

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 732 668**

51 Int. Cl.:

**G10L 19/008** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.10.2015 PCT/EP2015/075022**

87 Fecha y número de publicación internacional: **06.05.2016 WO16066705**

96 Fecha de presentación y número de la solicitud europea: **28.10.2015 E 15787573 (3)**

97 Fecha y número de publicación de la concesión europea: **03.04.2019 EP 3213322**

54 Título: **Mezcla paramétrica de señales de audio**

30 Prioridad:

**31.10.2014 US 201462073462 P**  
**28.05.2015 US 201562167711 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**25.11.2019**

73 Titular/es:

**DOLBY INTERNATIONAL AB (100.0%)**  
**Apollo Building, 3E, Herikerbergweg 1-35**  
**1101 CN Amsterdam Zuid-Oost, NL**

72 Inventor/es:

**VILLEMoes, LARS;**  
**PURNHAGEN, HEIKO y**  
**LEHTONEN, HEIDI-MARIA**

74 Agente/Representante:

**ELZABURU, S.L.P**

**ES 2 732 668 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 paramétrica de señales de audio

**Campo técnico**

5 La invención descrita en la presente memoria se refiere de manera general a codificación y decodificación de señales de audio, y en particular a mezcla de canales de una señal de mezcla descendente basada en metadatos asociados.

**Antecedentes**

10 Los sistemas de reproducción de audio que comprenden múltiples altavoces se usan con frecuencia para reproducir una escena de audio representada por una señal de audio multicanal, en donde los canales respectivos de la señal de audio multicanal se reproducen en los altavoces respectivos. La señal de audio multicanal, por ejemplo, puede haber sido grabada a través de una pluralidad de transductores acústicos o puede haber sido generada por equipos de creación de audio. En muchas situaciones, hay limitaciones de ancho de banda para transmitir la señal de audio al equipo de reproducción y/o espacio limitado para almacenar la señal de audio en una memoria de ordenador o en un dispositivo de almacenamiento portátil. Existen sistemas de codificación de audio para la codificación paramétrica de señales de audio, para reducir el ancho de banda o el almacenamiento necesario. En el lado del codificador, estos sistemas típicamente mezclan de manera descendente la señal de audio multicanal en una señal de mezcla descendente, que típicamente es una mezcla descendente mono (un canal) o estéreo (dos canales), y extraen información del lado que describe las propiedades de los canales por medio de parámetros como diferencias de nivel y correlación cruzada. La mezcla descendente y la información del lado entonces se codifican y envían al lado del decodificador. En el lado del decodificador, la señal de audio multicanal se reconstruye, es decir, se aproxima, a partir de la mezcla descendente bajo el control de los parámetros de la información del lado.

25 En vista de la amplia gama de diferentes tipos de dispositivos y sistemas disponibles para reproducción de contenido de audio multicanal, incluyendo un segmento emergente dirigido a usuarios finales en sus hogares, hay una necesidad de formas nuevas y alternativas para codificar de manera eficiente contenido de audio multicanal, para reducir los requisitos de ancho de banda y/o el tamaño de memoria requerido para el almacenamiento, facilitar la reconstrucción de la señal de audio multicanal en el lado del decodificador, y/o aumentar la fidelidad de la señal de audio multicanal que se reconstruye en el lado del decodificador. También hay una necesidad para facilitar la reproducción de contenido de audio multicanal codificado en diferentes tipos de sistemas de altavoces, incluyendo sistemas con menos altavoces que el número de canales presentes en el contenido de audio multicanal original.

30 El documento WO 2014/126689 A1 describe la aplicación de un proceso de filtrado de descorrelación a datos de audio multicanal, en base a las características de audio. Dicho proceso causa una coherencia de señal de intercorrelación específica entre señales de descorrelación específicas de canal para al menos un par de canales. Se puede controlar la coherencia entre canales entre una pluralidad de pares de canales de audio.

35 HERRE JURGEN ET AL: MPEG Surround - The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding”, JAES, AES, 60 EAST 42<sup>nd</sup> STREET, ROOM 2520 NUEVA YORK 10165-2520, EE.UU., volumen 56, n° 11, 1 de noviembre de 2008, páginas 932-9355, XP040508729 describe una codificación eficiente y compatible con versiones anteriores de sonido multicanal de alta calidad usando técnicas de codificación paramétrica.

40 El documento US 2006/0165184 A1 describe señales de reconstrucción multicanal de manera que los canales reconstruidos están al menos parcialmente descorrelacionados unos de otros usando una señal de mezcla descendente a partir de una señal multicanal original y un conjunto de señales descorrelacionadas proporcionadas por un descorrelacionador.

**Breve descripción de los dibujos**

En lo que sigue, se describirán ejemplos de realizaciones con mayor detalle y con referencia a los dibujos que se acompañan, en los que:

45 La Fig. 1 es un diagrama de bloques generalizado de una sección de codificación para codificar una señal de M canales como una señal de mezcla descendente de dos canales y metadatos asociados, según un ejemplo de realización;

La Fig. 2 es un diagrama de bloques generalizado de un sistema de codificación de audio que comprende la sección de codificación representada en la Fig. 1, según un ejemplo de realización;

50 La Fig. 3 es un diagrama de flujo de un método de codificación de audio para codificar una señal de audio de M canales como una señal de mezcla descendente de dos canales y metadatos asociados, según un ejemplo de realización;

Las Fig. 4-6 ilustran formas alternativas para dividir una señal de audio de canales 11.1 (o canales 7.1 + 4 o canales 7.1.4) en grupos de canales representados por canales de mezcla descendente respectivos, según ejemplos de realizaciones;

5 La Fig. 7 es un diagrama de bloques generalizado de una sección de decodificación para proporcionar una señal de salida de dos canales en base a una señal de mezcla descendente de dos canales y parámetros de mezcla ascendente asociados, según un ejemplo de realización;

La Fig. 8 es un diagrama de bloques generalizado de un sistema de decodificación de audio que comprende la sección de decodificación representada en la Fig. 7, según un ejemplo de realización;

10 La Fig. 9 es un diagrama de bloques generalizado de una sección de decodificación para proporcionar una señal de salida de dos canales en base a una señal de mezcla descendente de dos canales y parámetros de mezcla asociados, según un ejemplo de realización;

La Fig. 10 es un diagrama de flujo de un método de decodificación de audio para proporcionar una señal de salida de dos canales en base a una señal de mezcla descendente de dos canales y metadatos asociados, según un ejemplo de realización;

15 La Fig. 11 ilustra esquemáticamente un medio legible por ordenador, según un ejemplo de realización;

La Fig. 12 es un diagrama de bloques generalizado de una sección de decodificación para proporcionar una señal de salida de K canales en base a una señal de mezcla descendente de dos canales y parámetros de mezcla ascendente asociados, según un ejemplo de realización;

20 Las Fig. 13-14 ilustran formas alternativas para dividir una señal de audio de canales 11.1 (o canales 7.1 + 4 o canales 7.1.4) en grupos de canales, según ejemplos de realizaciones; y

Las Fig. 15-16 ilustran formas alternativas para dividir una señal de audio de canales 13.1 (o canales 9.1 + 4 o canales 9.1.4) en grupos de canales, según ejemplos de realizaciones.

Todas las figuras son esquemáticas y, en general, solamente muestran las partes que son necesarias con el fin de dilucidar la invención, mientras que otras partes se pueden omitir o meramente sugerir.

## 25 Descripción de ejemplos de realizaciones

Como se usa en la presente memoria, una señal de audio puede ser una señal de audio autónoma, una parte de audio de una señal audiovisual o una señal multimedia o cualquiera de éstas en combinación con metadatos.

Como se usa en la presente memoria, un canal es una señal de audio asociada con una posición/orientación espacial predefinida/fija o una posición espacial indefinida tal como "izquierda" o "derecha".

### 30 1. Descripción general - lado del decodificador

Según un primer aspecto, los ejemplos de realizaciones proponen sistemas de decodificación de audio, métodos de decodificación de audio y productos de programas de ordenador asociados. Los sistemas de decodificación, métodos y productos de programa de ordenador propuestos, según el primer aspecto, pueden compartir de manera general las mismas características y ventajas.

35 Según ejemplos de realizaciones, se proporciona un método de decodificación de audio que comprende recibir una señal de mezcla descendente de dos canales. La señal de mezcla descendente está asociada con metadatos que comprenden parámetros de mezcla ascendente para la reconstrucción paramétrica de una señal de audio de M canales en base a la señal de mezcla descendente, donde  $M \geq 4$ . Un primer canal de la señal de mezcla descendente corresponde a una combinación lineal de un primer grupo de uno o más canales de la señal de audio de M canales, y un segundo canal de la señal de mezcla descendente corresponde a una combinación lineal de un segundo grupo de uno o más canales de la señal de audio de M canales. El primer y segundo grupos constituyen una división de los M canales de la señal de audio de M canales. El método de decodificación de audio comprende además: recibir al menos una parte de los metadatos; generar una señal descorrelacionada en base a al menos un canal de la señal de mezcla descendente; determinar un conjunto de coeficientes de mezcla en base a los metadatos recibidos; y formar una señal de salida de dos canales como una combinación lineal de la señal de mezcla descendente y la señal descorrelacionada según los coeficientes de mezcla. Los coeficientes de mezcla se determinan de manera que un primer canal de la señal de salida se aproxime a una combinación lineal de un tercer grupo de uno o más canales de la señal de audio de M canales, y de manera que un segundo canal de la señal de salida se aproxime a una combinación lineal de un cuarto grupo de uno o más canales de la señal de audio de M canales. Los coeficientes de mezcla también se determinan de manera que el tercer y cuarto grupos constituyan una división de los M canales de la señal de audio de M canales, y de manera que ambos del tercer y cuarto grupos comprendan al menos un canal del primer grupo.

40

45

50

- 5 La señal de audio de M canales se ha codificado como la señal de mezcla descendente de dos canales y los parámetros de mezcla ascendente para la reconstrucción paramétrica de la señal de audio de M canales. Cuando se codifica la señal de audio de M canales en un lado del codificador, se puede elegir el formato de codificación, por ejemplo, para facilitar la reconstrucción de la señal de audio de M canales a partir de la señal de mezcla descendente, para mejorar la fidelidad de la señal de audio de M canales como reconstruida a partir de la señal de mezcla descendente, y/o para mejorar la eficiencia de codificación de la señal de mezcla descendente. Esta elección del formato de codificación se puede realizar seleccionando el primer y segundo grupos y formando los canales de las señales de mezcla descendente como combinaciones lineales respectivas de los canales en los grupos respectivos.
- 10 Los inventores se han dado cuenta de que aunque el formato de codificación elegido puede facilitar la reconstrucción de la señal de audio de M canales a partir de la señal de mezcla descendente, la señal de mezcla descendente puede no ser adecuada en sí misma para la reproducción usando una configuración particular de dos altavoces. La señal de salida, correspondiente a una división diferente de la señal de audio de M canales en el tercer y cuarto grupos, puede ser más adecuada para un ajuste de reproducción de dos canales particular que la señal de mezcla descendente. Proporcionar la señal de salida en base a la señal de mezcla descendente y los metadatos recibidos puede mejorar, por lo tanto, la calidad de reproducción de dos canales que se percibe por un oyente, y/o mejorar la fidelidad de la reproducción de dos canales a un campo de sonido representado por la señal de audio de M canales.
- 15 Los inventores se han dado cuenta además de que, en lugar de primero reconstruir la señal de audio de M canales a partir de la señal de mezcla descendente y entonces generar una representación de dos canales alternativa de la señal de audio de M canales (por ejemplo, mediante mezcla aditiva), la representación de dos canales alternativa proporcionada por la señal de salida se puede generar más eficientemente a partir de la señal de mezcla descendente y los metadatos recibidos explotando el hecho de que algunos canales de la señal de audio de M canales se agrupan juntos de manera similar en ambas de las representaciones de dos canales. La formación de la señal de salida como una combinación lineal de la señal de mezcla descendente y la señal descorrelacionada puede reducir, por ejemplo, la complejidad de cálculo en el lado del decodificador y/o reducir el número de componentes o pasos de procesamiento empleados para obtener una representación de dos canales alternativa de la señal de audio de M canales.
- 20 El primer canal de la señal de mezcla descendente puede haber sido formado, por ejemplo, en un lado del codificador, como una combinación lineal del primer grupo de uno o más canales. De manera similar, el segundo canal de la señal de mezcla descendente puede haber sido formado, por ejemplo, en un lado del codificador, como una combinación lineal del segundo grupo de uno o más canales.
- 25 Los canales de la señal de audio de M canales pueden formar, por ejemplo, un subconjunto de un número mayor de canales que representan juntos un campo de sonido.
- 30 Se apreciará que dado que ambos del tercer y cuarto grupos comprenden al menos un canal del primer grupo, la división proporcionada por el tercer y cuarto grupos es diferente de la división proporcionada por el primer y segundo grupos.
- 35 La señal descorrelacionada sirve para aumentar la dimensionalidad del contenido de audio de la señal de mezcla descendente, que se percibe por un oyente. Generar la señal descorrelacionada puede incluir por ejemplo aplicar un filtro lineal a uno o más canales de la señal de mezcla descendente.
- 40 La formación de la señal de salida puede incluir, por ejemplo, aplicar al menos algunos de los coeficientes de mezcla a los canales de la señal de mezcla descendente, y al menos algunos de los coeficientes de mezcla a uno o más canales de la señal descorrelacionada.
- 45 En un ejemplo de realización, los metadatos recibidos pueden incluir los parámetros de mezcla ascendente, y los coeficientes de mezcla se pueden determinar procesando los parámetros de mezcla ascendente, por ejemplo, realizando operaciones matemáticas (por ejemplo, incluyendo operaciones aritméticas) sobre los parámetros de mezcla ascendente. Los parámetros de mezcla ascendente típicamente ya están determinados en el lado del codificador y se proporcionan junto con la señal de mezcla descendente para la reconstrucción paramétrica de la señal de audio de M canales en el lado del decodificador. Los parámetros de mezcla ascendente transportan información acerca de la señal de audio de M canales que puede emplear para proporcionar la señal de salida en base a la señal de mezcla descendente. Determinar, en el lado del decodificador, los coeficientes de mezcla en base a los parámetros de mezcla ascendente, reduce la necesidad de que se generen metadatos adicionales en el lado del codificador y permite una reducción de los datos transmitidos desde el lado del codificador.
- 50 En un ejemplo de realización, los metadatos recibidos pueden incluir parámetros de mezcla distintos de los parámetros de mezcla ascendente. En el presente ejemplo de realización, los coeficientes de mezcla se pueden determinar en base a los metadatos recibidos y, por ello, en base a los parámetros de mezcla. Los parámetros de mezcla se pueden determinar ya en el lado del codificador y transmitir al lado del decodificador para facilitar la determinación de los coeficientes de mezcla. Además, el uso de parámetros de mezcla para determinar los
- 55

