' se apaga.

- 20 En el caso de que el primero de dichos canales de salida 37.1' y el segundo de dichos canales de salida 37.2' sean mezclados en un canal común 31.3 de la señal de audio de salida 31, la descorrelación en el decodificador de núcleo 6 puede omitirse para los canales de salida primero y segundo 37.1', 37.2'. De esta manera, la complejidad computacional y las fallas que se originan a partir del proceso de descorrelación al igual que del proceso de mezcla descendente pueden reducirse significativamente. De este modo, puede evitarse una descorrelación innecesaria.

- 25 En una realización más avanzada, puede observarse el primer factor de escala para la mezcla del primero de dichos canales de salida 37.1' de la señal de salida del procesador 37'. De la misma manera, puede usarse un segundo factor de escala para la mezcla del segundo de dichos canales de salida 37.2' de la señal de salida del procesador 37'. En el presente documento, un factor de escala es un valor numérico, habitualmente, entre cero y uno, que describe la relación entre la intensidad de señal en el canal original (canal de salida 37.1', 37.2' de la señal de salida del procesador 37') y la intensidad de señal de la señal resultante en el canal mezclado (canal común 31.1 de la señal de audio de salida 31). Los factores de escala pueden estar contenidos en una matriz de mezcla descendente. Usando un primer umbral para el primer factor de escala y/o usando un segundo umbral para el segundo factor de escala, puede garantizarse que la descorrelación para el primer canal de salida 37.1' y el segundo canal de salida 37.2' solo se apague si por lo menos una parte determinada del primer canal de salida 37.1' y/o por lo menos una parte determinada del segundo canal de salida 37.2' se mezclan en el canal común 31.3. A modo de ejemplo, los umbrales pueden establecerse en cero.

- 40 En la realización de la figura 1, los canales de salida del decodificador 13.3 y 13.4 están mezclados en un canal común 31.3 de la señal de audio de salida 31. Los factores de escala primero y segundo pueden ser 0.7071. Como los umbrales primero y segundo en esta realización se establecen en cero, su descorrelacionador 39' está apagado.

- 45 En realizaciones preferidas, el dispositivo de control 46 está configurado para recibir un conjunto de pautas 47 desde el dispositivo conversor de formato 9, 10 según las cuales el dispositivo conversor de formato 9, 10 mezcla los canales 37.1, 37.2, 37.1', 37.2' de la señal de salida del procesador 37, 37' con los canales 31.1, 31.2, 31.3 de la señal de audio de salida 31 según el contexto de altavoz objetivo 45, donde el dispositivo de control 46 está configurado para controlar los procesadores 36, 36' según el conjunto de pautas 47 recibidas recibido. En el presente documento, el control de los procesadores 36, 36' puede incluir el control de los descorrelacionadores 39, 39' y/o de las mezcladoras 40, 40'. Mediante esta característica, puede garantizarse que el dispositivo de control 46 controla los procesadores 36, 36' de manera precisa.

- 50 Mediante el conjunto de pautas 47, la información sobre si los canales de salida de un procesador 36, 36' son combinados por una posterior etapa de conversión de formato puede ser proporcionada al dispositivo de control 9, 10. Las pautas recibidas por el dispositivo de control 46 son típicamente en forma de una matriz de mezcla descendente que define factores de escala para cada canal de salida del decodificador de núcleo 13.1, 13.2, 13.3, 13.4 a cada canal de salida de audio 31.1, 31.2, 31.3 utilizado por el dispositivo conversor de formato 9, 10. En una siguiente etapa, las pautas de control para el control de los descorrelacionadores pueden calcularse por el dispositivo de control a partir de las pautas de mezcla descendente. Estas pautas de control pueden estar contenidas en una denominada matriz de mezcla, que puede generarse por el dispositivo de control 46 según el contexto de altavoz objetivo 45. Estas pautas de control pueden entonces usarse para el control de los descorrelacionadores 39, 39' y/o las mezcladoras 40, 40'. Como resultado, el dispositivo de control 46 puede adaptarse a diferentes contextos de altavoz objetivos 45 sin intervención manual.

En la figura 1, el conjunto de pautas 47 puede contener la información de que los canales de salida del decodificador 13.3 y 13.4 se mezclan en un canal común 31.3 de la señal de audio de salida 31. Esto puede efectuarse en la realización de la figura 1, ya que el altavoz envolvente izquierdo y el altavoz envolvente derecho del contexto de

altavoz de referencia 42 son reemplazados por un altavoz envolvente central en el contexto de altavoz objetivo 45.

En realizaciones preferidas, el dispositivo de control 46 está configurado para controlar los descorrelacionadores 39, 39' del decodificador de núcleo 6 de modo tal que una cantidad de canales incoherentes de la señal de salida del decodificador de núcleo 13 es igual a la cantidad de altavoces del contexto de altavoz objetivo 45. En este caso, la complejidad computacional y las fallas que se originan a partir del proceso de descorrelación al igual que del proceso de mezcla descendente pueden reducirse significativamente.

Por ejemplo, en la figura 1, existen tres canales incoherentes: el primero es el canal de salida del decodificador 13.1, el segundo es el canal de salida del decodificador 13.2, y el tercero es cada uno de los canales de salida del decodificador 13.3 y 13.4, ya que los canales de salida del decodificador 13.3 y 13.4 son coherentes debido a la omisión del descorrelacionador 39'.

En realizaciones, por ejemplo, en la realización de la figura 1, el dispositivo conversor de formato 9, 10 comprende una mezcladora descendente 10 para la mezcla descendente de la señal de salida del decodificador de núcleo 13. La mezcladora descendente 10 puede producir directamente la señal de audio de salida 31, como se muestra en la figura 1. Sin embargo, en algunas realizaciones, la mezcladora descendente 10 puede estar conectada a otro elemento del conversor de formato 10, tal como un renderizador binaural 9, que entonces produce la señal de audio de salida 31.

La figura 2 muestra un diagrama de bloques de una segunda realización de un decodificador según la invención. A continuación, solo se describirán las diferencias con respecto a la primera realización. En la figura 2, el conversor de formato 9, 10 comprende un renderizador binaural 9. Los renderizadores binaurales 9 se usan generalmente para convertir una señal de multicanal en una señal estéreo adaptada para usarse con auriculares estéreo. El renderizador binaural 9 produce una mezcla descendente binaural LB y RB de la señal de multicanal alimentada, de modo que cada canal de esta señal es representado por una fuente de sonido virtual. La señal de multicanal puede tener hasta 32 canales o más. Sin embargo, en la figura 2, se muestra una señal de cuatro canales, por razones de simplicidad. El procesamiento puede llevarse a cabo a modo de trama en un dominio de filtro espejo en cuadratura (QMF). La binauralización se realiza basándose en respuestas de impulsos ambientales binaurales medidas y causa una complejidad computacional extremadamente alta, que se correlaciona con la cantidad de canales incoherentes/descorrelacionados de la señal alimentada al renderizador binaural 9. Con el fin de reducir la complejidad computacional, por lo menos uno de los descorrelacionadores 39, 39' puede apagarse.

En la realización de la figura 2, la señal de salida del decodificador de núcleo 13 se alimenta al renderizador binaural 9 como una señal de entrada de renderizador binaural 13. En este caso, el dispositivo de control 46 habitualmente está configurado para controlar los procesadores del decodificador de núcleo 6 de tal manera que una cantidad de los canales 13.1, 13.2, 13.3, 13.4 de la señal de salida del decodificador de núcleo 13 es mayor que la cantidad de altavoces de los auriculares. Esto puede ser deseable, por ejemplo, dado que el renderizador binaural 9 puede utilizar la información de sonido espacial contenida en los canales para el ajuste de las características de frecuencia de la señal estéreo alimentada a los auriculares con el fin de generar una impresión de audio tridimensional.

En realizaciones no mostradas, una señal de salida de mezcladora descendente de la mezcladora descendente 10 se alimenta al renderizador binaural 9 como una señal de entrada de renderizador binaural. En el caso de que la señal de audio de salida de la mezcladora descendente 10 sea alimentada al renderizador binaural 9, la cantidad de canales de su señal de entrada es significativamente menor que en los casos donde la señal de salida del decodificador de núcleo 13 se alimenta al renderizador binaural 9, de modo que se reduce la complejidad computacional.

En realizaciones ventajosas, el procesador 36 es una herramienta de decodificación de una entrada y dos salidas (OTT) 36, como se muestra en la figura 3 y la figura 4.

Como se muestra en la figura 3, el descorrelacionador 39 está configurado para crear una señal descorrelacionada 48 mediante la descorrelación de por lo menos un canal 38.1 de la señal de entrada del procesador 38, donde la mezcladora 40 mezcla la señal de audio de entrada del procesador 48 y la señal descorrelacionada 48 basándose en una señal de diferencia de nivel de canal (CLD) 49 y/o una señal de coherencia entre canales (ICC) 50, de modo que una señal de salida del procesador 37 consiste en dos canales de salida incoherentes 37.1, 37.2.

Dicha herramienta de decodificación de una entrada a salida 36 permite la creación de una señal de salida del procesador 37 con un par de canales 37.1, 37.2 que tienen la amplitud y coherencia correctas entre sí de una forma fácil. Típicamente, un descorrelacionador (filtro de descorrelación) consiste en un prerretardo dependiente de la frecuencia seguido de secciones todo paso (IIR).

En algunas realizaciones, el dispositivo de control está configurado para apagar el descorrelacionador 39 de uno de los procesadores 36 mediante el establecimiento de la señal de audio descorrelacionada 48 en cero, evitando que la

mezcladora mezcle la señal descorrelacionada 48 con una señal de salida del procesador 37 del procesador 36 respectivo. Ambos métodos permiten el apagado del descorrelacionador 39 de manera fácil.

5 Algunas realizaciones pueden ser definidas para un decodificador de multicanal 2 basándose en la "codificación de audio y voz unificadas según la norma ISO/IEC IS 23003-3".

La codificación de multicanal USAC se compone de diferentes elementos de canales. Un ejemplo para canales de audio 5.1 se proporciona a continuación.

10 Ejemplo de carga de transmisión de datos de corriente de bits simple

	N.º de elementos	elemIdx	tipo elemento usac[elemIdx]
señal de salida de canal 5.1	4	1	ID_USAC_SCE
		2	ID_USAC_CPE
		3	ID_USAC_CPE
		4	ID_USAC_LFE

15 Cada elemento estéreo ID_USAC_CPE puede estar configurado para usar MPEG Surround para la mezcla ascendente de mono con estéreo por una OTT 36. Como se representa a continuación, cada elemento genera dos canales de salida 37.1, 37.2 con las indicaciones espaciales correctas mediante la mezcla de una señal de entrada mono con la salida de un descorrelacionador 39 que se alimenta con dicha señal de entrada mono [2] [3].

20 Un bloque de construcción importante es el descorrelacionador 39 que se usa para sintetizar la coherencia/correlación correcta de los canales de salida 37.1, 37.2. Típicamente, los filtros de descorrelación consisten en un prerretardo dependiente de frecuencia seguido de secciones de todo paso (IIR).

25 En el caso de que los canales de salida 37.1, 37.2 de un bloque de decodificación de OTT 36 se mezclen de manera descendente por una etapa de conversión de formato posterior, la síntesis de la correlación correcta se torna perceptualmente irrelevante. En consecuencia, para estos bloques de mezcla ascendente, el descorrelacionador 39 puede omitirse. Esto puede lograrse de la siguiente manera.

30 Puede establecerse una interacción entre la conversión de formato 9, 10 y la decodificación, como se muestra en la figura 5. La información puede generarse si los canales de salida de un bloque de decodificación de OTT 36 se mezclan de manera descendente por una etapa de conversión de formato 9, 10 posterior. Esta información está contenida en una denominada matriz de mezcla, que se genera por una calculadora de matriz 46 y se pasa al decodificador USAC 6. La información procesada por la calculadora de matriz es típicamente la matriz de mezcla descendente proporcionada por el módulo de conversión de formato 9, 10.

35 El bloque de procesamiento de conversión de formato 9, 10 convierte los datos de audio de modo que son adecuados para la reproducción en un contexto de altavoz 45, que puede diferir del contexto de altavoz de referencia 42. Este contexto se denomina contexto de altavoz objetivo 45.

40 La mezcla descendente describe el caso donde se usa una cantidad menor de altavoces que la presente en el contexto de altavoz de referencia 42, en el contexto de altavoz objetivo 45.

45 En la figura 6, se muestra un decodificador de núcleo 6, que proporciona una señal de salida de decodificador de núcleo que comprende los canales de salida 13.1 a 13.6 adecuados para un contexto de altavoz de referencia 5.1 42, que comprende un canal L de altavoz frontal izquierdo, un canal R de altavoz frontal derecho, un canal LS de altavoz envolvente izquierdo, un canal RS de altavoz envolvente derecho, un canal C de altavoz frontal central y un canal LFE de altavoz de mejora de baja frecuencia. Los canales de salida 13.1 y 13.2 se crean por el procesador 36 basándose en elementos de pares de canales (ID_USAC_CPE), que se alimentan al procesador 36, como los canales descorrelacionados 13.1 y 13.2, cuando el descorrelacionador 39 del procesador 36 se enciende.

50 El canal L de altavoz frontal izquierdo, el canal R del altavoz frontal derecho, el canal LS de altavoz envolvente izquierdo, el canal RS de altavoz envolvente derecho y el canal C de altavoz frontal central son canales principales, mientras que el canal LFE de altavoz de mejora de baja frecuencia es opcional.

55 De la misma manera, los canales de salida 13.3 y 13.4 se crean por el procesador 36' basándose en elementos de pares de canales (ID_USAC_CPE), que se alimentan al procesador 36' como canales descorrelacionados 13.3 y 13.4, cuando el descorrelacionador 39' del procesador 36' se enciende.

El canal de salida 13.5 es basándose en elementos de canales únicos (ID_USAC_SCE), mientras que el canal de

salida 13.6 es basándose en elementos de mejora de baja frecuencia ID_USAC_LFE.

En el caso de que se encuentren disponibles seis altavoces adecuados, la señal de salida del decodificador de núcleo 13 puede usarse para la reproducción sin ninguna mezcla descendente. Sin embargo, en el caso de que solo esté disponible un conjunto de altavoz estéreo, la señal de salida del decodificador de núcleo 13 puede mezclarse de manera descendente.

Típicamente, el procesamiento de mezcla descendente puede ser descrito por una matriz de mezcla descendente que define factores de escala para cada canal de fuente a cada canal objetivo. Por ejemplo, ITU BS775 define la siguiente matriz de mezcla descendente para la mezcla descendente de canales principales 5.1 a estéreo, que mapea los canales L, R, C, LS y RS a los canales estéreo L' y R'.

$$M_{DMX} = \begin{pmatrix} 1,0 & 0,0 & 0,7071 & 0,701 & 0,0 \\ 0,0 & 1,0 & 0,7071 & 0,0 & 0,7071 \end{pmatrix}$$

La matriz de mezcla descendente tiene la dimensión m×n donde n es la cantidad de canales fuente y m es la cantidad de canales de destino.

A partir de la matriz de mezcla descendente M_{DMX}, se deduce una denominada matriz de mezcla M_{Mix} en el bloque de procesamiento de calculadora de matriz, que describe qué canales fuente se están combinando. Tiene la dimensión n×n.

$$M_{Mix}(i,j) = \begin{cases} 1, & \text{si se combinan el canal y el canal mediante mezcla descendente} \\ 0 & \text{de otro modo} \end{cases}$$

Observar que M_{Mix} es una matriz simétrica.

Para el ejemplo anterior de mezcla descendente de canales 5 a estéreo, la matriz de mezcla M_{Mix} es la siguiente:

$$M_{Mix} = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Un método para la obtención de la matriz de mezcla es proporcionado por el siguiente pseudocódigo:

M_{Mix} = cero n x n matriz

para i = 1 a m

para j = 1 a n

establecer _j = 0

si M_{DMX}(i, j) >thr

establecer _j = 1

fin

para k = 1 a n

establecer _k = 0

si M_{DMX}(i, k) >thr

establecer _k = 1

fin

si establecer_j == 1 establecer_k == 1

$M_{\text{Mix}}(j, k) = 1$

fin

fin

fin

fin

A modo de ejemplo, el umbral thr puede establecerse a cero.

Cada bloque de decodificación de OTT produce dos canales de salida correspondientes al canal número i y j. Si la matriz de mezcla $M_{\text{Mix}}(i, j)$ equivale a uno, la descorrelación se apaga para este bloque de decodificación.

Con el fin de omitir el descorrelacionador 39, los elementos $q^{l,m}$ se establecen a cero. Alternativamente, la vía de descorrelación puede omitirse, como se representa a continuación.

Esto da como resultado que los elementos $H12_{OTT}^{l,m}$ y $H22_{OTT}^{l,m}$ de la matriz de mezcla ascendente $R_2^{l,m}$ se establezcan a cero o se omitan, respectivamente (véase “6.5.3.2 derivación de elemento de matriz arbitraria” de ref. [2] para detalles).

En otra realización preferida, los elementos $H11_{OTT}^{l,m}$ y $H21_{OTT}^{l,m}$ de la matriz de mezcla ascendente $R_2^{l,m}$ se calcularán estableciendo $\text{ICC}^{l,m} = 1$.

La figura 7 ilustra la mezcla descendente de los canales principales L, R, LS, LR y C a canales estéreo L' y R'. Como los canales L y R creados por el procesador 36 no se mezclan en un canal común de la señal de audio de salida 31, el descorrelacionador 39 del procesador 36 permanece encendido. De la misma manera, el descorrelacionador 39' del procesador 36' permanece encendido cuando los canales LS y RS creados por el procesador 36' no se mezclan en un canal común de la señal de audio de salida 31. El canal LFE de altavoz de mejora de baja frecuencia podría usarse de manera opcional.

La figura 8 ilustra una mezcla descendente del contexto de altavoz de referencia 5.1 42 que se muestra en la figura 6 a un contexto de altavoz objetivo 4.0 45. Como los canales L y R creados por el procesador 36 no se mezclan en un canal común de la señal de audio de salida 31, el descorrelacionador 39 del procesador 36 permanece encendido. Sin embargo, los canales 13.3 (LS en la figura 6) y 13.4 (RS en la figura 6) creados por el procesador 36' se mezclan en un canal común 31.3 de la señal de audio de salida 31 con el fin de formar un canal CS de altavoz envolvente central. Por lo tanto, el descorrelacionador 39' del procesador 36' se apaga, de modo que el canal 13.3 es un canal CS' de altavoz envolvente central y de modo que el canal 13.4 es un canal CS'' de altavoz envolvente central. Al hacer esto, se genera un contexto de altavoz de referencia modificado 42'. Obsérvese que los canales CS' y CS'' están correlacionados, pero no son idénticos.

Por razones de exhaustividad, debe añadirse que los canales 13.5 (C) y 13.6 (LFE) se mezclan en un canal común 31.4 de la señal de audio de salida 31 con el fin de formar un canal C de altavoz frontal central.

En la figura 9, se muestra un decodificador de núcleo 6 que proporciona una señal de salida de decodificador de núcleo 13 que comprende los canales de salida 13.1 a 13.10 adecuados para un contexto de altavoz 9.1 de referencia 42, que comprende un canal L de altavoz frontal izquierdo, un canal LC de altavoz central frontal izquierdo, un canal LS del altavoz envolvente izquierdo, un LVR posterior de altura vertical envolvente izquierdo, un canal R del altavoz frontal derecho, un canal RS del altavoz envolvente derecho, un canal RC de altavoz central frontal derecho, un canal RS del altavoz envolvente derecho, un RVR posterior de altura vertical envolvente izquierdo, un canal C del altavoz frontal central y un canal LFE de altavoz de mejora de baja frecuencia.

Los canales de salida 13.1 y 13.2 se crean por el procesador 36 basándose en elementos de pares de canales (ID_USAC_CPE), que se alimentan al procesador 36, como canales descorrelacionados 13.1 y 13.2, cuando el descorrelacionador 39 del procesador 36 se enciende.

De manera análoga, los canales de salida 13.3 y 13.4 se crean por el procesador 36' basándose en elementos de pares de canales (ID_USAC_CPE), que se alimentan al procesador 36', como canales descorrelacionados 13.3 y 13.4, cuando el descorrelacionador 39' del procesador 36' se enciende.

Adicionalmente, los canales de salida 13.5 y 13.6 se crean por el procesador 36'' basándose en elementos de pares de canales (ID_USAC_CPE), que se alimentan al procesador 36'', como los canales descorrelacionados 13.5 y 13.6, cuando el descorrelacionador 39'' del procesador 36'' se enciende.

5 Aún más, los canales de salida 13.7 y 13.8 se crean por el procesador 36''' basándose en elementos de pares de canales (ID_USAC_CPE), que se alimentan al procesador 36''', como canales descorrelacionados 13.7 y 13.8, cuando el descorrelacionador 39''' del procesador 36''' se enciende.

10 El canal de salida 13.9 se basa en elementos de canales únicos (ID_USAC_SCE), mientras que el canal de salida 13.10 se basa en elementos de mejora de baja frecuencia ID_USAC_LFE.

15 La figura 10 ilustra una mezcla descendente del contexto de altavoz 9.1 de referencia 42 que se muestra en la figura 9, a un contexto de altavoz objetivo 5.1 45. Como los canales 13.1 y 13.2 creados por el procesador 36 se mezclan en un canal común 31.1 de la señal de audio de salida 31 con el fin de formar un canal L' de altavoz frontal izquierdo, el descorrelacionador 39 del procesador 36 se apaga, de modo que el canal 13.1 es un canal L' de altavoz frontal izquierdo, y de modo que el canal 13.2 es un canal L'' de altavoz frontal izquierdo.