- coeficientes de mezcla permite el control de los coeficientes de mezcla desde el lado del codificador. Dado que la señal de audio de M canales original está disponible en el lado del codificador, los parámetros de mezcla se pueden sintonizar, por ejemplo, en el lado del codificador para aumentar la fidelidad de la señal de salida de dos canales como una representación de dos canales de la señal de audio de M canales. Los parámetros de mezcla pueden ser, por ejemplo, los coeficientes de mezcla en sí mismos, o los parámetros de mezcla pueden proporcionar una representación más compacta de los coeficientes de mezcla. Los coeficientes de mezcla se pueden determinar, por ejemplo, procesando los parámetros de mezcla, por ejemplo, según una regla predefinida. Los parámetros de mezcla pueden incluir, por ejemplo, tres parámetros asignables independientemente.
- En un ejemplo de realización, los coeficientes de mezcla se pueden determinar independientemente de cualquier valor de los parámetros de mezcla ascendente, lo que permite la sintonización de los coeficientes de mezcla independientemente de los parámetros de mezcla ascendente, y permite aumentar la fidelidad de la señal de salida de dos canales como una representación de dos canales de la señal de audio de M canales.
- En un ejemplo de realización, puede mantenerse que  $M = 5$ , es decir, la señal de audio de M canales puede ser una señal de audio de cinco canales. El método de decodificación de audio del presente ejemplo de realización se puede emplear, por ejemplo, para los cinco canales regulares de uno de los formatos de audio 5.1 actualmente establecidos, o para cinco canales en el lado izquierdo o derecho en una señal de audio multicanal 11.1. Alternativamente, puede mantener que  $M = 4$ , o  $M \geq 6$ .
- En un ejemplo de realización, cada ganancia que controla una contribución desde un canal de la señal de audio de M canales a una de las combinaciones lineales, a la que corresponden los canales de la señal de mezcla descendente correspondiente, puede coincidir con una ganancia que controla una contribución del canal de la señal de audio de M canales a una de las combinaciones lineales aproximadas por los canales de la señal de salida. El hecho de que estas ganancias coincidan en el presente ejemplo de realización permite simplificar la provisión de la señal de salida en base a la señal de mezcla descendente. En particular, es posible reducir el número de canales descorrelacionados empleados para aproximar las combinaciones lineales del tercer y cuarto grupos en base a la señal de mezcla descendente.
- Se pueden emplear, por ejemplo, diferentes ganancias para diferentes canales de la señal de audio de M canales.
- En un primer ejemplo, todas las ganancias pueden tener el valor 1. En el primer ejemplo, el primer y segundo canales de la señal de mezcla descendente pueden corresponder a sumas no ponderadas del primer y segundo grupos, respectivamente, y el primer y segundo canales de la señal de salida pueden aproximar a sumas no ponderadas del tercer y cuarto conjuntos, respectivamente.
- En un segundo ejemplo, al menos algunas de las ganancias pueden tener valores diferentes de 1. En el segundo ejemplo, el primer y segundo canales de la señal de mezcla descendente pueden corresponder a sumas ponderadas del primer y segundo grupos, respectivamente, y el primer y segundo canales de la señal de salida pueden aproximar las sumas ponderadas del primer y segundo conjuntos, respectivamente.
- En un ejemplo de realización, el método de decodificación puede comprender además: recibir un flujo de bits que representa la señal de mezcla descendente y los metadatos; y extraer, del flujo de bits, la señal de mezcla descendente y la parte recibida de los metadatos. En otras palabras, los metadatos recibidos empleados para determinar los coeficientes de mezcla se pueden haber extraído primero del flujo de bits. Todos los metadatos, incluyendo los parámetros de mezcla ascendente, se pueden extraer, por ejemplo, del flujo de bits. En un ejemplo alternativo, solamente los metadatos necesarios para determinar los coeficientes de mezcla se pueden extraer del flujo de bits, y la extracción de metadatos adicionales, por ejemplo, se puede inhibir.
- En un ejemplo de realización, la señal descorrelacionada puede ser una señal de un único canal y la señal de salida se puede formar incluyendo no más de un canal de señal descorrelacionada en la combinación lineal de la señal de mezcla descendente y la señal descorrelacionada, es decir, en la combinación lineal a partir de la cual se obtiene la señal de salida. Los inventores se han dado cuenta que no hay necesidad de reconstruir la señal de audio de M canales con el fin de proporcionar la señal de salida de dos canales, y que dado que la señal de audio de M canales completa no necesita ser reconstruida, se puede reducir el número de canales de señal descorrelacionada.
- En un ejemplo de realización, los coeficientes de mezcla se pueden determinar de manera que los dos canales de la señal de salida reciban contribuciones de igual magnitud (por ejemplo, igual amplitud). Las contribuciones de la señal descorrelacionada al canal respectivo de la señal de salida pueden tener signos opuestos. En otras palabras, los coeficientes de mezcla se pueden determinar de manera que una suma de un coeficiente de mezcla que controla una contribución de un canal de la señal descorrelacionada al primer canal de la señal de salida, y un coeficiente de mezcla que controla una contribución del mismo canal de la señal descorrelacionada al segundo canal de la señal de salida, tenga el valor 0.
- En el presente ejemplo de realización, la cantidad (por ejemplo, la amplitud) de contenido de audio que se origina a partir de una señal descorrelacionada (es decir, el contenido de audio para aumentar la dimensionalidad de la señal de mezcla descendente) puede ser, por ejemplo, igual en ambos canales de la señal de salida.

En un ejemplo de realización, la formación de la señal de salida puede equivaler a una proyección de tres canales a dos canales, es decir, una proyección de los dos canales de la señal de mezcla descendente y un canal de señal descorrelacionada a los dos canales de la señal de salida. Por ejemplo, la señal de salida se puede obtener directamente como una combinación lineal de la señal de mezcla descendente y la señal descorrelacionada sin reconstruir primero los M canales completos de la señal de audio de M canales.

En un ejemplo de realización, los coeficientes de mezcla se pueden determinar de manera que una suma de un coeficiente de mezcla que controla una contribución del primer canal de la señal de mezcla descendente al primer canal de la señal de salida, y un coeficiente de mezcla que controla una contribución del primer canal de la señal de mezcla descendente al segundo canal de la señal de salida, tenga el valor uno. En particular, uno de los coeficientes de mezcla es derivable a partir de los parámetros de mezcla ascendente (por ejemplo, enviados como un valor explícito u obtenible a partir de los parámetros de mezcla ascendente después de realizar cálculos en una representación compacta, como se explica en otras secciones de esta descripción) y el otro se puede calcular fácilmente requiriendo que la suma de ambos coeficientes de mezcla sea igual a uno.

Además, o alternativamente, los coeficientes de mezcla se pueden determinar de manera que una suma de un coeficiente de mezcla que controla una contribución desde el segundo canal de la señal de mezcla descendente al primer canal de la señal de salida, y un coeficiente de mezcla que controla una contribución del segundo canal de la señal de mezcla descendente al segundo canal de la señal de salida, tenga el valor uno.

En un ejemplo de realización, el primer grupo puede consistir en dos o tres canales. Un canal de la señal de mezcla descendente correspondiente a una combinación lineal de dos o tres canales, en lugar de corresponder a una combinación lineal de cuatro o más canales, puede aumentar la fidelidad de la señal de audio de M canales a medida que se reconstruye mediante un decodificador que realiza una reconstrucción paramétrica de todos de los M canales. El método de decodificación del presente ejemplo de realización puede ser compatible con tal formato de codificación.

En un ejemplo de realización, la señal de audio de M canales puede comprender tres canales que representan diferentes direcciones horizontales en un entorno de reproducción para la señal de audio de M canales, y dos canales que representan direcciones separadas verticalmente de las de los tres canales en el entorno de reproducción. En otras palabras, la señal de audio de M canales puede comprender tres canales destinados a reproducción mediante fuentes de audio situadas sustancialmente a la misma altura que un oyente (o una oreja de un oyente) y/o propagación sustancialmente horizontal, y dos canales destinados a la reproducción mediante fuentes de audio situadas a otras alturas y/o propagación (sustancialmente) no horizontalmente. Los dos canales pueden representar, por ejemplo, direcciones elevadas.

En un ejemplo de realización, el primer grupo puede consistir en los tres canales que representan diferentes direcciones horizontales en un entorno de reproducción para la señal de audio de M canales, y el segundo grupo puede consistir en los dos canales que representan direcciones separadas verticalmente de los de los tres canales en el entorno de reproducción. La división vertical de la señal de audio de M canales proporcionada por el primer y segundo grupos en el presente ejemplo de realización puede aumentar la fidelidad de la señal de audio de M canales a medida que se reconstruye por un decodificador que realiza la reconstrucción paramétrica de todos de los M canales, por ejemplo, en casos donde la dimensión vertical es importante para la impresión general del campo de sonido representado por la señal de audio de M canales. El método de decodificación del presente ejemplo de realización puede ser compatible con un formato de codificación que proporciona esta división vertical.

En un ejemplo de realización, uno del tercer y cuarto grupos puede comprender ambos de los dos canales que representan direcciones separadas verticalmente de las de los tres canales en el entorno de reproducción. Alternativamente, cada uno del tercer y cuarto grupos puede comprender uno de los dos canales que representan direcciones separadas verticalmente de las de los tres canales en el entorno de reproducción, es decir, el tercer y cuarto grupos pueden comprender uno de estos dos canales.

En un ejemplo de realización, la señal descorrelacionada se puede obtener procesando una combinación lineal de los canales de la señal de mezcla descendente, por ejemplo, incluyendo la aplicación de un filtro lineal a la combinación lineal de los canales de los canales de señal de mezcla descendente. Alternativamente, la señal descorrelacionada se puede obtener en base a no más de uno de los canales de la señal de mezcla descendente, por ejemplo, procesando un canal de la señal de mezcla descendente (por ejemplo, incluyendo la aplicación de un filtro lineal). Si, por ejemplo, el segundo grupo de canales consiste en un único canal y el segundo canal de la señal de mezcla descendente corresponde a este único canal, entonces la señal descorrelacionada se puede obtener, por ejemplo, procesando solamente el primer canal de la señal de mezcla descendente.

En un ejemplo de realización, el primer grupo puede consistir en N canales, donde  $N \geq 3$ , y el primer grupo puede ser reconstruible como una combinación lineal del primer canal de la señal de mezcla descendente y una señal descorrelacionada de (N - 1) canales aplicando coeficientes de mezcla ascendente de un primer tipo, a los que se hace referencia en la presente memoria como coeficientes de mezcla ascendente seca, al primer el canal de la señal de mezcla descendente y coeficientes de mezcla ascendente de un segundo tipo, a los que se hace referencia en la presente memoria como coeficientes de mezcla ascendente húmeda, a los canales de la señal descorrelacionada de

(N - 1) canales. En el presente ejemplo de realización, los metadatos recibidos pueden incluir parámetros de mezcla ascendente de un primer tipo, a los que se hace referencia en la presente memoria como parámetros de mezcla ascendente seca, y parámetros de mezcla ascendente de un segundo tipo, a los que se hace referencia en la presente memoria como parámetros de mezcla ascendente húmeda. La determinación de los coeficientes de mezcla puede comprender: determinar, en base a los parámetros de mezcla ascendente seca, los coeficientes de mezcla ascendente seca; poblar una matriz intermedia que tenga más elementos que el número de parámetros de mezcla ascendente húmeda recibidos, en base a los parámetros de mezcla ascendente húmeda recibidos y saber que la matriz intermedia pertenece a una clase de matriz predefinida; obtener los coeficientes de mezcla ascendente húmeda multiplicando la matriz intermedia por una matriz predefinida, en donde los coeficientes de mezcla ascendente húmeda corresponden a la matriz resultante de la multiplicación e incluye más coeficientes que el número de elementos en la matriz intermedia; y procesar los coeficientes de mezcla ascendente húmeda y seca.

En el presente ejemplo de realización, el número de coeficientes de mezcla ascendente húmeda para reconstruir el primer grupo de canales es mayor que el número de parámetros de mezcla ascendente húmeda recibidos. Explotando el conocimiento de la matriz predefinida y la clase de matriz predefinida para obtener los coeficientes de mezcla ascendente húmeda a partir de los parámetros de mezcla ascendente húmeda recibidos, se puede reducir la cantidad de información necesaria para la reconstrucción paramétrica del primer grupo de canales, permitiendo una reducción de la cantidad de metadatos transmitidos junto con la señal de mezcla descendente desde un lado del codificador. Reduciendo la cantidad de datos necesarios para la reconstrucción paramétrica, se puede reducir el ancho de banda requerido para la transmisión de una representación paramétrica de la señal de audio de M canales, y/o el tamaño de memoria requerido para almacenar tal representación.

La señal descorrelacionada de (N - 1) canales se puede generar en base al primer canal de la señal de mezcla descendente y sirve para aumentar la dimensionalidad del contenido del primer grupo de canales reconstruido, como se percibe por un oyente.

La clase de matriz predefinida se puede asociar con propiedades conocidas de al menos algunos elementos de la matriz que son válidos para todas las matrices de la clase, tales como ciertas relaciones entre algunos de los elementos de la matriz, o algunos elementos de la matriz que son cero. El conocimiento de estas propiedades permite poblar la matriz intermedia en base a menos parámetros de mezcla ascendente húmeda que el número total de elementos de la matriz en la matriz intermedia. El lado del decodificador tiene conocimiento al menos de las propiedades de, y las relaciones entre, los elementos que necesita para calcular todos los elementos de la matriz sobre la base de menos parámetros de mezcla húmeda.