20 Adicionalmente, los canales 13.3 y 13.4 creados por el procesador 36' se mezclan en un canal común 31.2 de la señal de audio de salida 31 con el fin de formar un canal LS de altavoz envolvente izquierdo. Por lo tanto, el descorrelacionador 39' del procesador 36' se apaga, de modo que el canal 13.3 es un canal LS' de altavoz envolvente izquierdo y de modo que el canal 13.4 es un canal LS'' de altavoz envolvente izquierdo.

25 Como los canales 13.5 y 13.6 creados por el procesador 36'' están mezclados en un canal común 31.3 de la señal de audio de salida 31 con el fin de formar un canal L de altavoz frontal derecho, el descorrelacionador 39'' del procesador 36'' se apaga, de modo que el canal 13.5 es un canal R' de altavoz frontal derecho y de modo que el canal 13.2 es un canal R'' de altavoz frontal derecho.

30 Además, los canales 13.7 y 13.8 creados por el procesador 36''' se mezclan en un canal común 31.4 de la señal de audio de salida 31 con el fin de formar un canal RS de altavoz envolvente derecho. Por lo tanto, el descorrelacionador 39''' del procesador 36''' se apaga, de modo que el canal 13.7 es un canal RS' de altavoz envolvente derecho y de modo que el canal 13.8 es un canal RS'' de altavoz envolvente derecho.

35 Al hacer esto, se genera un contexto de altavoz de referencia modificado 42', donde la cantidad de canales incoherentes de la señal de salida del decodificador de núcleo 13 es igual a la cantidad de los canales de altavoz del contexto objetivo 45.

40 Debe observarse que este procesamiento solo se aplicará a bandas de frecuencia donde se aplica la descorrelación. Las bandas de frecuencia donde se usa la codificación residual no se ven afectadas.

Como se menciona anteriormente, la invención puede aplicarse para la renderización binaural. La reproducción binaural se produce, típicamente, en auriculares y/o dispositivos móviles. En ese caso, pueden existir limitaciones, que limitan el decodificador y la complejidad de la renderización.

45 Puede realizarse la reducción u omisión del procesamiento de descorrelacionador. En el caso de que la señal de audio sea eventualmente procesada para la reproducción binaural, se propone omitir o reducir la descorrelación en la totalidad o parte de los bloques de decodificación de OTT.

50 Esto evita fallas de las señales de audio de mezcla descendente que se descorrelacionaron en el decodificador.

La cantidad de canales de salida decodificados para la renderización binaural puede reducirse. Además de omitir la descorrelación, puede ser deseable decodificar una cantidad menor de canales de salida incoherentes, lo que entonces da como resultado una menor cantidad de canales de entrada incoherentes para la renderización binaural. Por ejemplo, el material de canal original 22.2, que decodifica a 5.1, y la renderización binaural de solo 5 canales en lugar de 22, si la decodificación tiene lugar en un dispositivo móvil.

Con el fin de reducir la complejidad general del decodificador, se propone aplicar el siguiente procesamiento:

60 A) Definir un contexto de altavoz objetivo con una cantidad menor de canales que la configuración de canal original. La cantidad de canales objetivo depende de las limitaciones de calidad y complejidad.

Con el fin de lograr el contexto de altavoz objetivo, existen dos posibilidades, B1 y B2, que también pueden combinarse:

B1) Decodificar hasta una cantidad menor de canales, es decir, saltando el bloque de procesamiento de OTT completo en el decodificador. Esto requiere una vía de información procedente del renderizador binaural hasta el decodificador de núcleo (USAC) para controlar el procesamiento del decodificador.

- 5 B2) Aplicar una etapa de conversión de formato (es decir, realizando una mezcla descendente) desde la configuración de canal de altavoz original o una configuración de canal intermedia hasta el contexto de altavoz objetivo. Esto puede realizarse en una etapa de posprocesamiento después del decodificador de núcleo (USAC), y no requiere un proceso de decodificación alterado.

- 10 Finalmente, se realiza la etapa C):

C) Realizar la renderización binaural de una cantidad menor de canales.

Aplicación para la decodificación de SAOC.

- 15 Los métodos descritos anteriormente también pueden aplicarse al procesamiento de codificación de objeto paramétrico (SAOC).

- 20 Puede realizarse la conversión de formato con reducción/omisión de procesamiento de descorrelacionador. Si se aplica la conversión de formato después de la decodificación SAOC, se transmite la información del conversor de formato al decodificador SAOC. Con dicha información, se controla la correlación dentro del decodificador SAOC para reducir la cantidad de señales descorrelacionadas artificialmente. Esta información puede ser la matriz de mezcla descendente total, o información derivada.

- 25 Adicionalmente, puede ejecutarse la renderización binaural con reducción/omisión del procesamiento de descorrelacionador. En el caso de la codificación de objeto paramétrico (SAOC), la descorrelación se aplica en el proceso de decodificación. El procesamiento de descorrelación dentro del decodificador SAOC debe omitirse o reducirse si sigue la renderización binaural.

- 30 Además, puede ejecutarse la renderización binaural con un número reducido de canales. Si se aplica la reproducción binaural después de la decodificación SAOC, el decodificador SAOC puede configurarse para renderizar hasta una menor cantidad de canales, usando una matriz de mezcla descendente que se construye basándose en la información del conversor de formato.

- 35 Como la filtración de descorrelación requiere una complejidad computacional sustancial, la carga de trabajo de decodificación general puede reducirse en gran medida por el método propuesto.

- 40 Si bien los filtros paso todo están diseñados de tal manera para tener un mínimo impacto sobre la calidad de sonido subjetiva, no siempre puede evitarse la introducción de fallas audibles, por ejemplo, embarrado de transientes debido a distorsiones de fases o "zumbido" de ciertos componentes de frecuencia. Por lo tanto, puede lograrse una mejora de la calidad de sonido del audio, ya que los efectos secundarios del proceso de descorrelacionador se omiten. Además, se evita cualquier desenmascarado de dichas fallas de descorrelacionador por la posterior mezcla descendente, mezcla ascendente o procesamiento binaural.

- 45 Adicionalmente, se han descrito los métodos para la reducción de la complejidad en el caso de la renderización binaural en combinación con un decodificador de núcleo (USAC) o un decodificador SAOC.

Con respecto al decodificador y codificador y los métodos de las realizaciones descritas, se menciona lo siguiente:

- 50 Si bien se han descrito algunos aspectos en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a una etapa de método o una característica de una etapa de método. De manera análoga, los aspectos descritos en el contexto de una etapa de método también representan una descripción de un bloque o artículo o característica correspondiente de un aparato correspondiente.

- 55 Según determinados requisitos de implementación, las realizaciones de la invención pueden implementarse en hardware o software. La implementación puede realizarse usando un medio de almacenamiento digital, por ejemplo, un disquete, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene señales de control electrónicamente legibles almacenadas en la misma, que actúan conjuntamente (o pueden actuar conjuntamente) con un sistema informático programable de manera que se realiza el método respectivo.

- 60

Algunos ejemplos que no se encuentran dentro del alcance de las reivindicaciones comprenden un portador de datos que tiene señales de control electrónicamente legibles, que pueden actuar conjuntamente con un sistema informático programable, de manera que se realiza uno de los métodos descritos en el presente documento.

En general, los ejemplos que no se encuentran dentro del alcance de las reivindicaciones pueden implementarse como un producto de programa informático con un código de programa, donde el código de programa es operativo para llevar a cabo uno de los métodos cuando se ejecuta el producto de programa informático en un ordenador. El código de programa puede almacenarse, por ejemplo, en un portador legible por máquina. Otras realizaciones comprenden el programa informático para llevar a cabo uno de los métodos descritos en el presente documento, que pueden almacenarse en un portador legible por máquina o en un medio de almacenamiento no transitorio. En otras palabras, una realización del método de la invención, por lo tanto, es un programa informático que tiene un código de programa para la realización de uno de los métodos descritos en el presente documento, cuando se ejecuta el programa informático en un ordenador.

Un ejemplo adicional que no se encuentra dentro del alcance de las reivindicaciones es un portador de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para llevar a cabo uno de los métodos descritos en el presente documento.

Un ejemplo adicional que no se encuentra dentro del alcance de las reivindicaciones es una corriente de datos o una secuencia de señales que representan el programa informático para llevar a cabo uno de los métodos que se describen en el presente documento. La corriente de datos o la secuencia de señales, por ejemplo, puede estar configurada para ser transferida por medio de una conexión de comunicación de datos, por ejemplo, por medio de Internet.

Un ejemplo adicional que no se encuentra dentro del alcance de las reivindicaciones comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado o adaptado para llevar a cabo uno de los métodos que se describen en el presente documento.

Un ejemplo adicional que no se encuentra dentro del alcance de las reivindicaciones comprende un ordenador que tiene instalado en el mismo el programa informático para llevar a cabo uno de los métodos que se describen en el presente documento.

En algunas realizaciones, puede usarse un dispositivo lógico programable (por ejemplo, una matriz de puertas programables de campo) para realizar algunas o la totalidad de las funcionalidades de los métodos que se describen en el presente documento. En algunas realizaciones, una matriz de puertas programables de campo puede actuar conjuntamente con un microprocesador con el fin de llevar a cabo uno de los métodos que se describen en el presente documento. En general, los métodos son llevados a cabo, ventajosamente, por cualquier aparato de hardware.

Referencias

- [1] Surround Sound Explained - parte 5. publicado en: soundonsound magazine, diciembre de 2001.
- [2] norma ISO/IEC IS 23003-1, MPEG audio technologies - parte 1: MPEG Surround
- [3] norma ISO/IEC IS 23003-3, MPEG audio technologies - parte 3: Unified speech and audio coding.