Cómo determinar y emplear la matriz predefinida y la clase de matriz predefinida se describe con más detalle en la página 16, línea 15 a página 20, línea 2 en la solicitud de patente provisional de EE.UU. nº 61/974544; primer inventor nombrado: Lars Villemoes; fecha de presentación: 3 de abril de 2014. Véase en particular la ecuación (9) en la misma para ejemplos de la matriz predefinida.

En un ejemplo de realización, los metadatos recibidos pueden incluir  $N(N - 1)/2$  parámetros de mezcla ascendente húmeda. En el presente ejemplo de realización, poblar la matriz intermedia puede incluir obtener valores para  $(N - 1)^2$  elementos de la matriz en base a los  $N(N - 1)/2$  parámetros de mezcla ascendente húmeda recibidos y saber que la matriz intermedia pertenece a la clase de matriz predefinida. Esto puede incluir la inserción de los valores de los parámetros de mezcla ascendente húmeda inmediatamente como elementos de la matriz, o el procesamiento de los parámetros de mezcla ascendente húmeda de una manera adecuada para derivar valores para los elementos de la matriz. En el presente ejemplo de realización, la matriz predefinida puede incluir  $N(N - 1)$  elementos, y el conjunto de coeficientes de mezcla ascendente húmeda puede incluir  $N(N - 1)$  coeficientes. Por ejemplo, los metadatos recibidos pueden incluir no más de  $N(N - 1)/2$  parámetros de mezcla ascendente húmeda asignables independientemente y/o el número de parámetros de mezcla ascendente húmeda puede ser no más de la mitad del número de coeficientes de mezcla ascendente húmeda para reconstruir el primer grupo de canales.

En un ejemplo de realización, los metadatos recibidos pueden incluir (N - 1) parámetros de mezcla ascendente seca. En el presente ejemplo de realización, los coeficientes de mezcla ascendente seca pueden incluir N coeficientes, y los coeficientes de mezcla ascendente seca se pueden determinar en base a los (N - 1) parámetros de mezcla ascendente seca recibidos y en base a una relación predefinida entre los coeficientes de mezcla ascendente seca. Por ejemplo, los metadatos recibidos pueden incluir no más de (N - 1) parámetros de mezcla ascendente seca asignables independientemente.

En un ejemplo de realización, la clase de matriz predefinida puede ser una de: matrices triangulares inferiores o superiores, en donde las propiedades conocidas de todas las matrices en la clase incluyen elementos de la matriz predefinidos que son cero; matrices simétricas, en donde las propiedades conocidas de todas las matrices en la clase incluyen elementos de la matriz predefinidos (a cada lado de la diagonal principal) que son iguales; y productos de una matriz ortogonal y una matriz diagonal, en donde las propiedades conocidas de todas las matrices en la clase incluyen relaciones conocidas entre elementos de la matriz predefinidos. En otras palabras, la clase de matriz predefinida puede ser la clase de matrices triangulares inferiores, la clase de matrices triangulares superiores, la clase de matrices simétricas o la clase de productos de una matriz ortogonal y una matriz diagonal. Una propiedad

común de cada una de las clases anteriores es que su dimensionalidad es menor que el número total de elementos de la matriz.

En un ejemplo de realización, el método de decodificación puede comprender además: recibir señalización que indica uno (seleccionado) de al menos dos formatos de codificación de la señal de audio de M canales, los formatos de codificación correspondientes a las diferentes divisiones respectivas de los canales de la señal de audio de M canales en el primer y segundo grupos respectivos asociados con los canales de la señal de mezcla descendente. En el presente ejemplo de realización, el tercer y cuarto grupos pueden estar predefinidos, y los coeficientes de mezcla se pueden determinar de manera que una única división de la señal de audio de M canales en el tercer y cuarto grupos de canales, aproximada por los canales de la señal de salida, se mantenga para (es decir, es común a) los al menos dos formatos de codificación.

En el presente ejemplo de realización, la señal descorrelacionada se puede determinar, por ejemplo, en base al formato de codificación indicado y a al menos un canal de la señal de mezcla descendente.

En el presente ejemplo de realización, los al menos dos formatos de codificación diferentes pueden haber sido empleados en el lado del codificador cuando se determina la señal de mezcla descendente y los metadatos, y el método de decodificación puede manejar diferencias entre los formatos de codificación ajustando los coeficientes de mezcla, y opcionalmente también la señal descorrelacionada. En caso de que se detecte un cambio de un primer formato de codificación a un segundo formato de codificación, el método de decodificación puede incluir, por ejemplo, realizar interpolación de los parámetros de mezcla asociados con el primer formato de codificación a los parámetros de mezcla asociados con el segundo formato de codificación.

En un ejemplo de realización, el método de decodificación puede comprender además: pasar la señal de mezcla descendente a través como la señal de salida, en respuesta a la señalización que indica un formato de codificación particular. En el presente ejemplo de realización, el formato de codificación particular puede corresponder a una división de los canales de la señal de audio de M canales que coincide con una división que definen el tercer y cuarto grupos. En el presente ejemplo de realización, la división proporcionada por los canales de la señal de mezcla descendente puede coincidir con la división a ser proporcionada por los canales de la señal de salida, y puede no ser necesario procesar la señal de mezcla descendente. La señal de mezcla descendente, por lo tanto, se puede pasar a través como la señal de salida.

En un ejemplo de realización, el método de decodificación puede comprender: suprimir la contribución de la señal descorrelacionada a la señal de salida, en respuesta a la señalización que indica un formato de codificación particular. En el presente ejemplo de realización, el formato de codificación particular puede corresponder a una división de los canales de la señal de audio de M canales que coincide con una división que definen el tercer y cuarto grupos. En el presente ejemplo de realización, la división proporcionada por los canales de la señal de mezcla descendente puede coincidir con la división a ser proporcionada por los canales de la señal de salida, y puede no ser necesaria para una descorrelación.

En un ejemplo de realización, en un primer formato de codificación, el primer grupo puede consistir en tres canales que representan diferentes direcciones horizontales en un entorno de reproducción para la señal de audio de M canales, y el segundo grupo de canales pueden consistir en dos canales que representan direcciones separadas verticalmente de las de los tres canales en el entorno de reproducción. En un segundo formato de codificación, cada uno del primer y segundo grupos puede comprender uno de los dos canales.

Según ejemplos de realizaciones, se proporciona un sistema de decodificación de audio que comprende una sección de decodificación configurada para recibir una señal de mezcla descendente de dos canales. La señal de mezcla descendente está asociada con metadatos que comprenden parámetros de mezcla ascendente para la reconstrucción paramétrica de una señal de audio de M canales en base a la señal de mezcla descendente, donde  $M \geq 4$ . Un primer canal de la señal de mezcla descendente corresponde a una combinación lineal de un primer grupo de uno o más canales de la señal de audio de M canales, y un segundo canal de la señal de mezcla descendente corresponde a una combinación lineal de un segundo grupo de uno o más canales de la señal de audio de M canales. El primer y segundo grupo constituyen una división de los M canales de la señal de audio de M canales. La sección de decodificación está configurada además para: recibir al menos una parte de los metadatos; y proporcionar una señal de salida de dos canales en base a la señal de mezcla descendente y los metadatos recibidos. La sección de decodificación comprende una sección de descorrelación configurada para recibir al menos un canal de la señal de mezcla descendente y emitir, en base al mismo, una señal descorrelacionada. La sección de decodificación comprende además una sección de mezcla configurada para: determinar un conjunto de coeficientes de mezcla en base a los metadatos recibidos, y formar la señal de salida como una combinación lineal de la señal de mezcla descendente y la señal descorrelacionada según los coeficientes de mezcla. La sección de mezcla está configurada para determinar los coeficientes de mezcla de manera que un primer canal de la señal de salida se aproxime a una combinación lineal de un tercer grupo de uno o más canales de la señal de audio de M canales, y de manera que un segundo canal de la señal de salida se aproxime a una combinación lineal de un cuarto grupo de uno o más canales de la señal de audio de M canales. La sección de mezcla está configurada además para determinar los coeficientes de mezcla de manera que el tercer y cuarto grupos constituyan una división de los M canales de la

señal de audio de M canales, y de manera que ambos del tercer y cuarto grupos comprendan al menos un canal del primer grupo.

En un ejemplo de realización, el sistema de decodificación de audio puede comprender además una sección de decodificación adicional configurada para recibir una señal de mezcla descendente de dos canales adicional. La señal de mezcla descendente adicional se puede asociar con metadatos adicionales que comprenden parámetros de mezcla ascendente adicionales para la reconstrucción paramétrica de una señal de audio de M canales adicional en base a la señal de mezcla descendente adicional. Un primer canal de la señal de mezcla descendente adicional puede corresponder a una combinación lineal de un primer grupo de uno o más canales de la señal de audio de M canales adicional, y un segundo canal de la señal de mezcla descendente adicional puede corresponder a una combinación lineal de un segundo grupo de uno o más canales de la señal de audio de M canales adicional. El primer y segundo grupos de canales de la señal de audio de M canales adicional pueden constituir una división de los M canales de la señal de audio de M canales adicional. La sección de decodificación adicional se puede configurar además para: recibir al menos una parte de los metadatos adicionales; y proporcionar una señal de salida de dos canales adicional en base a la señal de mezcla descendente adicional y los metadatos recibidos adicionales. La sección de decodificación adicional puede comprender una sección de descorrelación adicional configurada para recibir al menos un canal de la señal de mezcla descendente adicional y para emitir, en base al mismo, una señal descorrelacionada adicional. La sección de decodificación adicional puede comprender además una sección de mezcla adicional configurada para: determinar un conjunto de coeficientes de mezcla adicionales en base a los metadatos adicionales recibidos, y formar la señal de salida adicional como una combinación lineal de la señal de mezcla descendente adicional y la señal descorrelacionada adicional según los coeficientes de mezcla adicionales. La sección de mezcla adicional se puede configurar para determinar los coeficientes de mezcla adicionales de manera que un primer canal de la señal de salida adicional se aproxime a una combinación lineal de un tercer grupo de uno o más canales de la señal de audio de M canales adicional, y de manera que un segundo canal de la señal de salida adicional se aproxime a una combinación lineal de un cuarto grupo de uno o más canales de la señal de audio de M canales adicional. La sección de mezcla adicional se puede configurar además para determinar los coeficientes de mezcla adicionales de manera que el tercer y cuarto grupos de canales de la señal de audio de M canales adicional constituyan una división de los M canales de la señal de audio de M canales adicional, y de manera que ambos del tercer y cuarto grupos de canales de la señal de audio de M canales adicional comprenda al menos un canal del primer grupo de canales de la señal de audio de M canales adicional.

En el presente ejemplo de realización, la sección de decodificación adicional, la sección de descorrelación adicional y la sección de mezcla adicional pueden ser, por ejemplo, funcionalmente equivalentes a (o configuradas de manera análoga como) la sección de decodificación, la sección de descorrelación y la sección de mezcla, respectivamente. Alternativamente, al menos una de la sección de decodificación adicional, la sección de descorrelación adicional y la sección de mezcla adicional, por ejemplo, se pueden configurar para realizar al menos un tipo diferente de cálculo y/o interpolación que la realizada por la sección correspondiente de la sección de decodificación, la sección de descorrelación y la sección de mezcla.

En el presente ejemplo de realización, la sección de decodificación adicional, la sección de descorrelación adicional y la sección de mezcla adicional pueden ser, por ejemplo, operables independientemente de la sección de decodificación, la sección de descorrelación y la sección de mezcla.

En un ejemplo de realización, el sistema de decodificación puede comprender además un demultiplexor configurado para extraer, de un flujo de bits: la señal de mezcla descendente, la al menos una parte de los metadatos, y un canal de audio codificado de manera discreta. El sistema de decodificación puede comprender además una sección de decodificación de un único canal operable para decodificar el canal de audio codificado de manera discreta. El canal de audio codificado de manera discreta se puede codificar, por ejemplo, en el flujo de bits usando un códec de audio perceptivo, tal como Dolby Digital o MPEG AAC, y la sección de decodificación de un único canal puede comprender, por ejemplo, un decodificador central para decodificar el canal de audio codificado de manera discreta. La sección de decodificación de un único canal puede ser, por ejemplo, operable para decodificar el canal de audio codificado de manera discreta independientemente de la sección de decodificación.

Según ejemplos de realizaciones, se proporciona un producto de programa de ordenador que comprende un medio legible por ordenador con instrucciones para realizar cualquiera de los métodos del primer aspecto.

Según ejemplos de realizaciones del sistema de decodificación de audio, del método y del producto de programa de ordenador del primer aspecto, descrito anteriormente, la señal de salida puede ser una señal de K canales, donde  $2 \leq K < M$ , en lugar de una señal de dos canales, y los K canales de la señal de salida pueden corresponder a una división de la señal de audio de M canales en K grupos, en lugar de dos canales de la señal de salida correspondientes a una división de la señal de M canales en dos grupos.

Más específicamente, según ejemplos de realizaciones, se proporciona un método de codificación de audio que comprende recibir una señal de mezcla descendente de dos canales. La señal de mezcla descendente se asocia con metadatos que comprenden parámetros de mezcla ascendente para la reconstrucción paramétrica de una señal de audio de M canales en base a la señal de mezcla descendente, donde  $M \geq 4$ . Un primer canal de la señal de mezcla descendente corresponde a una combinación lineal de un primer grupo de uno o más canales de la señal de

5 audio de M canales, y un segundo canal de la señal de mezcla descendente corresponde a una combinación lineal de un segundo grupo de uno o más canales de la señal de audio de M canales. El primer y segundo grupos constituyen una división de los M canales de la señal de audio de M canales. El método de decodificación de audio puede comprender además: recibir al menos una parte de los metadatos; generar una señal descorrelacionada en base a al menos un canal de la señal de mezcla descendente; determinar un conjunto de coeficientes de mezcla en base a los metadatos recibidos; y formar una señal de salida de K canales como una combinación lineal de la señal de mezcla descendente y la señal descorrelacionada según los coeficientes de mezcla, en donde  $2 \leq K < M$ . Los coeficientes de mezcla se pueden determinar de manera que cada uno de los K canales de la señal de salida se aproxime a una combinación lineal de un grupo de uno o más canales de la señal de audio de M canales (y cada uno de los K canales de la señal de salida corresponde, por lo tanto, a un grupo de uno o más canales de la señal de audio de M canales), los grupos correspondientes a los canales respectivos de la señal de salida constituyen una división de los M canales de la señal de audio de M canales en K grupos de uno o más canales; y al menos dos de los K grupos comprenden al menos un canal del primer grupo.

10 La señal de audio de M canales se ha codificado como la señal de mezcla descendente de dos canales y los parámetros de mezcla ascendente para la reconstrucción paramétrica de la señal de audio de M canales. Cuando se codifica la señal de audio de M canales en un lado del codificador, se puede elegir el formato de codificación por ejemplo para facilitar la reconstrucción de la señal de audio de M canales a partir de la señal de mezcla descendente, para mejorar la fidelidad de la señal de audio de M canales como reconstruida a partir de la señal de mezcla descendente, y/o para mejorar la eficiencia de codificación de la señal de mezcla descendente. Esta elección del formato de codificación se puede realizar seleccionando el primer y segundo grupos y formando los canales de las señales de mezcla descendente como combinaciones lineales respectivas de los canales en los grupos respectivos.

15 Los inventores se han dado cuenta de que aunque el formato de codificación elegido puede facilitar la reconstrucción de la señal de audio de M canales a partir de la señal de mezcla descendente, la señal de mezcla descendente puede no ser adecuada en sí misma para la reproducción usando una configuración particular de K altavoces. La señal de salida de K canales, correspondiente a una división de la señal de audio de M canales en los K grupos, puede ser más adecuada para un ajuste de reproducción de K canales particular que la señal de mezcla descendente. Proporcionar la señal de salida en base a la señal de mezcla descendente y los metadatos recibidos puede mejorar, por lo tanto, la calidad de reproducción de K canales como se percibe por un oyente, y/o mejorar la fidelidad de la reproducción de K canales a un campo de sonido representado por la señal de audio de M canales.

20 Los inventores se han dado cuenta además de que, en lugar de primero reconstruir la señal de audio de M canales a partir de la señal de mezcla descendente y entonces generar la representación de K canales de la señal de audio de M canales (por ejemplo, mediante mezcla aditiva), la representación de K canales proporcionada por la señal de salida se puede generar de manera más eficiente a partir de la señal de mezcla descendente y los metadatos recibidos explotando el hecho de que algunos canales de la señal de audio de M canales se agrupan juntos de manera similar en la representación de dos canales proporcionada por la señal de mezcla descendente y la representación de K canales a ser proporcionada. Formar la señal de salida como una combinación lineal de la señal de mezcla descendente y la señal descorrelacionada, por ejemplo, puede reducir la complejidad de cálculo en el lado del decodificador y/o reducir el número de componentes o pasos de procesamiento empleados para obtener una representación de K canales de la señal de audio de M canales.

25 Mediante los K grupos que constituyen una división de los canales de la señal de audio de M canales se entiende que los K grupos están separados y juntos incluyen todos los canales de la señal de audio de M canales.

30 Formar la señal de salida de K canales puede incluir, por ejemplo, aplicar al menos algunos de los coeficientes de mezcla a los canales de la señal de mezcla descendente, y al menos algunos de los coeficientes de mezcla a uno o más canales de la señal descorrelacionada.

35 El primer y segundo canales de la señal de mezcla descendente pueden corresponder, por ejemplo, a sumas (ponderadas o no ponderadas) de los canales en el primer y segundo grupos de uno o más canales, respectivamente.

40 Los K canales de la señal de salida, por ejemplo, pueden aproximar las sumas (ponderadas o no ponderadas) de los canales en los K grupos de uno o más canales, respectivamente.

45 En algunos ejemplos de realizaciones,  $K = 2$ ,  $K = 3$  o  $K = 4$ .

En algunos ejemplos de realizaciones,  $M = 5$ , o  $M = 6$ .

50 En un ejemplo de realización, la señal descorrelacionada puede ser una señal de dos canales, y la señal de salida se puede formar incluyendo no más de dos canales de señal descorrelacionada en la combinación lineal de la señal de mezcla descendente y la señal descorrelacionada, es decir, en la combinación lineal a partir de la cual se obtiene la señal de salida. Los inventores se han dado cuenta de que no es necesario reconstruir la señal de audio de M canales con el fin de proporcionar la señal de salida de dos canales, y que dado que la señal de audio de M canales completa no necesita ser reconstruida, se puede reducir el número de canales de señal descorrelacionada.

En un ejemplo de realización,  $K = 3$  y formar la señal de salida puede equivaler a una proyección de cuatro canales a tres canales, es decir, una proyección de los dos canales de la señal de mezcla descendente y dos canales de señal descorrelacionada a los tres canales de la señal de salida. Por ejemplo, la señal de salida se puede obtener directamente como una combinación lineal de la señal de mezcla descendente y la señal descorrelacionada sin reconstruir primero los  $M$  canales completos de la señal de audio de  $M$  canales.

En un ejemplo de realización, los coeficientes de mezcla se pueden determinar de manera que un par de canales de la señal de salida reciban contribuciones de igual magnitud (por ejemplo, igual amplitud) de un canal de la señal descorrelacionada. Las contribuciones de este canal de la señal descorrelacionada al canal respectivo del par pueden tener signos opuestos. En otras palabras, los coeficientes de mezcla se pueden determinar de manera que una suma de un coeficiente de mezcla que controla una contribución de un canal de la señal descorrelacionada a un (por ejemplo, un primer) canal de la señal de salida, y un coeficiente de mezcla que controla una contribución del mismo canal de la señal descorrelacionada a otro (por ejemplo, un segundo) canal de la señal de salida, tenga el valor 0. La señal de salida de  $K$  canales puede incluir, por ejemplo, uno o más canales que no reciben ninguna contribución de este canal en particular de la señal descorrelacionada.

En un ejemplo de realización, los coeficientes de mezcla se pueden determinar de manera que una suma de un coeficiente de mezcla que controla una contribución del primer canal de la señal de mezcla descendente a un (por ejemplo, un primer) canal de la señal de salida, y un coeficiente de mezcla que controla una contribución del primer canal de la señal de mezcla descendente a otro (por ejemplo, un segundo) canal de la señal de salida, tenga el valor 1. En particular, uno de los coeficientes de mezcla, por ejemplo, puede ser derivable a partir de los parámetros de mezcla (por ejemplo, enviado como un valor explícito u obtenible a partir de los parámetros de mezcla ascendente después de realizar cálculos en una representación compacta, como se explica en otras secciones de esta descripción) y el otro se puede calcular fácilmente requiriendo que la suma de ambos coeficientes de mezcla sea igual a uno. La señal de salida de  $K$  canales puede incluir, por ejemplo, uno o más canales que no reciben ninguna contribución del primer canal de la señal de mezcla descendente.

En un ejemplo de realización, los coeficientes de mezcla se pueden determinar de manera que una suma de un coeficiente de mezcla que controla una contribución del segundo canal de la señal de mezcla descendente a un (por ejemplo, un primer) canal de la señal de salida, y un coeficiente de mezcla que controla una contribución del segundo canal de la señal de mezcla descendente a otro (por ejemplo, un segundo) canal de la señal de salida, tenga el valor uno. La señal de salida de  $K$  canales puede incluir, por ejemplo, uno o más canales que no reciben ninguna contribución del segundo canal de la señal de mezcla descendente.

En un ejemplo de realización, el método puede comprender recibir señalización que indique uno (uno seleccionado) de al menos dos formatos de codificación de la señal de audio de  $M$  canales. Los formatos de codificación pueden corresponder a diferentes divisiones respectivas de los canales de la señal de audio de  $M$  canales en el primer y segundo grupos respectivos asociados con los canales de la señal de mezcla descendente. Los  $K$  grupos pueden estar predefinidos. Los coeficientes de mezcla se pueden determinar de manera que una única división de la señal de audio de  $M$  canales en los  $K$  grupos de canales, aproximada por los canales de la señal de salida, se mantenga para (es decir, sea común a) los al menos dos formatos de codificación.

En un ejemplo de realización, la señal descorrelacionada puede comprender dos canales. Se puede obtener un primer canal de la señal descorrelacionada en base al primer canal de la señal de mezcla descendente, por ejemplo, procesando no más que el primer canal de la señal de mezcla descendente. Se puede obtener un segundo canal de la señal descorrelacionada en base al segundo canal de la señal de mezcla descendente, por ejemplo, procesando no más que el segundo canal de la señal de mezcla descendente.

## II. Descripción general - lado del codificador

Según un segundo aspecto, los ejemplos de realizaciones proponen sistemas de codificación de audio así como métodos de codificación de audio y productos de programas de ordenador asociados. Los sistemas de codificación, métodos y productos de programas de ordenador propuestos, según el segundo aspecto, generalmente pueden compartir las mismas características y ventajas. Además, las ventajas presentadas anteriormente para las características de los sistemas de decodificación, los métodos y los productos de programas de ordenador, según el primer aspecto, pueden ser generalmente válidas para las características correspondientes de los sistemas de codificación, los métodos y los productos de programas de ordenador según el segundo aspecto.

Según ejemplos de realizaciones, se proporciona un método de codificación de audio que comprende: recibir una señal de audio de  $M$  canales, donde  $M \geq 4$ ; y calcular una señal de mezcla descendente de dos canales en base a la señal de audio de  $M$  canales. Un primer canal de la señal de mezcla descendente se forma como una combinación lineal de un primer grupo de uno o más canales de la señal de audio de  $M$  canales, y un segundo canal de la señal de mezcla descendente se forma como una combinación lineal de un segundo grupo de uno o más canales de la señal de audio de  $M$  canales. El primer y segundo grupos constituyen una división de los  $M$  canales de la señal de audio de  $M$  canales. El método de codificación comprende además: determinar los parámetros de mezcla ascendente para la reconstrucción paramétrica de la señal de audio de  $M$  canales a partir de la señal de mezcla descendente; y determinar los parámetros de mezcla para obtener, en base a la señal de mezcla descendente, una

5 señal de salida de dos canales, en donde un primer canal de la señal de salida se aproxima a una combinación lineal de un tercer grupo de uno o más canales de la señal de audio de M canales, y en donde un segundo canal de la señal de salida se aproxima a una combinación lineal de un cuarto grupo de uno o más canales de la señal de audio de M canales. El tercer y cuarto grupos constituyen una división de los M canales de la señal de audio de M canales, y ambos del tercer y cuarto grupos comprenden al menos un canal del primer grupo. El método de codificación comprende además: emitir la señal de mezcla descendente y los metadatos para el almacenamiento o la transmisión conjunta, en donde los metadatos comprenden los parámetros de mezcla ascendente y los parámetros de mezcla.

10 Los canales de la señal de mezcla descendente corresponden a una división de los M canales de la señal de audio de M canales en el primer y segundo grupos y, por ejemplo, pueden proporcionar una representación de dos canales un poco eficiente de la señal de audio de M canales y/o una representación de dos canales que permita una reconstrucción paramétrica de alta fidelidad de la señal de audio de M canales.

15 Los inventores se han dado cuenta de que aunque la representación de dos canales empleada puede facilitar la reconstrucción de la señal de audio de M canales a partir de la señal de mezcla descendente, la señal de mezcla descendente puede no ser adecuada en sí misma para la reproducción usando una disposición particular de dos altavoces. Los parámetros de mezcla, emitidos junto con la señal de mezcla descendente y los parámetros de mezcla ascendente, permiten obtener la señal de salida de dos canales en base a la señal de mezcla descendente. La señal de salida, correspondiente a una división diferente de la señal de audio de M canales en el tercer y cuarto grupos de canales, puede ser más adecuada para un ajuste de reproducción de dos canales particular que la señal de mezcla descendente. Proporcionar la señal de salida en base a la señal de mezcla descendente y los parámetros de mezcla puede mejorar, por lo tanto, la calidad de reproducción de dos canales como se percibe por un oyente, y/o mejorar la fidelidad de la reproducción de dos canales para un campo de sonido representado por la señal de audio de M canales.

25 El primer canal de la señal de mezcla descendente se puede formar, por ejemplo, como una suma de los canales en el primer grupo, o como una escala de los mismos. En otras palabras, el primer canal de la señal de mezcla descendente se puede formar, por ejemplo, como una suma de los canales (es decir, una suma del contenido de audio de los canales respectivos, por ejemplo, formar mediante la mezcla aditiva sobre una base por muestra o por coeficiente de transformación) en el primer grupo, o como una versión reescalada de tal suma (por ejemplo, obtenida sumando los canales y multiplicando la suma por un factor de reescalado). De manera similar, el segundo canal de la señal de mezcla descendente se puede formar, por ejemplo, como una suma de los canales en el segundo grupo, o como una escala de los mismos. El primer canal de la señal de salida, por ejemplo, puede aproximarse a la suma de los canales del tercer grupo, o una escala de los mismos, y el segundo canal de la señal de salida, por ejemplo, puede aproximarse a la suma de los canales en el cuarto grupo, o una escala de los mismos.

35 Por ejemplo, la señal de audio de M canales puede ser una señal de audio de cinco canales. El método de codificación de audio se puede emplear, por ejemplo, para los cinco canales regulares de uno de los formatos de audio 5.1 establecidos actualmente, o para cinco canales en el lado izquierdo o derecho en una señal de audio multicanal 11.1. Alternativamente, se puede mantener que  $M = 4$ , o  $M \geq 6$ .

40 En un ejemplo de realización, los parámetros de mezcla pueden controlar las contribuciones respectivas de la señal de mezcla descendente y de una señal descorrelacionada a la señal de salida. Al menos algunos de los parámetros de mezcla se pueden determinar minimizando una contribución de la señal descorrelacionada entre tales parámetros de mezcla que hacen que los canales de la señal de salida sean aproximaciones que conservan la covarianza de las combinaciones lineales (o sumas) del primer y segundo grupos de canales, respectivamente. La contribución de la señal descorrelacionada se puede minimizar, por ejemplo, en el sentido de que se minimiza la energía de la señal o la amplitud de esta contribución.