REIVINDICACIONES

1. Dispositivo decodificador de audio para la decodificación de una señal de audio de entrada comprimida que comprende

5 por lo menos un decodificador de núcleo (6, 24) que tiene uno o más procesadores (36, 36') para la generación de una señal de salida del procesador (37) basándose en una señal de entrada del procesador (38, 38'), donde una cantidad de canales de salida (37.1, 37.2, 37.1', 37.2') de la señal de salida del procesador (37, 37') es mayor que una cantidad de canales de entrada (38.1, 38.1') de la señal de entrada del procesador (38, 38'), donde cada uno de los uno o más procesadores (36, 36') comprende un descorrelacionador (39, 39') y una mezcladora (40, 40'), donde una señal de salida del decodificador de núcleo (13) que tiene una pluralidad de canales (13.1, 13.2, 13.3, 13.4) comprende la señal de salida del procesador (37, 37'), y donde la señal de salida del decodificador de núcleo (13) es adecuada para un contexto de altavoz de referencia (42);

15 por lo menos un dispositivo conversor de formato (9, 10) configurado para convertir la señal de salida del decodificador de núcleo (13) en una señal de audio de salida (31), que es adecuada para un contexto de altavoz objetivo (45); y

20 un dispositivo de control (46) configurado para el control de por lo menos uno o más procesadores (36, 36') de tal manera que el descorrelacionador (39, 39') del al menos un procesador (36, 36') puede ser controlado independientemente de la mezcladora (40, 40') de dicho al menos un procesador (36, 36'),

25 caracterizado porque el dispositivo de control (46) está configurado para el control de por lo menos uno de los descorrelacionadores (39, 39') de los uno o más procesadores (36, 36') de tal manera que, según el contexto de altavoz objetivo (45), la mezcladora (40, 40') del procesador (36, 36') esté operativa cuando el descorrelacionador (39, 39') del procesador (36, 36') está apagado.
2. Dispositivo de decodificador según la reivindicación 1, donde el dispositivo de control (46) está configurado para desactivar por lo menos uno o más procesadores (36, 36') de modo que los canales de entrada (38.1, 38.1') de la señal de entrada del procesador (38, 38') se alimentan a los canales de salida (37.1, 37.2, 37.1', 37.2') de la señal de salida del procesador (37, 37') de una forma no procesada.
3. Dispositivo de decodificador según la reivindicación 1 o 2, donde el procesador (36, 36') es una herramienta de decodificación de una entrada y dos salidas, donde el descorrelacionador (39, 39') está configurado para crear una señal descorrelacionada (48) mediante la descorrelación de por lo menos uno de los canales (38.1, 38.1') de la señal de entrada del procesador (38, 38'), donde la mezcladora (40, 40') mezcla una señal de entrada del procesador (38) y la señal descorrelacionada (46) basándose en una señal de diferencia de nivel de canal (49) y/o una señal de coherencia entre canales (50), de manera que una señal de salida del procesador (37, 37') consiste en dos canales de salida incoherentes (37.1, 37.2, 37.1', 37.2').
4. Dispositivo de decodificador según la reivindicación 3, donde el dispositivo de control está configurado para apagar el descorrelacionador (39, 39') de uno de los procesadores (36, 36') estableciendo la señal descorrelacionada (48) en cero o impidiendo que la mezcladora (40, 40') mezcle la señal descorrelacionada (48) en una señal de salida del procesador (37) del procesador (36, 36') respectivo.
5. Dispositivo de decodificador según una de las reivindicaciones 1 a 4, donde el decodificador de núcleo (6) es un decodificador tanto para música como para voz, donde una señal de entrada del procesador (38) de por lo menos uno de los procesadores (36, 36') contiene elementos de pares de canales.
6. Dispositivo de decodificador según una de las reivindicaciones 1 a 5, donde el decodificador de núcleo (24) es un codificador de objeto paramétrico.
7. Dispositivo de decodificador según una de las reivindicaciones 1 a 6, donde la cantidad de altavoces del contexto de altavoz de referencia (42) es mayor que una cantidad de altavoces del contexto de altavoz objetivo (45).
8. Dispositivo de decodificador según una de las reivindicaciones 1 a 7, donde el dispositivo de control (46) está configurado para apagar los descorrelacionadores (39') para por lo menos uno primero de dichos canales de salida (37.1') de la señal de salida del procesador (37') y uno segundo de dichos canales de salida (37.2') de la señal de salida del procesador (37'), si el primero de dichos canales de salida (37.1') y el segundo de dichos canales de salida (37.2'), según el contexto de altavoz objetivo, están mezclados en un canal común (31.2) de la señal de audio de salida (31), siempre y cuando un primer factor de escala para la mezcla del primero de dichos canales de salida (37,1') en el canal común (31.2) exceda un primer umbral,

y/o un segundo factor de escala para la mezcla del segundo de dichos canales de salida (37.2') en el canal común (31.2) exceda un segundo umbral.

- 5 9. Dispositivo de decodificador según una de las reivindicaciones 1 a 8, donde el dispositivo de control (46) está configurado para recibir un conjunto de pautas (47) desde el dispositivo conversor de formato (9, 10) según las cuales el dispositivo conversor de formato (9, 10) mezcla los canales (13.1, 13.2, 13.3, 13.4) de la señal de salida del decodificador de núcleo (13) en los canales (31.1, 31.2, 31.3) de la señal de audio de salida (31) según el contexto de altavoz objetivo (45), donde el dispositivo de control (46) está configurado para controlar por lo menos uno de los procesadores (36, 36') según el conjunto de pautas recibidas (47).
- 10 10. Dispositivo de decodificador según una de las reivindicaciones 1 a 9, donde el dispositivo de control (46) está configurado para controlar los descorrelacionadores (39, 39') de los procesadores (36, 36') de tal manera que una cantidad de canales incoherentes de la señal de salida del decodificador de núcleo (13) es igual a la cantidad de los canales (31.1, 31.2, 31.3) de la señal de audio de salida (31).
- 15 11. Dispositivo de decodificador según una de las reivindicaciones 1 a 10, donde el dispositivo conversor de formato (9, 10) comprende una mezcladora descendente (10) para la mezcla descendente de la señal de salida del decodificador de núcleo (13).
- 20 12. Dispositivo de decodificador según una de las reivindicaciones 1 a 11, donde el dispositivo conversor de formato (9, 10) comprende un renderizador binaural (10).
- 25 13. Dispositivo de decodificador según la reivindicación 12, donde la señal de salida del decodificador de núcleo (13) se alimenta al renderizador binaural (9) como una señal de entrada de renderizador binaural.
- 30 14. Dispositivo de decodificador según la reivindicación 11 y una de las reivindicaciones 12 a 13, donde una señal de salida de mezcladora descendente, de la mezcladora descendente (9), se alimenta al renderizador binaural (10) como una señal de entrada de renderizador binaural.
- 35 15. Método para la decodificación de una señal de audio de entrada comprimida, donde el método comprende las etapas de:

proporcionar por lo menos un decodificador de núcleo (6, 24) que tiene uno o más procesadores (36, 36') para la generación de una señal de salida del procesador (37) basándose en una señal de entrada del procesador (38), donde una cantidad de canales de salida (37.1, 37.2, 37.1', 37.2') de la señal de salida del procesador (37, 37') es mayor que una cantidad de canales de entrada (38.1, 38.1') de la señal de entrada del procesador (38, 38'), donde cada uno de los uno o más procesadores (36, 36') comprende un descorrelacionador (39, 39') y una mezcladora (40, 40'), donde una señal de salida de decodificador de núcleo (13) que tiene una pluralidad de canales (13.1, 13.2, 13.3, 13.4) comprende la señal de salida del procesador (37, 37'), y donde la señal de salida de decodificador de núcleo (13) es adecuada para un contexto de altavoz de referencia (42);

proporcionar por lo menos un dispositivo conversor de formato (9, 10) configurado para convertir la señal de salida del decodificador de núcleo (13) en una señal de audio de salida (31), que es adecuada para un contexto de altavoz objetivo (45); y

proporcionar un dispositivo de control (46) configurado para el control de por lo menos uno o más procesadores (36, 36') de tal manera que el descorrelacionador (39, 39') del procesador (36, 36') puede ser controlado de forma independiente de la mezcladora (40, 40') de dicho al menos un procesador (36, 36'),

caracterizado porque

el dispositivo de control (46) está configurado para el control de por lo menos uno de los descorrelacionadores (39, 39') de los uno o más procesadores (36, 36') de tal manera que, según el contexto de altavoz objetivo (45), la mezcladora (40, 40') del procesador (36, 36') esté operativa cuando el descorrelacionador (39, 39') del procesador (36, 36') está apagado.
- 60 16. Programa informático para la implementación del método según la reivindicación 15 cuando se ejecuta en un ordenador o en un procesador de señales.