45 La combinación lineal del tercer grupo, que el primer canal de la señal de salida ha de aproximar, y la combinación lineal del cuarto grupo, que el segundo canal de la señal de salida ha de aproximar, pueden corresponder, por ejemplo, a una señal de audio de dos canales que tiene una primera matriz de covarianza. Los canales de la señal de salida que son aproximaciones que conservan la covarianza de las combinaciones lineales del primer y segundo grupos de canales, respectivamente, pueden corresponder, por ejemplo, a que una matriz de covarianza de la señal de salida coincida (o al menos coincida sustancialmente) con la primera matriz de covarianza.

50 Entre las aproximaciones que conservan la covarianza, un tamaño reducido (por ejemplo, energía o amplitud) de la contribución de la señal descorrelacionada puede ser indicativo de un aumento de la fidelidad de la aproximación que se percibe por un oyente durante la reproducción. Emplear los parámetros de mezcla que disminuyen la contribución de la señal descorrelacionada puede mejorar la fidelidad de la señal de salida como una representación de dos canales de la señal de audio de M canales.

55 En un ejemplo de realización, el primer grupo de canales puede consistir en N canales, donde  $N \geq 3$ , y al menos algunos de los parámetros de mezcla ascendente pueden ser adecuados para la reconstrucción paramétrica del primer grupo de canales del primer canal de la señal de mezcla descendente y una señal descorrelacionada de  $(N - 1)$  canales determinada en base al primer canal de la señal de mezcla descendente. En el presente ejemplo de

realización, la determinación de los parámetros de mezcla ascendente puede incluir: determinar un conjunto de coeficientes de mezcla ascendente de un primer tipo, a los que se hace referencia como coeficientes de mezcla ascendente seca, con el fin de definir una correspondencia lineal del primer canal de la señal de mezcla descendente que se aproxima al primer grupo de canales; y determinar una matriz intermedia en base a una diferencia entre una covarianza del primer grupo de canales que se recibe, y una covarianza del primer grupo de canales que se aproxima por la correspondencia lineal del primer canal de la señal de mezcla descendente. Cuando se multiplica por una matriz predefinida, la matriz intermedia puede corresponder a un conjunto de coeficientes de mezcla ascendente de un segundo tipo, a los que se hace referencia como coeficientes de mezcla ascendente húmeda, que definen una correspondencia lineal de la señal descorrelacionada como parte de la reconstrucción paramétrica del primer grupo de canales. El conjunto de coeficientes de mezcla ascendente húmeda puede incluir más coeficientes que el número de elementos en la matriz intermedia. En el presente ejemplo de realización, los parámetros de mezcla ascendente pueden incluir un primer tipo de parámetros de mezcla ascendente, a los que se hace referencia como parámetros de mezcla ascendente seca, a partir de los cuales es derivable el conjunto de coeficientes de mezcla ascendente seca, y un segundo tipo de parámetros de mezcla ascendente, a los que se hace referencia como parámetros de mezcla ascendente húmeda, que definen de manera única la matriz intermedia siempre que la matriz intermedia pertenezca a una clase de matriz predefinida. La matriz intermedia puede tener más elementos que el número de parámetros de mezcla ascendente húmeda.

En el presente ejemplo de realización, una copia de reconstrucción paramétrica del primer grupo de canales en el lado del decodificador incluye, como una contribución, una señal de mezcla ascendente seca formada por la correspondencia lineal del primer canal de la señal de mezcla descendente y, como una contribución adicional, una señal de mezcla ascendente húmeda formada por la correspondencia lineal de la señal descorrelacionada. El conjunto de coeficientes de mezcla ascendente seca define la correspondencia lineal del primer canal de la señal de mezcla descendente y el conjunto de coeficientes de mezcla ascendente húmeda define la correspondencia lineal de la señal descorrelacionada. Emitiendo parámetros de mezcla ascendente húmeda que sean menores que el número de coeficientes de mezcla ascendente húmeda, y a partir de los cuales los coeficientes de mezcla ascendente húmeda son derivables en base a la matriz predefinida y la clase de matriz predefinida, se puede reducir la cantidad de información enviada al lado del decodificador para habilitar la reconstrucción de la señal de audio de M canales. Reduciendo la cantidad de datos necesarios para la reconstrucción paramétrica, se puede reducir el ancho de banda requerido para la transmisión de una representación paramétrica de la señal de audio de M canales, y/o el tamaño de memoria requerido para almacenar tal representación.

La matriz intermedia se puede determinar, por ejemplo, de manera que una covarianza de la señal obtenida por la correspondencia lineal de la señal descorrelacionada complementa la covarianza del primer grupo de canales que se aproxima por la correspondencia lineal del primer canal de la señal de mezcla descendente.

Cómo determinar y emplear la matriz predefinida y la clase de matriz predefinida se describe con más detalle en la página 16, línea 15 a la página 20, línea 2 en la solicitud de patente provisional de EE.UU. N° 61/974544; primer inventor nombrado: Lars Villemoes; fecha de presentación: 3 de abril de 2014. Véase en particular la ecuación (9) en la misma para ejemplos de la matriz predefinida.

En un ejemplo de realización, determinar la matriz intermedia puede incluir determinar la matriz intermedia de manera que una covarianza de la señal obtenida por la correspondencia lineal de la señal descorrelacionada, definida por el conjunto de coeficientes de mezcla ascendente húmeda, se aproxima, o coincide sustancialmente con, la diferencia entre la covarianza del primer grupo de canales que se recibe y la covarianza del primer grupo de canales que se aproxima por la correspondencia lineal del primer canal de la señal de mezcla descendente. En otras palabras, la matriz intermedia se puede determinar de manera que una copia de la reconstrucción del primer grupo de canales, obtenida como una suma de una señal de mezcla ascendente seca formada por la correspondencia lineal del primer canal de la señal de mezcla descendente y una señal de mezcla ascendente húmeda formada por la correspondencia lineal de la señal descorrelacionada por completo, o al menos aproximadamente, reinstaura la covarianza del primer grupo de canales como se recibe.

En un ejemplo de realización, los parámetros de mezcla ascendente húmeda pueden incluir no más de  $N(N - 1)/2$  parámetros de mezcla ascendente húmeda asignables independientemente. En el presente ejemplo de realización, la matriz intermedia puede tener  $(N - 1)^2$  elementos de la matriz y se puede definir de manera única mediante los parámetros de mezcla ascendente húmeda, siempre que la matriz intermedia pertenezca a la clase de matriz predefinida. En el presente ejemplo de realización, el conjunto de coeficientes de mezcla ascendente húmeda puede incluir  $N(N - 1)$  coeficientes.

En un ejemplo de realización, el conjunto de coeficientes de mezcla ascendente seca puede incluir N coeficientes. En el presente ejemplo de realización, los parámetros de mezcla ascendente seca pueden incluir no más de N - 1 parámetros de mezcla ascendente seca, y el conjunto de coeficientes de mezcla ascendente seca puede ser derivable a partir de los N - 1 parámetros de mezcla ascendente seca usando una regla predefinida.

En un ejemplo de realización, el conjunto determinado de coeficientes de mezcla ascendente seca puede definir una correspondencia lineal del primer canal de la señal de mezcla descendente correspondiente a una aproximación de error cuadrático medio mínimo del primer grupo de canales, es decir, entre el conjunto de correspondencias lineales

del primer canal de la señal de mezcla descendente, el conjunto determinado de coeficientes de mezcla ascendente seca puede definir la correspondencia lineal que mejor se aproxima al primer grupo de canales en un sentido de cuadrado medio mínimo.

5 En un ejemplo de realización, el método de codificación puede comprender además seleccionar uno de al menos dos formatos de codificación, en donde los formatos de codificación corresponden a diferentes divisiones respectivas de los canales de la señal de audio de M canales en el primer y segundo grupos respectivos asociados con los canales de la señal de mezcla descendente. El primer y segundo canales de la señal de mezcla descendente se pueden formar como combinaciones lineales de un primer y un segundo grupo de uno o más canales, respectivamente, de la señal de audio de M canales, según el formato de codificación seleccionado. Los parámetros de mezcla ascendente y los parámetros de mezcla se pueden determinar en base al formato de codificación seleccionado. El método de codificación puede comprender además proporcionar una señalización que indique el formato de codificación seleccionado. La señalización se puede emitir, por ejemplo, para almacenamiento y/o transmisión conjunta con la señal de mezcla descendente y los metadatos.

15 La señal de audio de M canales reconstruida en base a la señal de mezcla descendente y los parámetros de mezcla ascendente puede ser una suma de: una señal de mezcla ascendente seca formada aplicando coeficientes de mezcla ascendente seca a la señal de mezcla descendente; y una señal de mezcla ascendente húmeda formada aplicando coeficientes de mezcla ascendente húmeda a una señal descorrelacionada determinada en base a la señal de mezcla descendente. La selección de un formato de codificación se puede hacer, por ejemplo, en base a una diferencia entre una covarianza de la señal de audio de M canales que se recibe y una covarianza de la señal de audio de M canales que se aproxima por la señal de mezcla ascendente seca, para los formatos de codificación respectivos. La selección de un formato de codificación se puede hacer, por ejemplo, en base a los coeficientes de mezcla ascendente húmeda para los formatos de codificación respectivos, por ejemplo, en base a las sumas de cuadrados respectivas de los coeficientes de mezcla ascendente húmeda para los formatos de codificación respectivos. El formato de codificación seleccionado se puede asociar, por ejemplo, con una mínima de las sumas de cuadrados de los formatos de codificación respectivos.

Según ejemplos de realizaciones, se proporciona un sistema de codificación de audio que comprende una sección de codificación configurada para codificar una señal de audio de M canales como una señal de mezcla descendente de dos canales y metadatos asociados, donde  $M \geq 4$ , y para emitir la señal de mezcla descendente y los metadatos para almacenamiento o transmisión conjuntos. La sección de codificación comprende una sección de mezcla descendente configurada para calcular la señal de mezcla descendente en base a la señal de audio de M canales. Un primer canal de la señal de mezcla descendente se forma como una combinación lineal de un primer grupo de uno o más canales de la señal de audio de M canales, y un segundo canal de la señal de mezcla descendente se forma como una combinación lineal de un segundo grupo de uno o más canales de la señal de audio de M canales. El primer y segundo grupos constituyen una división de los M canales de la señal de audio de M canales. La sección de codificación comprende además una sección de análisis configurada para determinar: parámetros de mezcla ascendente para la reconstrucción paramétrica de la señal de audio de M canales a partir de la señal de mezcla descendente; y parámetros de mezcla para obtener, en base a la señal de mezcla descendente, una señal de salida de dos canales. Un primer canal de la señal de salida se aproxima a una combinación lineal de un tercer grupo de uno o más canales de la señal de audio de M canales, y un segundo canal de la señal de salida se aproxima a una combinación lineal de un cuarto grupo de uno o más canales de la señal de audio de M canales. El tercer y cuarto grupos constituyen una división de los M canales de la señal de audio de M canales. Ambos del tercer y cuarto grupos comprenden al menos un canal del primer grupo. Los metadatos comprenden los parámetros de mezcla ascendente y los parámetros de mezcla.

45 Según ejemplos de realizaciones, se proporciona un producto de programa de ordenador que comprende un medio legible por ordenador con instrucciones para realizar cualquiera de los métodos del segundo aspecto.

Según ejemplos de realizaciones del sistema de codificación de audio, del método y del producto de programa de ordenador del segundo aspecto, descrito anteriormente, la señal de salida puede ser una señal de K canales, donde  $2 \leq K < M$ , en lugar de una señal de dos canales, y los K canales de la señal de salida pueden corresponder a una división de la señal de audio de M canales en K grupos, en lugar de dos canales de la señal de salida correspondientes a una división de la señal de M canales en dos grupos.

Más específicamente, según ejemplos de realizaciones, se proporciona un método de codificación de audio que comprende: recibir una señal de audio de M canales, donde  $M \geq 4$ ; y calcular una señal de mezcla descendente de dos canales en base a la señal de audio de M canales. Un primer canal de la señal de mezcla descendente se forma como una combinación lineal de un primer grupo de uno o más canales de la señal de audio de M canales, y un segundo canal de la señal de mezcla descendente se forma como una combinación lineal de un segundo grupo de uno o más canales de la señal de audio de M canales. El primer y segundo grupos constituyen una división de los M canales de la señal de audio de M canales. El método de codificación puede comprender además: determinar parámetros de mezcla ascendente para la reconstrucción paramétrica de la señal de audio de M canales a partir de la señal de mezcla descendente; y determinar los parámetros de mezcla para obtener, en base a la señal de mezcla descendente, una señal de salida de K canales, en donde  $2 \leq K < M$ , en donde cada uno de los K canales de la señal de salida se aproxima a una combinación lineal de un grupo de uno o más canales de la señal de audio de M

canales. Los grupos correspondientes a los canales respectivos de la señal de salida pueden constituir una división de los M canales de la señal de audio de M canales en K grupos de uno o más canales, y al menos dos de los K grupos pueden comprender al menos un canal del primer grupo. El método de codificación puede comprender además la salida de la señal de mezcla descendente y los metadatos para almacenamiento o transmisión conjunta, en donde los metadatos comprenden los parámetros de mezcla ascendente y los parámetros de mezcla.

En un ejemplo de realización, los parámetros de mezcla pueden controlar las contribuciones respectivas de la señal de mezcla descendente y de una señal descorrelacionada a la señal de salida. Al menos algunos de los parámetros de mezcla se pueden determinar minimizando la contribución de la señal descorrelacionada entre tales parámetros de mezcla que hacen que los canales de la señal de salida sean aproximaciones que conservan la covarianza de las combinaciones lineales (o sumas) del uno o más canales de los K grupos de canales respectivos. La contribución de la señal descorrelacionada se puede minimizar, por ejemplo, en el sentido de que se minimice la energía de la señal o la amplitud de esta contribución.

Las combinaciones lineales de los canales de los K grupos, que los K canales de la señal de salida han de aproximar, pueden corresponder, por ejemplo, a una señal de audio de K canales que tiene una primera matriz de covarianza. Los canales de la señal de salida que son aproximaciones que conservan la covarianza de las combinaciones lineales de los canales de los K grupos de canales, respectivamente, pueden corresponder, por ejemplo, a que una matriz de covarianza de la señal de salida coincida (o al menos coincida sustancialmente) con la primera matriz de covarianza.

Entre las aproximaciones que conservan la covarianza, un tamaño reducido (por ejemplo, energía o amplitud) de la contribución de la señal descorrelacionada puede ser indicativo de un aumento de fidelidad de la aproximación que se percibe por un oyente durante la reproducción. Emplear parámetros de mezcla que disminuyan la contribución de la señal descorrelacionada puede mejorar la fidelidad de la señal de salida como una representación de K canales de la señal de audio de M canales.

### III. Descripción general - medio legible por ordenador

Según un tercer aspecto, los ejemplos de realizaciones proponen medios legibles por ordenador. Las ventajas presentadas anteriormente para las características de los sistemas, métodos y productos de programa de ordenador, según el primer y/o segundo aspectos, pueden ser generalmente válidas para las características correspondientes de medios legibles por ordenador según el tercer aspecto.

Según ejemplos de realizaciones, se proporciona una portadora de datos que representa: una señal de mezcla descendente de dos canales; y parámetros de mezcla ascendente que permiten la reconstrucción paramétrica de una señal de audio de M canales en base a la señal de mezcla descendente, donde  $M \geq 4$ . Un primer canal de la señal de mezcla descendente corresponde a una combinación lineal de un primer grupo de uno o más canales de la señal de audio de M canales, y un segundo canal de la señal de mezcla descendente corresponde a una combinación lineal de un segundo grupo de uno o más canales de la señal de audio de M canales. El primer y segundo grupos constituyen una división de los M canales de la señal de audio de M canales. La portadora de datos representa además los parámetros de mezcla que permiten la provisión de una señal de salida de dos canales en base a la señal de mezcla descendente. Un primer canal de la señal de salida se aproxima a una combinación lineal de un tercer grupo de uno o más canales de la señal de audio de M canales, y un segundo canal de la señal de salida se aproxima a una combinación lineal de un cuarto grupo de uno o más canales de la señal de audio de M canales. El tercer y cuarto grupos constituyen una división de los M canales de la señal de audio de M canales. Ambos del tercer y cuarto grupos comprenden al menos un canal del primer grupo.

En un ejemplo de realización, los datos representados por la portadora de datos se pueden organizar en tramas de tiempo y se pueden estratificar de manera que, para una trama de tiempo dada, la señal de mezcla descendente y los parámetros de mezcla asociados para esa trama de tiempo se pueden extraer independientemente de los parámetros de mezcla ascendente asociados. Por ejemplo, la portadora de datos se puede estratificar de manera que la señal de mezcla descendente y los parámetros de mezcla asociados para esa trama de tiempo se puedan extraer sin extraer y/o acceder a los parámetros de mezcla ascendente asociados. Según ejemplos de realizaciones del medio legible por ordenador (o portadora de datos) del tercer aspecto, descrito anteriormente, la señal de salida puede ser una señal de K canales, donde  $2 \leq K < M$ , en lugar de una señal de dos canales, y los K canales de la señal de salida pueden corresponder a una división de la señal de audio de M canales en K grupos, en lugar de dos canales de la señal de salida correspondientes a una división de la señal de M canales en dos grupos.

Más específicamente, según ejemplos de realizaciones, se proporciona un medio legible por ordenador (o portadora de datos) que representa: una señal de mezcla descendente de dos canales; y parámetros de mezcla ascendente que permiten la reconstrucción paramétrica de una señal de audio de M canales en base a la señal de mezcla descendente, donde  $M \geq 4$ . Un primer canal de la señal de mezcla descendente corresponde a una combinación lineal de un primer grupo de uno o más canales de la señal de audio de M canales, y un segundo canal de la señal de mezcla descendente corresponde a una combinación lineal de un segundo grupo de uno o más canales de la señal de audio de M canales. El primer y segundo grupos constituyen una división de los M canales de la señal de audio de M canales. La portadora de datos puede representar además parámetros de mezcla que permiten la

provisión de una señal de salida de K canales en base a la señal de mezcla descendente, donde  $2 \leq K < M$ . Cada canal de la señal de salida puede aproximarse a una combinación lineal (por ejemplo, suma ponderada o no ponderada) de un grupo de uno o más canales de la señal de audio de M canales. Los grupos correspondientes a los canales respectivos de la señal de salida pueden constituir una división de los M canales de la señal de audio de M canales en K grupos de uno o más canales. Al menos dos de los K grupos pueden comprender al menos un canal del primer grupo.

Ejemplos de realizaciones adicionales se definen en las reivindicaciones dependientes. Se observa que los ejemplos de realizaciones incluyen todas las combinaciones de características, incluso si se enumeran en reivindicaciones mutuamente diferentes.

10 IV. Ejemplos de realizaciones

Las Fig. 4-6 ilustra formas alternativas para dividir una señal de audio de canales 11.1 en grupos de canales para la codificación paramétrica de la señal de audio de canales 11.1 como una señal de audio de canales 5.1, o para la reproducción de la señal de audio de canales 11.1 en un sistema de altavoces que comprende cinco altavoces y un altavoz de graves.

15 La señal de audio de canales 11.1 comprende los canales L (izquierdo), LS (lado izquierdo), LB (parte posterior izquierda), TFL (parte superior delantera izquierda), TBL (parte superior posterior izquierda), R (derecho), RS (lado derecho), RB (parte posterior derecha), TFR (parte superior delantera derecha), TBR (parte superior posterior derecha), C (centro) y LFE (efectos de baja frecuencia). Los cinco canales L, LS, LB, TFL y TBL forman una señal de audio de cinco canales que representa un espacio de la mitad izquierda en un entorno de reproducción de la señal de audio de canales 11.1. Los tres canales L, LS y LB representan diferentes direcciones horizontales en el entorno de reproducción y los dos canales TFL y TBL representan direcciones separadas verticalmente de las de los tres canales L, LS y LB. Los dos canales TFL y TBL, por ejemplo, se pueden destinar a la reproducción en altavoces de techo. De manera similar, los cinco canales R, RS, RB, TFR y TBR forman una señal de audio de cinco canales adicional que representa un espacio de la mitad derecha del entorno de reproducción, los tres canales R, RS y RB que representan diferentes direcciones horizontales en el entorno de reproducción y los dos canales TFR y TBR que representan direcciones separadas verticalmente de las de los tres canales R, RS y RB.

Con el fin de representar la señal de audio de canales 11.1 como una señal de audio de canales 5.1, la colección de canales L, LS, LB, TFL, TBL, R, RS, RB, TFR, TBR, C, y LFE se puede dividir en grupos de canales representados por canales de mezcla descendente respectivos y metadatos asociados. La señal de audio de cinco canales L, LS, LB, TFL, TBL se puede representar por una señal de mezcla descendente de dos canales  $L_1$ ,  $L_2$  y los metadatos asociados, mientras que la señal de audio de cinco canales R, RS, RB, TFR, TBR adicional se puede representar mediante una señal de mezcla descendente de dos canales  $R_1$ ,  $R_2$  adicional y metadatos adicionales asociados. Los canales C y LFE se pueden mantener como canales separados también en la representación de canales 5.1 de la señal de audio de canales 11.1.

35 La Fig. 4 ilustra un primer formato de codificación  $F_1$ , en el cual la señal de audio de cinco canales L, LS, LB, TFL, TBL se divide en un primer grupo 401 de canales L, LS, LB y un segundo grupo 402 de canales TFL, TBL, y en el que la señal de audio de cinco canales R, RS, RB, TFR, TBR adicional se divide en un primer grupo 403 adicional de canales R, RS, RB y un segundo grupo 404 adicional de canales TFR, TBR. En el primer formato de codificación  $F_1$ , el primer grupo de canales 401 se representa por un primer canal  $L_1$  de la señal de mezcla descendente de dos canales, y el segundo grupo 402 de canales se representa por un segundo canal  $L_2$  de la señal de mezcla descendente de dos canales. El primer canal  $L_1$  de la señal de mezcla descendente puede corresponder a una suma del primer grupo 401 de canales según

$$L_1 = L + LS + LB,$$

45 y el segundo canal  $L_2$  de la señal de mezcla descendente puede corresponder a una suma del segundo grupo 402 de canales según

$$L_2 = TFL + TBL.$$

En algunos ejemplos de realizaciones, algunos o todos los canales se pueden reescalar antes de sumar, de modo que el primer canal  $L_1$  de la señal de mezcla descendente pueda corresponder a una combinación lineal del primer grupo 401 de canales según  $L_1 = c_1L + c_2LS + c_3LB$ , y el segundo canal  $L_2$  de la señal de mezcla descendente pueda corresponder a una combinación lineal del segundo grupo 402 de canales según  $L_2 = c_4TFL + c_5TBL$ . Las ganancias  $c_2$ ,  $c_3$ ,  $c_4$ ,  $c_5$ , por ejemplo, pueden coincidir, mientras que la ganancia  $c_1$  puede tener, por ejemplo, un valor diferente; por ejemplo,  $c_1$  puede corresponder a ningún reescalado en absoluto. Por ejemplo, se pueden usar los valores  $c_1 = 1$  y  $c_2 = c_3 = c_4 = c_5 = 1/\sqrt{2}$ . No obstante, siempre que las ganancias  $c_1, \dots, c_5$  aplicadas a los respectivos canales L, LS, LB, TFL, TBL para el primer formato de codificación  $F_1$  coincidan con las ganancias aplicadas a estos canales en los otros formatos de codificación  $F_2$  y  $F_3$ , descritos a continuación con referencia a las Fig. 5 y 6, estas ganancias no afectan a los cálculos que se describen a continuación. De ahí que las ecuaciones y la aproximación derivada a continuación para los canales L, LS, LB, TFL, TBL se apliquen también a las versiones

reescaladas  $c_1L, c_2LS, c_3LB, c_4TFL, c_5TBL$  de estos canales. Si, por otra parte, se emplean diferentes ganancias en los diferentes formatos de codificación, al menos algunos de los cálculos realizados a continuación pueden tener que ser modificados; por ejemplo, se puede considerar la opción de incluir descorrelacionadores adicionales, con el interés de proporcionar aproximaciones más fieles.

- 5 De manera similar, el primer grupo adicional de canales 403 se representa por un primer canal  $R_1$  de la señal de mezcla descendente adicional, y el segundo grupo 404 adicional de canales se representa por un segundo canal  $R_2$  de la señal de mezcla descendente adicional.

10 El primer formato de codificación  $F_1$  proporciona canales de mezcla descendente dedicados  $L_2$  y  $R_2$  para representar los canales del techo TFL, TBL, TFR y TBR. El uso del primer formato de codificación  $F_1$ , por lo tanto, puede permitir la reconstrucción paramétrica de la señal de audio de canales 11.1 con fidelidad relativamente alta en los casos donde, por ejemplo, una dimensión vertical en el entorno de reproducción es importante para la impresión general de la señal de audio de canales 11.1.

15 La Fig. 5 ilustra un segundo formato de codificación  $F_2$ , en el que la señal de audio de cinco canales L, LS, LB, TFL, TBL se divide en el tercer 501 y el cuarto 502 grupos de canales representados por los canales  $L_1$  y  $L_2$  respectivos, donde los canales  $L_1$  y  $L_2$  corresponden a sumas de los grupos de canales respectivos, por ejemplo, empleando las mismas ganancias  $c_1, \dots, c_5$  para el reescalado como en el primer formato de codificación  $F_1$ . De manera similar, la señal de audio de cinco canales R, RS, RB, TFR, TBR adicional se divide en un tercer 503 y cuarto 504 grupos adicionales de canales representados por los canales  $R_1$  y  $R_2$  respectivos.

20 El segundo formato de codificación  $F_2$  no proporciona canales de mezcla descendente dedicados para representar los canales del techo TFL, TBL, TFR y TBR, pero puede permitir la reconstrucción paramétrica de la señal de audio de canales 11.1 con una fidelidad relativamente alta, por ejemplo, en casos donde la dimensión vertical en el entorno de reproducción no es tan importante para la impresión general de la señal de audio de canales 11.1. El segundo formato de codificación  $F_2$  también puede ser más adecuado para la reproducción de canales 5.1 que el primer formato de codificación  $F_1$ .

25 La Fig. 6 ilustra un tercer formato de codificación  $F_3$ , en el que la señal de audio de cinco canales L, LS, LB, TFL, TBL se divide en un quinto 601 y sexto 602 grupos de canales representados por los canales  $L_1$  y  $L_2$  respectivos de la señal de mezcla descendente, donde los canales  $L_1$  y  $L_2$  corresponden a sumas de los grupos de canales respectivos, por ejemplo, empleando las mismas ganancias  $c_1, \dots, c_5$  para el reescalado como en el primer formato de codificación  $F_1$ . De manera similar, la señal de audio de cinco canales R, RS, RB, TFR, TBR adicional se divide en un quinto 603 y sexto 604 grupos de canales adicionales representados por los canales  $R_1$  y  $R_2$  respectivos.

En el tercer formato de codificación  $F_3$ , los cuatro canales LS, LB, TFL y TBL se representan por el segundo canal  $L_2$ . Aunque la reconstrucción paramétrica de alta fidelidad de la señal de audio de canales 11.1 puede ser potencialmente más difícil en el tercer formato de codificación  $F_3$  que en los otros formatos de codificación, el tercer formato de codificación  $F_3$  se puede emplear, por ejemplo, para la reproducción de canales 5.1.

35 Los inventores se han dado cuenta de que los metadatos asociados con una representación de canales 5.1 de la señal de audio de canales 11.1 según uno de los formatos de codificación  $F_1, F_2, F_3$  se pueden emplear para generar una representación de canales 5.1 según otro de los formatos de codificación  $F_1, F_2, F_3$  sin reconstruir primero la señal de canales 11.1 original. La señal de cinco canales L, LS, LB, TFL, TBL que representa el plano de la mitad izquierda de la señal de audio de canales 11.1, y la señal de cinco canales R, RS, RB, TFR, TBR adicional que representa el plano de la mitad derecha, se pueden tratar de manera análoga.