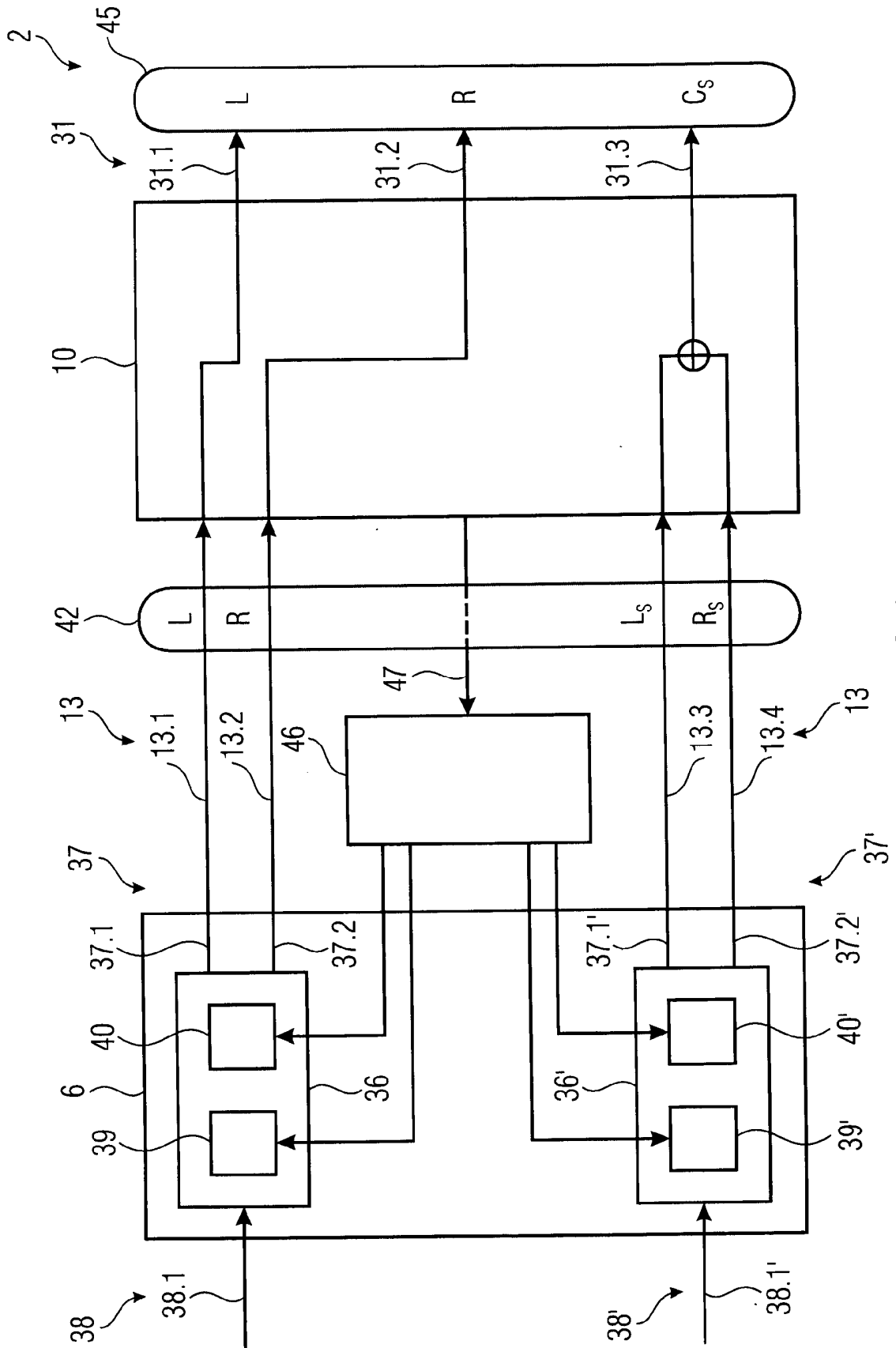


FIG 1

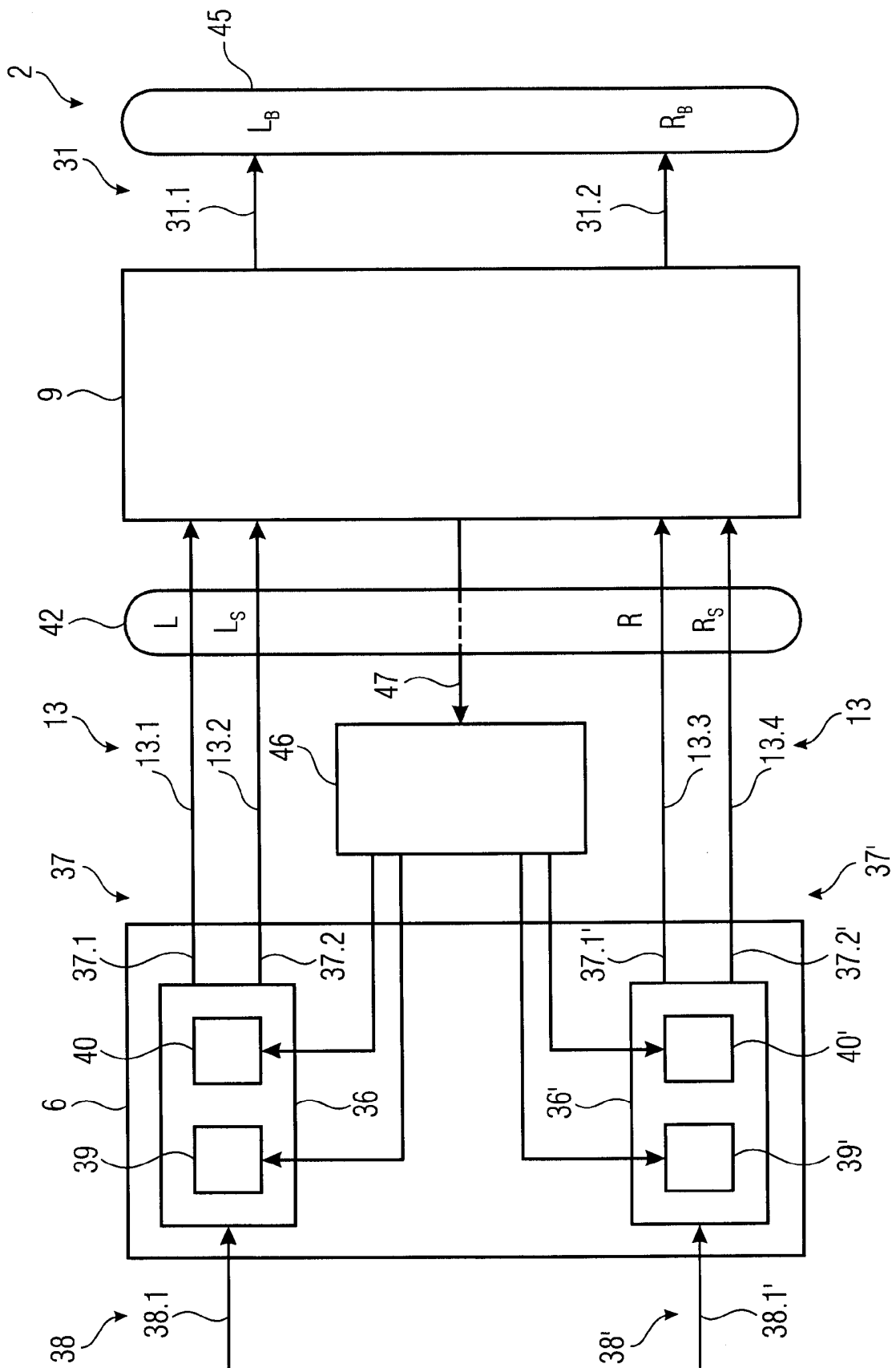


FIG 2

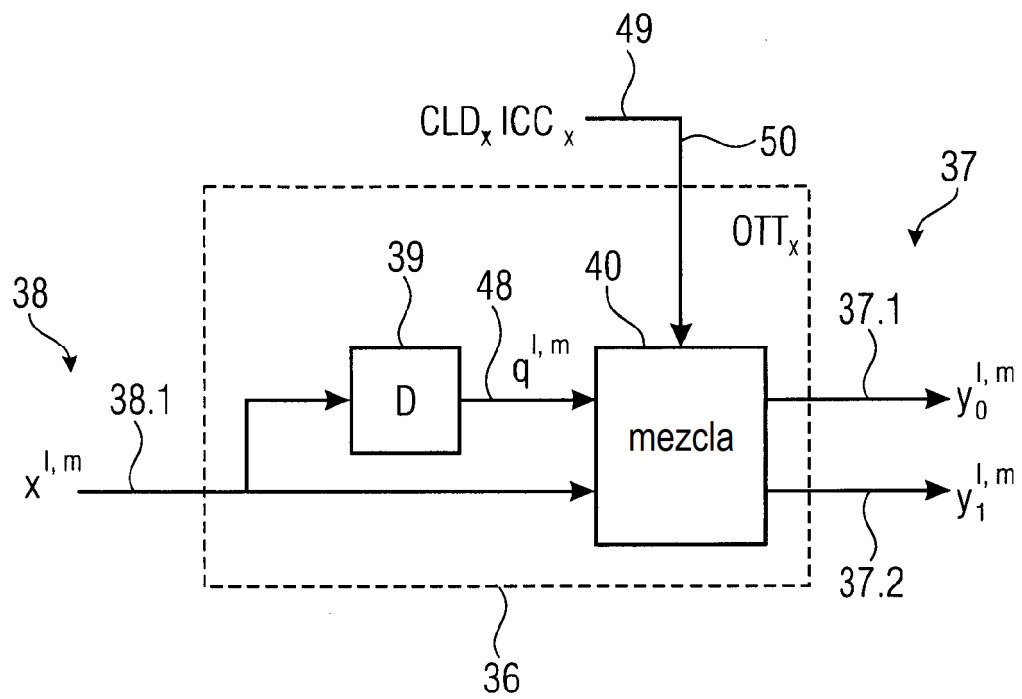


FIG 3