Supongamos que se han sumado tres canales  $x_1, x_2, x_3$  para formar el canal de mezcla descendente  $m_1$ , según  $m_1 = x_1 + x_2 + x_3$ , y que se han de reconstruir  $x_1$  y  $x_2 + x_3$ . Todos de los tres canales  $x_1, x_2, x_3$  se pueden reconstruir a partir del canal de mezcla descendente  $m_1$  como

$$\begin{bmatrix} x_1 \\ x_2 + x_3 \end{bmatrix} \approx \begin{bmatrix} c_1 \\ 1 - c_1 \end{bmatrix} m_1 + \begin{bmatrix} p_{11} & p_{12} \\ -p_{11} & -p_{12} \end{bmatrix} \begin{bmatrix} D_1(m_1) \\ D_2(m_1) \end{bmatrix},$$

45 empleando los parámetros de mezcla ascendente  $c_i, 1 \leq i \leq 3$ , y  $p_{ij}, 1 \leq i \leq 3, 1 \leq j \leq 2$  determinados en el lado del codificador, y descorrelacionadores  $D_1$  y  $D_2$  independientes. Suponiendo que los parámetros de mezcla ascendente empleados satisfacen  $c_1 + c_2 + c_3 = 1$  y  $p_{1k} + p_{2k} + p_{3k} = 0$ , para  $k = 1, 2$ , entonces las señales  $x_1$  y  $x_2 + x_3$  se pueden reconstruir como

$$\begin{bmatrix} x_1 \\ x_2 + x_3 \end{bmatrix} \approx \begin{bmatrix} c_1 \\ 1 - c_1 \end{bmatrix} m_1 + \begin{bmatrix} p_{11} & p_{12} \\ -p_{11} & -p_{12} \end{bmatrix} \begin{bmatrix} D_1(m_1) \\ D_2(m_1) \end{bmatrix},$$

50 que se puede expresar como

$$\begin{bmatrix} x_1 \\ x_2 + x_3 \end{bmatrix} \approx \begin{bmatrix} c_1 \\ 1 - c_1 \end{bmatrix} m_1 + \begin{bmatrix} p_1 \\ -p_1 \end{bmatrix} D_1(m_1), \quad (1)$$

donde los dos descorrelacionadores  $D_1$  y  $D_2$  se han sustituido por un único descorrelacionador  $D_1$ , y donde  $p_1^2 = p_{11}^2 + p_{12}^2$ . Si dos canales  $x_4$  y  $x_5$  se han sumado para formar un segundo canal de mezcla descendente  $m_2$  según  $m_2 = x_4 + x_5$ , entonces las señales  $x_1$  y  $x_2 + x_3 + x_4 + x_5$  se pueden reconstruir como

$$\begin{bmatrix} x_1 \\ x_2 + x_3 + x_4 + x_5 \end{bmatrix} \approx \begin{bmatrix} c_1 & 0 \\ 1 - c_1 & 1 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} + \begin{bmatrix} p_1 \\ -p_1 \end{bmatrix} D_1(m_1). \quad (2)$$

- 5 Como se describe a continuación, la ecuación (2) se puede emplear para generar señales conforme al tercer formato de codificación  $F_3$  en base a señales conforme al primer formato de codificación  $F_1$ .

Los canales  $x_4$  y  $x_5$  se pueden reconstruir como

$$\begin{bmatrix} x_4 \\ x_5 \end{bmatrix} \approx \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} m_2 + \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} D_3(m_2) = \begin{bmatrix} d_1 \\ 1 - d_1 \end{bmatrix} m_2 + \begin{bmatrix} q_1 \\ -q_1 \end{bmatrix} D_3(m_2) \quad (3)$$

- 10 empleando un descorrelacionador  $D_3$  y los parámetros de mezcla ascendente que satisfacen  $d_1 + d_2 = 1$  y  $q_1 + q_2 = 0$ . En base a las ecuaciones (1) y (3), las señales  $x_1 + x_4$  y  $x_2 + x_3 + x_5$  se pueden reconstruir como

$$\begin{bmatrix} x_1 + x_4 \\ x_2 + x_3 + x_5 \end{bmatrix} \approx \begin{bmatrix} c_1 & d_1 \\ 1 - c_1 & 1 - d_1 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} (p_1 D_1(m_1) + q_1 D_3(m_2)),$$

y como

$$\begin{bmatrix} x_1 + x_4 \\ x_2 + x_3 + x_5 \end{bmatrix} \approx \begin{bmatrix} c_1 & d_1 \\ 1 - c_1 & 1 - d_1 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} D_1(am_1 + bm_2), \quad (4)$$

- 15 donde las contribuciones de los dos descorrelacionadores  $D_1$  y  $D_3$  (es decir, descorrelacionadores de un tipo que conserva la energía de su señal de entrada) se han aproximado por una contribución de un único descorrelacionador  $D_1$  (es decir, un descorrelacionador de un tipo que conserva la energía de su señal de entrada). Esta aproximación se puede asociar con una pérdida de fidelidad percibida muy pequeña, particularmente si los canales de mezcla descendente  $m_1, m_2$  no están correlacionados y si los valores  $a = p_1$  y  $b = q_1$  se emplean para las ponderaciones  $a$  y  $b$ . El formato de codificación según el cual los canales de mezcla descendente  $m_1, m_2$  se generan en un lado del codificador, por ejemplo, se puede haber elegido en un esfuerzo por mantener baja la correlación entre los canales de mezcla descendente  $m_1, m_2$ . Como se describe a continuación, la ecuación (4) se puede emplear para generar señales conforme al segundo formato de codificación  $F_2$  en base a señales conforme al primer formato de codificación  $F_1$ .

La estructura de la ecuación (4) se puede modificar opcionalmente a

$$25 \begin{bmatrix} x_1 + x_4 \\ x_2 + x_3 + x_5 \end{bmatrix} \approx \begin{bmatrix} c_1 & d_1 \\ 1 - c_1 & 1 - d_1 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} + \begin{bmatrix} g \\ -g \end{bmatrix} D_1\left(\frac{a}{g}m_1 + \frac{b}{g}m_2\right),$$

donde un factor de ganancia  $g = (a^2 + b^2)^{1/2}$  se emplea para ajustar la potencia de la señal de entrada al descorrelacionador  $D_1$ . También se pueden emplear otros valores del factor de ganancia, tales como  $g = (a^2 + b^2)^{1/v}$ , para  $0 < v < 1$ .

- 30 Si el primer formato de codificación  $F_1$  se emplea para proporcionar una representación paramétrica de la señal de canales 11.1, y el segundo formato de codificación  $F_2$  se desea en el lado del decodificador para la representación del contenido de audio, entonces aplicando la aproximación de la ecuación (4) en ambos de los lados izquierdo y derecho, e indicando la naturaleza aproximada de algunas de las cantidades del lado izquierdo (cuatro canales de la señal de salida) por tildes, produce

$$\begin{bmatrix} \tilde{L}_1 \\ \tilde{R}_1 \\ \tilde{C} \\ \tilde{L}_2 \\ \tilde{R}_2 \end{bmatrix} = \begin{bmatrix} c_{1,L} & 0 & 0 & d_{1,L} & 0 & 1 & 0 \\ 0 & c_{1,R} & 0 & 0 & d_{1,R} & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 - c_{1,L} & 0 & 0 & 1 - d_{1,L} & 0 & -1 & 0 \\ 0 & 1 - c_{1,R} & 0 & 0 & 1 - d_{1,R} & 0 & -1 \end{bmatrix} \begin{bmatrix} L_1 \\ R_1 \\ C \\ L_2 \\ R_2 \\ S_L \\ S_R \end{bmatrix}, \quad (5)$$

- 35 donde, según el segundo formato de codificación  $F_2$ ,

$$\tilde{L}_1 \approx L + TFL \quad \text{y} \quad \tilde{L}_2 \approx LS + LB + TBL,$$

$$\widetilde{R}_1 \approx R + TFR \quad \text{y} \quad \widetilde{R}_2 \approx RS + RB + TBR,$$

donde  $S_L = D(a_L L_1 + b_L L_2)$  y  $S_R = D(a_R R_1 + b_R R_2)$ , donde  $c_{1,L}$ ,  $d_{1,L}$ ,  $a_L$ ,  $b_L$  y  $c_{1,R}$ ,  $d_{1,R}$ ,  $a_R$ ,  $b_R$  son las versiones del canal izquierdo y del canal derecho, respectivamente, de los parámetros  $c_1$ ,  $d_1$ ,  $a$ ,  $b$  de la ecuación (4), y donde  $D$  denota un operador de descorrelación. De ahí que una aproximación del segundo formato de codificación  $F_2$  se pueda obtener a partir del primer formato de codificación  $F_1$  en base a los parámetros de mezcla ascendente para la reconstrucción paramétrica de la señal de audio de canales 11.1, sin tener que reconstruir realmente la señal de audio de canales 11.1.

Si el primer formato de codificación  $F_1$  se emplea para proporcionar una representación paramétrica de la señal de canales 11.1, y se desea el tercer formato de codificación  $F_3$  en el lado del decodificador para la reproducción del contenido de audio, entonces aplicando la aproximación de la ecuación (2) tanto en el lado izquierdo como derecho, e indicando la naturaleza aproximada de algunas de las cantidades del lado izquierdo, produce:

$$\begin{bmatrix} \widetilde{L}_1 \\ \widetilde{R}_1 \\ C \\ \widetilde{L}_2 \\ \widetilde{R}_2 \end{bmatrix} = \begin{bmatrix} c_{1,L} & 0 & 0 & 0 & 0 & p_{1,L} & 0 \\ 0 & c_{1,R} & 0 & 0 & 0 & 0 & p_{1,R} \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 - c_{1,L} & 0 & 0 & 1 & 0 & -p_{1,L} & 0 \\ 0 & 1 - c_{1,R} & 0 & 0 & 1 & 0 & -p_{1,R} \end{bmatrix} \begin{bmatrix} L_1 \\ R_1 \\ C \\ L_2 \\ R_2 \\ D(L_1) \\ D(R_1) \end{bmatrix}, \quad (6)$$

donde, mediante el tercer formato de codificación  $F_3$ ,

$$\begin{aligned} \widetilde{L}_1 &\approx L & \text{y} & \quad \widetilde{L}_2 \approx LS + LB + TFL + TBL, \\ \widetilde{R}_1 &\approx R & \text{y} & \quad \widetilde{R}_2 \approx RS + RB + TFR + TBR, \end{aligned}$$

donde  $c_{1,L}$ ,  $p_{1,L}$  y  $c_{1,R}$ ,  $p_{1,R}$  son versiones del canal izquierdo y del canal derecho, respectivamente, de los parámetros  $c_1$  y  $p_1$  de la ecuación (2), y donde  $D$  denota un operador de descorrelación. De ahí que se pueda obtener una aproximación del tercer formato de codificación  $F_3$  a partir del primer formato de codificación  $F_1$  en base a los parámetros de mezcla ascendente para la reconstrucción paramétrica de la señal de audio de canales 11.1, sin tener que reconstruir realmente la señal de audio de canales 11.1.

Si el segundo formato de codificación  $F_2$  se emplea para proporcionar una representación paramétrica de la señal de audio de canales 11.1, y el primer formato de codificación  $F_1$  o el tercer formato de codificación  $F_3$  se desea en el lado del decodificador para la representación del contenido de audio, se pueden derivar relaciones similares a las presentadas en las ecuaciones (5) y (6) usando las mismas ideas.

Si el tercer formato de codificación  $F_3$  se emplea para proporcionar una representación paramétrica de la señal de audio de canales 11.1, y el primer formato de codificación  $F_1$  o el segundo formato de codificación  $F_2$  se desea en el lado del decodificador para la representación del contenido de audio, se pueden emplear al menos algunas las ideas descritas anteriormente. No obstante, como el sexto grupo 602 de canales, representado por el canal  $\widetilde{L}_2$ , incluye cuatro canales  $LS$ ,  $LB$ ,  $TFL$ ,  $TBL$ , se puede emplear, por ejemplo, más de un canal descorrelacionado para el lado izquierdo (y de manera similar para el lado derecho), y el otro canal  $\widetilde{L}_1$  que representa solamente el canal  $L$  puede, por ejemplo, no ser incluido como entrada a cualquiera de los descorrelacionadores.

Como se ha descrito anteriormente, los parámetros de mezcla ascendente para la reconstrucción paramétrica de la señal de audio de canales 11.1 a partir de una representación paramétrica de canales 5.1 (conforme a uno de los formatos de codificación  $F_1$ ,  $F_2$  y  $F_3$ ) se pueden emplear para obtener una representación de canales 5.1 alternativa de la señal de audio de canales 11.1 (conforme a cualquiera de los otros formatos de codificación  $F_1$ ,  $F_2$  y  $F_3$ ). En otros ejemplos de realizaciones, la representación de canales 5.1 alternativa se puede obtener en base a parámetros de mezcla específicamente determinados para este propósito en el lado del codificador. Ahora se describirá una forma para determinar tales parámetros de mezcla.