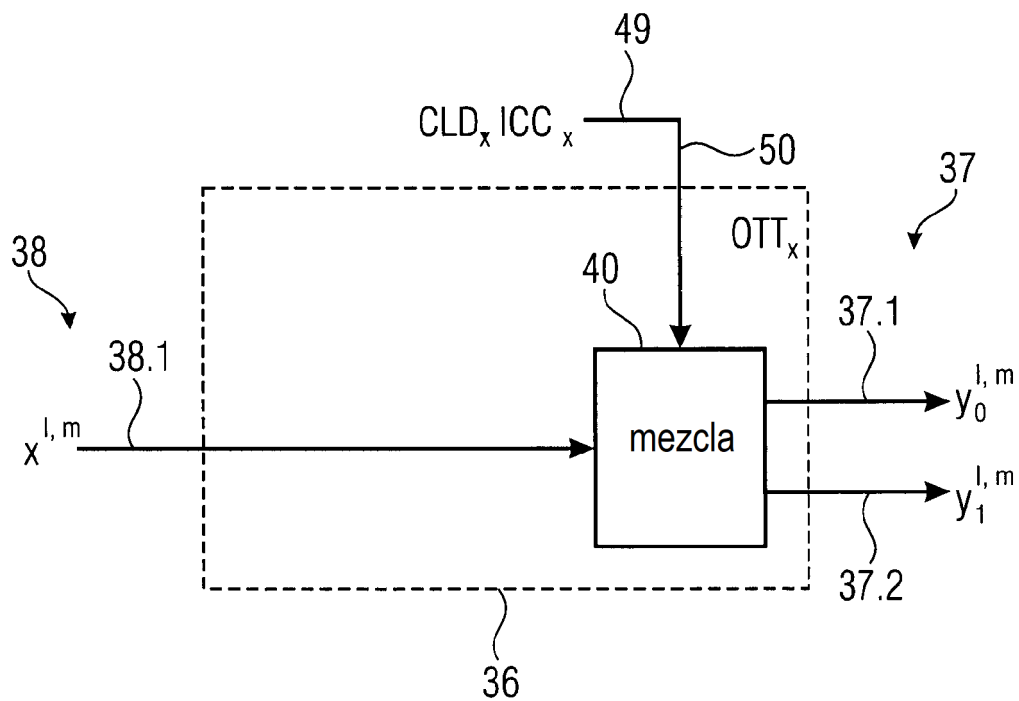


FIG 4

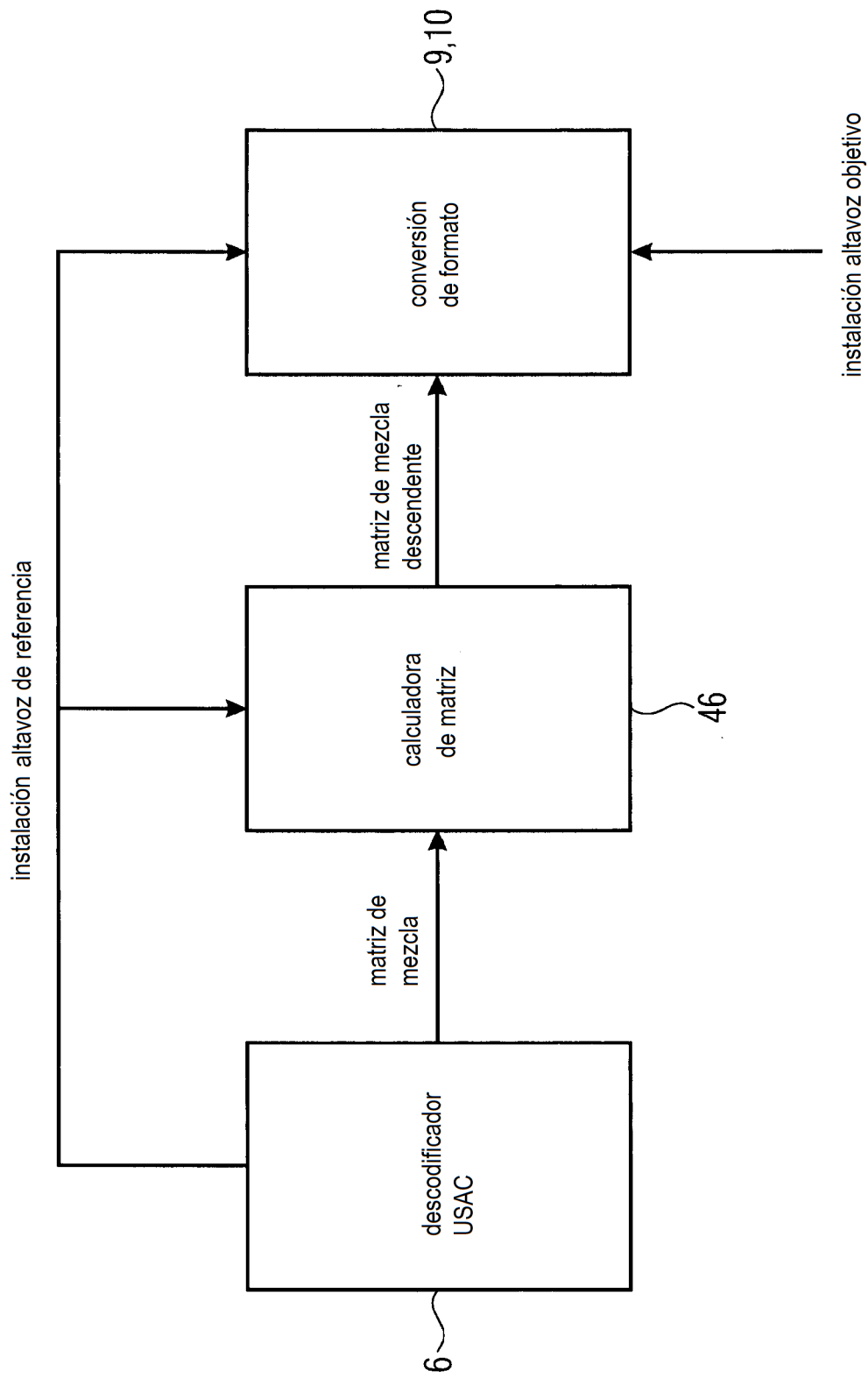


FIG 5

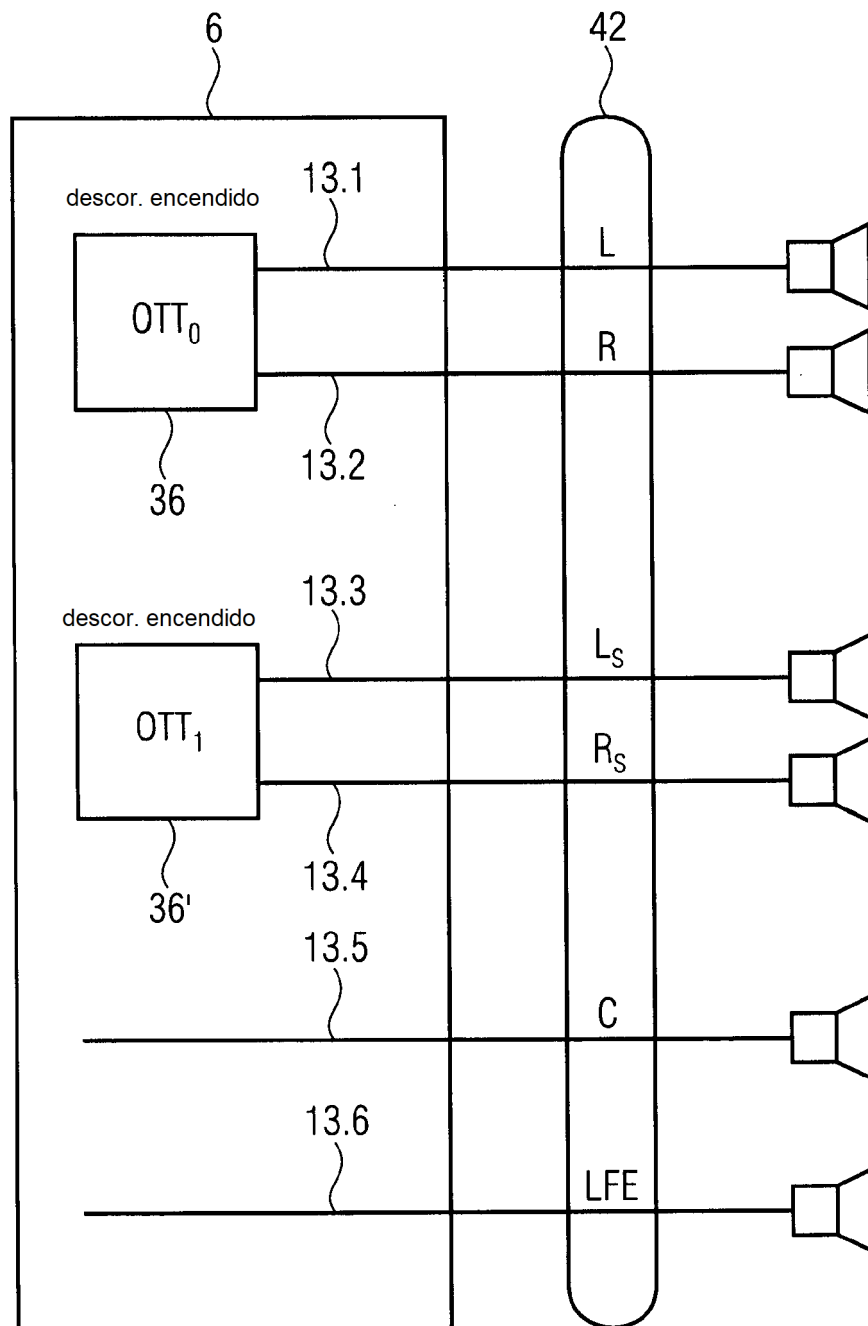


FIG 6

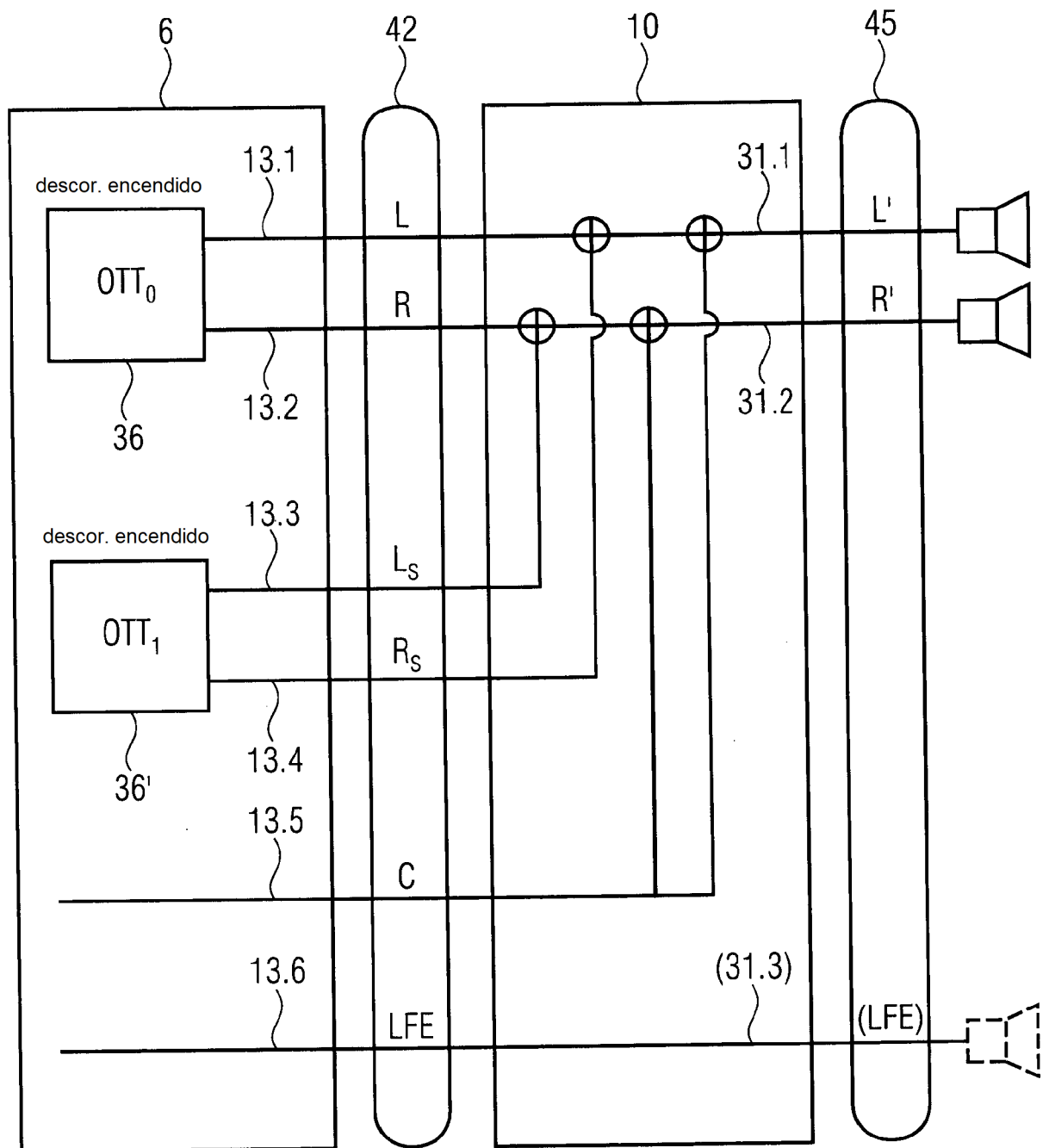


FIG 7

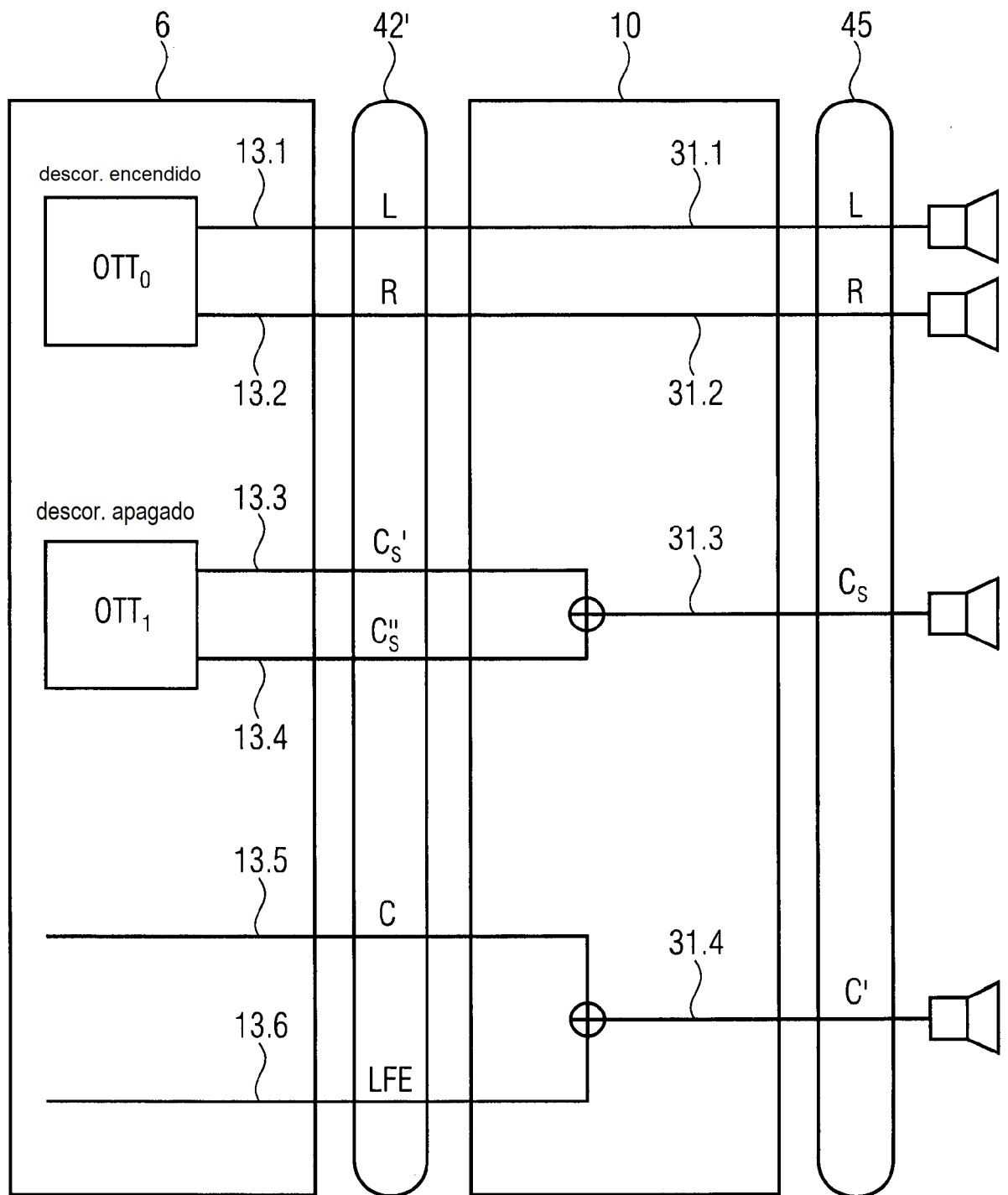


FIG 8

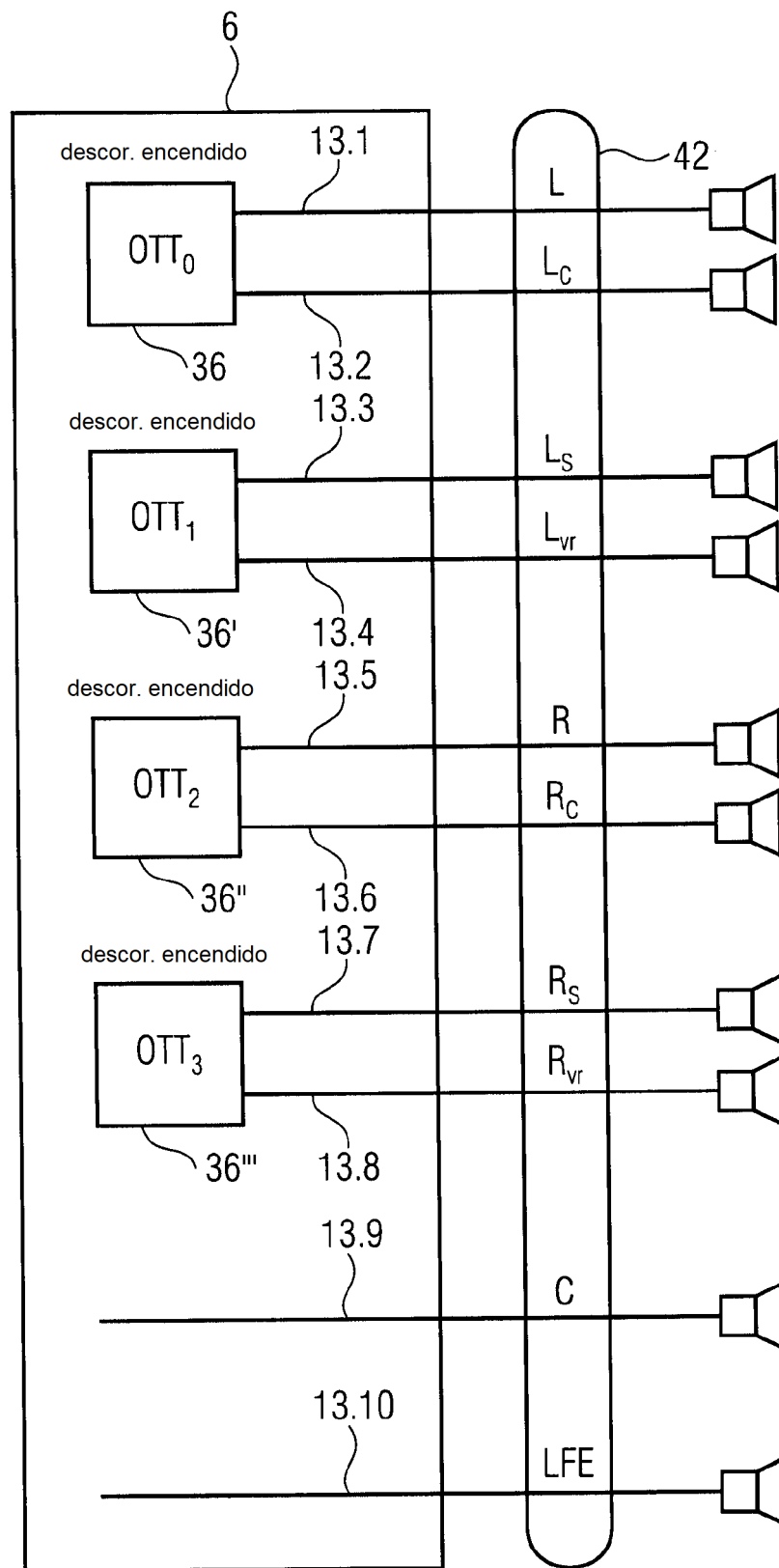


FIG 9

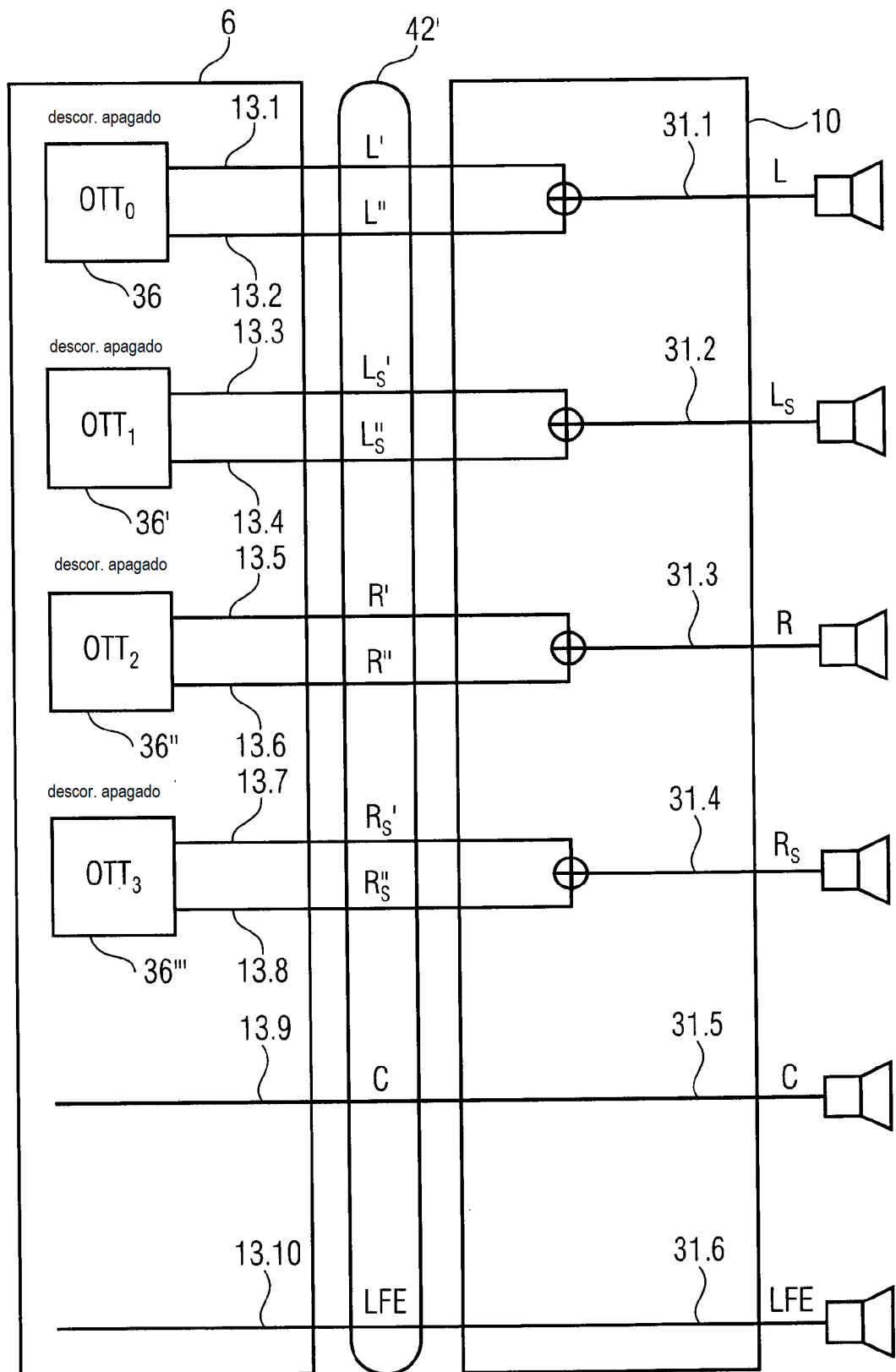


FIG 10

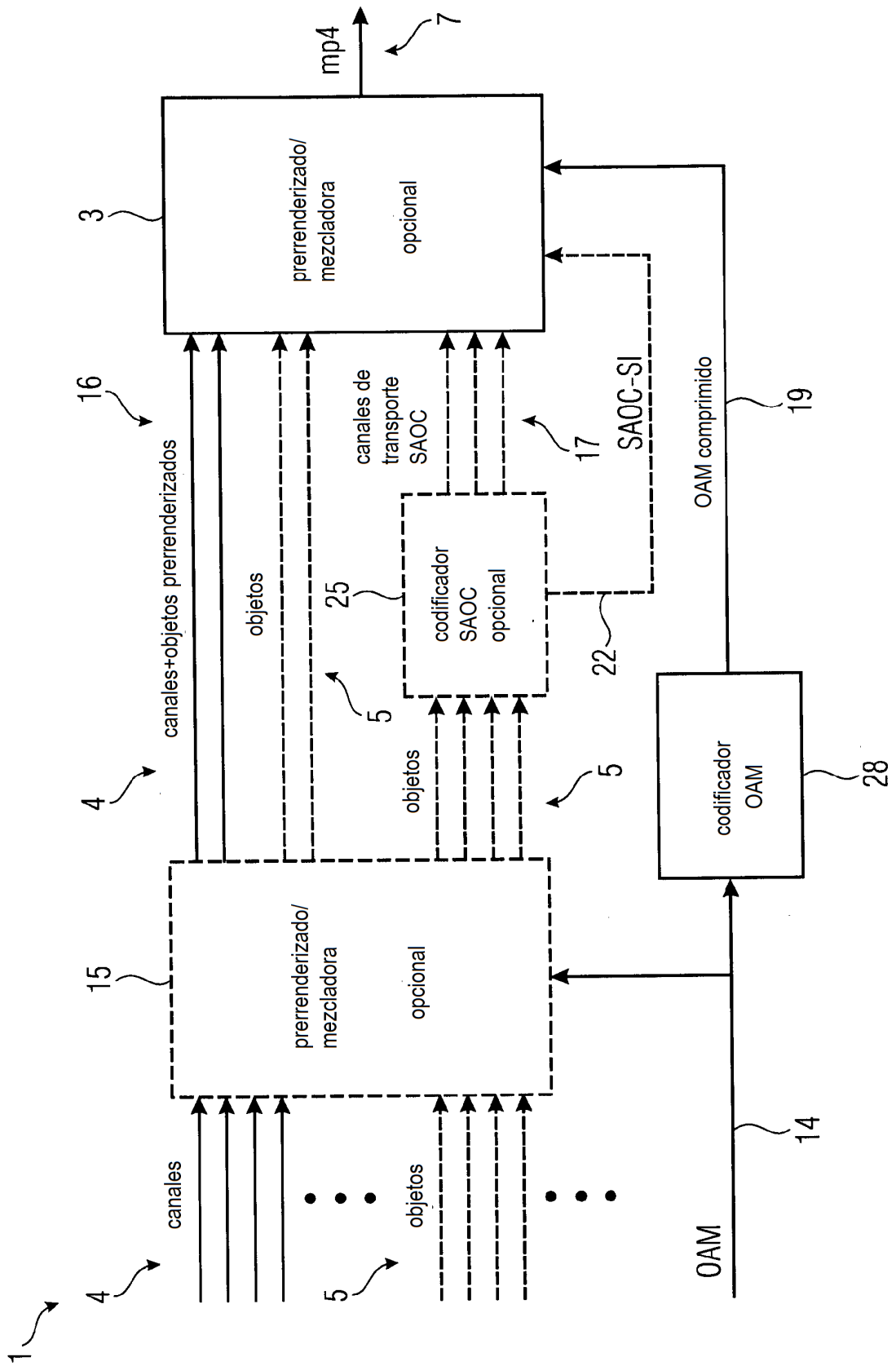


FIG 11

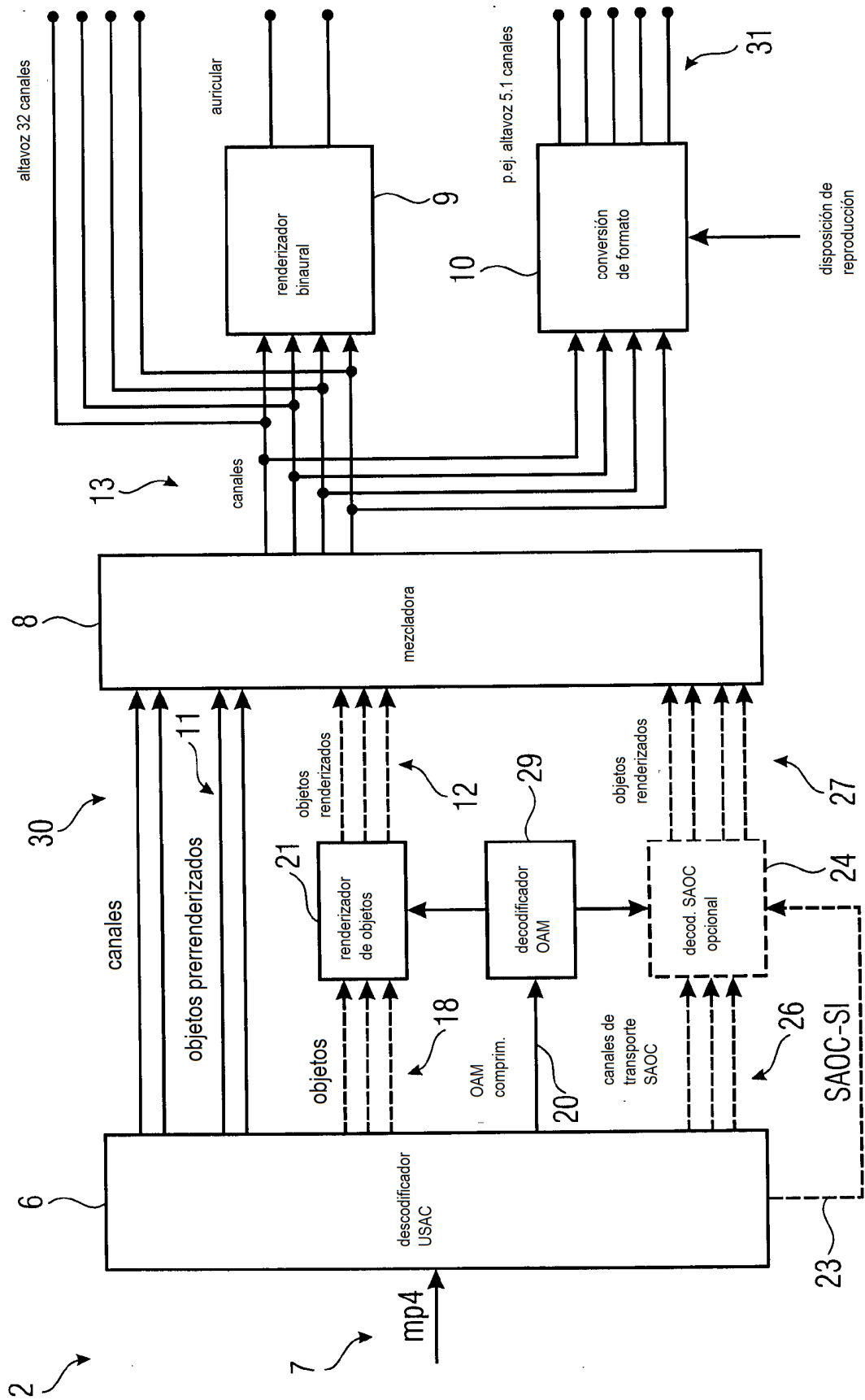


FIG 12

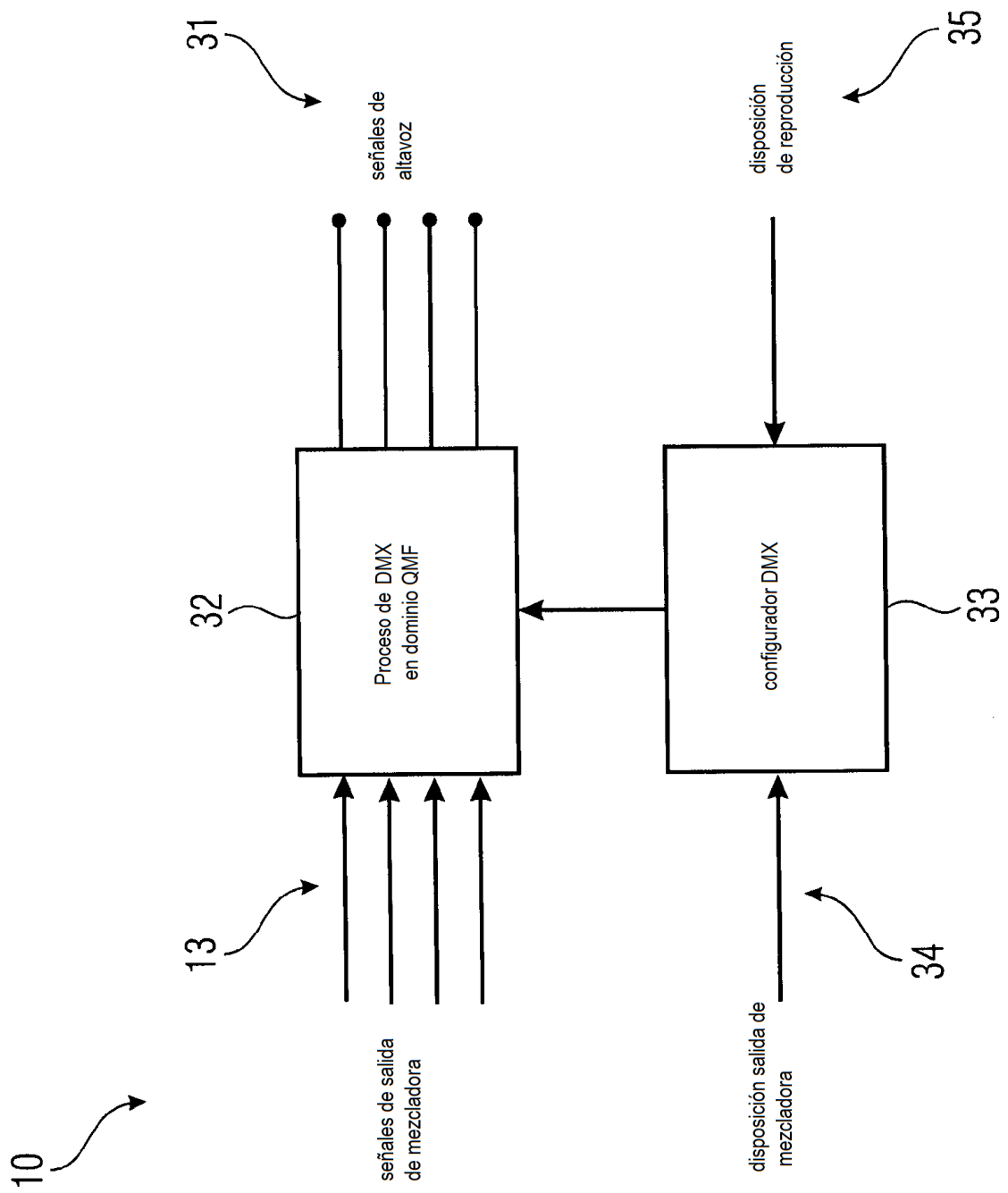


FIG 13