Dadas dos señales de audio  $y_1 = u_1 + u_2$  y  $y_2 = u_3 + u_4$  formadas a partir de cuatro señales de audio  $u_1$ ,  $u_2$ ,  $u_3$ ,  $u_4$ , se puede obtener una aproximación de las dos señales de audio  $z_1 = u_1 + u_3$  y  $z_2 = u_2 + u_4$ . La diferencia  $z_1 - z_2$  se puede estimar a partir de  $y_1$  e  $y_2$  como una estimación de mínimos cuadrados según

$$z_1 - z_2 = \alpha y_1 + \beta y_2 + r,$$

donde la señal de error  $r$  es ortogonal tanto para  $y_1$  como para  $y_2$ . Empleando que  $z_1 + z_2 = y_1 + y_2$ , entonces se puede derivar que

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \frac{1}{2} \left( \begin{bmatrix} 1 + \alpha \\ 1 - \alpha \end{bmatrix} y_1 + \begin{bmatrix} 1 + \beta \\ 1 - \beta \end{bmatrix} y_2 + \begin{bmatrix} 1 \\ -1 \end{bmatrix} r \right). \quad (7)$$

Con el fin de llegar a una aproximación que reinstaurando la estructura de covarianza correcta de las señales  $z_1$  y  $z_2$ , la señal de error  $r$  se puede sustituir por una señal descorrelacionada de la misma potencia, por ejemplo, de la forma  $\gamma D(y_1 + y_2)$ , donde  $D$  denota la descorrelación y donde el parámetro  $\gamma$  se ajusta para conservar la potencia de la señal. Empleando una parametrización diferente de la ecuación (7), la aproximación se puede expresar como

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \approx \begin{bmatrix} c \\ 1 - c \end{bmatrix} y_1 + \begin{bmatrix} d \\ 1 - d \end{bmatrix} y_2 + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \gamma D(y_1 + y_2). \quad (8)$$

Si el primer formato de codificación  $F_1$  se emplea para proporcionar una representación paramétrica de la señal de canales 11.1, y el segundo formato de codificación  $F_2$  se desea en el lado del decodificador para la representación del contenido de audio, entonces aplicando la aproximación de la ecuación (8) con  $z_1 = L + TFL$ ,  $z_2 = LS + LB + TBL$ ,  $y_1 = L + LS + LB$ , e  $y_2 = TFL + TBL$  en el lado izquierdo, y con  $z_1 = R + TFR$ ,  $z_2 = RS + RB + TBR$ ,  $y_1 = R + RS + RB$ , e  $y_2 = TFR + TBR$  en el lado derecho, e indicando la naturaleza aproximada de algunas de las cantidades del lado izquierdo por tildes, se produce:

$$\begin{bmatrix} \tilde{L}_1 \\ \tilde{R}_1 \\ C \\ \tilde{L}_2 \\ \tilde{R}_2 \end{bmatrix} = \begin{bmatrix} c_L & 0 & 0 & d_L & 0 & \gamma_L & 0 \\ 0 & c_R & 0 & 0 & d_R & 0 & \gamma_R \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 - c_L & 0 & 0 & 1 - d_L & 0 & -\gamma_L & 0 \\ 0 & 1 - c_R & 0 & 0 & 1 - d_R & 0 & -\gamma_R \end{bmatrix} \begin{bmatrix} L_1 \\ R_1 \\ C \\ L_2 \\ R_2 \\ r_L \\ r_R \end{bmatrix} \quad (9)$$

donde, mediante el primer formato de codificación  $F_1$ ,

$$\begin{aligned} \tilde{L}_1 &\approx L + TFL & \text{y} & & \tilde{L}_2 &\approx LS + LB + TBL, \\ \tilde{R}_1 &\approx R + TFR & \text{y} & & \tilde{R}_2 &\approx RS + RB + TBR, \end{aligned}$$

donde  $r_L = D(L_1 + L_2)$  y  $r_R = D(R_1 + R_2)$ , donde  $c_L$ ,  $d_L$ ,  $\gamma_L$  y  $c_R$ ,  $d_R$ ,  $\gamma_R$  son las versiones del canal izquierdo y del canal derecho, respectivamente, de los parámetros  $c$ ,  $d$ ,  $\gamma$  de la ecuación (8), y donde  $D$  denota descorrelación. De ahí que una aproximación del segundo formato de codificación  $F_2$  se pueda obtener a partir del primer formato de codificación  $F_1$  en base a los parámetros de mezcla  $c_L$ ,  $d_L$ ,  $\gamma_L$  y  $c_R$ ,  $d_R$ ,  $\gamma_R$ , por ejemplo, determinados en el lado del codificador para ese propósito y transmitidos juntos con las señales de mezcla descendente al lado del decodificador. El uso de parámetros de mezcla permite un aumento del control del lado del codificador. Dado que la señal de audio de canales 11.1 original está disponible en el lado del codificador, los parámetros de mezcla se pueden sintonizar, por ejemplo, en el lado del codificador para aumentar la fidelidad de la aproximación del segundo formato de codificación  $F_2$ .

De manera similar, una aproximación del tercer formato de codificación  $F_3$  se puede obtener a partir del primer formato de codificación  $F_1$  en base a parámetros de mezcla similares. Aproximaciones similares del primer formato de codificación  $F_1$  y del tercer formato de codificación  $F_3$  también se pueden obtener a partir del segundo formato de codificación  $F_2$ .

Como puede ver en la ecuación (9), los dos canales de la señal de salida  $\tilde{L}_1$ ,  $\tilde{L}_2$  reciben contribuciones de igual magnitud de la señal descorrelacionada  $r_L$ , pero de signos opuestos. La situación correspondiente se mantiene para las contribuciones de las señales descorrelacionadas  $S_L$  y  $D(L_1)$  en las ecuaciones (5) y (6), respectivamente.

Como se puede ver en la ecuación (9), la suma del coeficiente de mezcla  $c_L$  que controla una contribución del primer canal  $L_1$  de la señal de mezcla descendente al primer canal  $L_1$  de la señal de salida, y el coeficiente de mezcla  $1 - c_L$  que controla una contribución del primer canal  $L_1$  de la señal de mezcla descendente al segundo canal  $\tilde{L}_2$  de la señal de salida, tiene el valor 1. Las relaciones correspondientes también se mantienen en las ecuaciones (5) y (6).

La Fig. 1 es un diagrama de bloques generalizado de una sección de codificación 100 para codificar una señal de  $M$  canales como una señal de mezcla descendente de dos canales y metadatos asociados, según un ejemplo de realización.

La señal de audio de  $M$  canales se ejemplifica en la presente memoria mediante la señal de cinco canales  $L$ ,  $LS$ ,  $LB$ ,  $TFL$  y  $TBL$  descrita con referencia a la Fig. 4, y la señal de mezcla descendente se ejemplifica mediante el primer canal  $L_1$  y un segundo canal  $L_2$  calculado según el primer formato de codificación  $F_1$  descrito con referencia a la Fig. 4. Se pueden contemplar ejemplos de realizaciones en los que la sección de codificación 100 calcula una señal de

mezcla descendente según cualquiera de los formatos de codificación descritos con referencia a las Fig. 4 a 6. También se pueden contemplar ejemplos de realizaciones en las que la sección de codificación 100 calcula una señal de mezcla descendente en base a una señal de audio de M canales, donde  $M \geq 4$ . En particular, se apreciará que cálculos y aproximaciones similares a los descritos anteriormente, y que conducen a las ecuaciones (5), (6) y (9), se pueden realizar para ejemplos de realizaciones donde  $M = 4$ , o  $M \geq 6$ .

La sección de codificación 100 comprende una sección de mezcla descendente 110 y una sección de análisis 120. La sección de mezcla descendente 110 calcula la señal de mezcla descendente en base a la señal de audio de cinco canales formando el primer canal  $L_1$  de la señal de mezcla descendente como una combinación lineal (por ejemplo, como una suma) del primer grupo 401 de canales de la señal de audio de cinco canales, y formando el segundo canal  $L_2$  de la señal de mezcla descendente como una combinación lineal (por ejemplo, como una suma) del segundo grupo 402 de canales de la señal de audio de cinco canales. El primer y segundo grupos 401, 402 constituyen una división de los cinco canales L, LS, LB, TFL, TBL de la señal de audio de cinco canales. La sección de análisis 120 determina los parámetros de mezcla ascendente  $\alpha_{LU}$  para la reconstrucción paramétrica de la señal de audio de cinco canales a partir de la señal de mezcla descendente en un decodificador paramétrico. La sección de análisis 120 también determina los parámetros de mezcla  $\alpha_{LM}$  para obtener, en base a la señal de mezcla descendente, una señal de salida de dos canales.

En el presente ejemplo de realización, la señal de salida es una representación de dos canales de la señal de audio de cinco canales según el segundo formato de codificación  $F_2$  descrito con referencia a la Fig. 5. No obstante, también se pueden contemplar ejemplos de realizaciones en los cuales la señal de salida representa la señal de audio de cinco canales según cualquiera de los formatos de codificación descritos con referencia a las Fig. 4 a 6.

Un primer canal  $\widetilde{L}_1$  de la señal de salida se aproxima a una combinación lineal (por ejemplo, una suma) del tercer grupo 501 de canales de la señal de audio de cinco canales, y un segundo canal  $\widetilde{L}_2$  de la señal de salida se aproxima a una combinación lineal (por ejemplo, una suma) del cuarto grupo 502 de canales de la señal de audio de cinco canales. El tercer y cuarto grupos 501, 502 constituyen una división diferente de los cinco canales L, LS, LB, TFL, TBL de la señal de audio de cinco canales que se proporciona por el primer y segundo grupos 401, 402 de canales. En particular, el tercer grupo 501 comprende el canal L del primer grupo 401, mientras que el cuarto grupo 502 comprende los canales LS y LB del primer grupo 401.

La sección de codificación 100 emite la señal de mezcla descendente  $L_1$ ,  $L_2$  y los metadatos asociados para el almacenamiento y/o la transmisión conjunta al lado del decodificador. Los metadatos comprenden los parámetros de mezcla ascendente  $\alpha_{LU}$  y los parámetros de mezcla  $\alpha_{LM}$ . Los parámetros de mezcla  $\alpha_{LM}$  pueden transportar suficiente información para emplear la ecuación (9) para obtener la señal de salida  $\widetilde{L}_1$ ,  $\widetilde{L}_2$  en base a la señal de mezcla descendente  $L_1$ ,  $L_2$ . Los parámetros de mezcla  $\alpha_{LM}$ , por ejemplo, pueden incluir los parámetros  $c_L$ ,  $d_L$ ,  $\gamma_L$  o incluso todos los elementos de la matriz de más a la izquierda en la ecuación (9).

La Fig. 2 es un diagrama de bloques generalizado de un sistema de codificación de audio 200 que comprende la sección de codificación 100 descrita con referencia a la Fig. 1, según un ejemplo de realización. En el presente ejemplo de realización, un contenido de audio, por ejemplo, grabado por uno o más transductores acústicos 201, o generado por el equipo de creación de audio 201, se proporciona en forma de la señal de audio de canales 11.1 descrita con referencia a las Fig. 4 a 6. Una sección de análisis de filtro de espejo en cuadratura (QMF) 202 transforma la señal de audio de cinco canales L, LS, LB TFL, TBL, segmento de tiempo por segmento de tiempo, en un dominio de QMF para procesamiento por la sección de codificación 100 del audio de cinco canales en forma de cuadros de tiempo/frecuencia. El sistema de codificación de audio 200 comprende una sección de codificación 203 adicional análoga a la sección de codificación 100 y adaptada para codificar la señal de audio de cinco canales R, RS, RB, TFR y TBR adicional como la señal de mezcla descendente de dos canales  $R_1$ ,  $R_2$  adicional y metadatos asociados que comprenden parámetros de mezcla ascendente  $\alpha_{RU}$  adicionales y parámetros de mezcla  $\alpha_{RM}$  adicionales. Los parámetros de mezcla  $\alpha_{RM}$  adicionales pueden incluir, por ejemplo, los parámetros  $c_R$ ,  $d_R$  y  $\gamma_R$  de la ecuación (9). La sección de análisis de QMF 202 también transforma la señal de audio de cinco canales R, RS, RB, TFR y TBR adicional en un dominio de QMF para su procesamiento por la sección de codificación 203 adicional. La señal de mezcla descendente  $L_1$ ,  $L_2$  emitida por la sección de codificación 100 se transforma de vuelta del dominio de QMF por una sección de síntesis de QMF 204 y se transforma en un dominio de transformada de coseno discreta modificada (MDCT) por una sección de transformación 205. Las secciones de cuantificación 206 y 207 cuantifican los parámetros de mezcla ascendente  $\alpha_{LU}$  y los parámetros de mezcla  $\alpha_{LM}$ , respectivamente. Por ejemplo, se puede emplear una cuantificación uniforme con un tamaño de paso de 0.1 o 0.2 (adimensional), seguida por una codificación de entropía en forma de codificación de Huffman. Una cuantificación más gruesa con un tamaño de paso 0.2 se puede emplear, por ejemplo, para ahorrar ancho de banda de transmisión, y una cuantificación más fina con un tamaño de paso 0.1, por ejemplo, se puede emplear para mejorar la fidelidad de la reconstrucción en el lado del decodificador. De manera similar, la señal de mezcla descendente  $R_1$ ,  $R_2$  adicional emitida por la sección de codificación 203 adicional se transforma de nuevo desde el dominio de QMF mediante una sección de síntesis de QMF 208 y se transforma en un dominio de MDCT por una sección de transformación 209. Las secciones de cuantificación 210 y 211 cuantifican los parámetros de mezcla ascendente  $\alpha_{RU}$  adicionales y los parámetros de mezcla  $\alpha_{RM}$  adicionales, respectivamente. Los canales C y LFE también se transforman en un dominio de MDCT por

las secciones de transformación 214 y 215 respectivas. Las señales y canales de mezcla descendente transformados por MDCT, y los metadatos cuantificados, entonces se combinan en un flujo de bits B mediante un multiplexor 216, para la transmisión al lado del decodificador. El sistema de codificación de audio 200 también puede comprender un codificador central (no mostrado en la Fig. 2) configurado para codificar la señal de mezcla descendente  $L_1$ ,  $L_2$ , la señal de mezcla descendente  $R_1$ ,  $R_2$  adicional y los canales C y LFE usando un códec de audio perceptivo, tal como Dolby Digital o MPEG AAC, antes de que las señales de mezcla descendente y los canales C y LFE se proporcionen al multiplexor 216. Una ganancia de recorte, por ejemplo, correspondiente a -8.7 dB, se puede aplicar, por ejemplo, a la señal de mezcla descendente  $L_1$ ,  $L_2$ , la señal de mezcla descendente  $R_1$ ,  $R_2$  adicional y el canal C, antes de formar el flujo de bits B.

5  
10  
15

La Fig. 3 es un diagrama de flujo de un método de codificación de audio 300 realizado por el sistema de codificación de audio 200, según un ejemplo de realización. El método de codificación de audio 300 comprende: recibir 310 la señal de audio de cinco canales L, LS, LB, TFL, TBL; calcular 320 la señal de mezcla descendente de dos canales  $L_1$ ,  $L_2$  en base a la señal de audio de cinco canales; determinar 330 los parámetros de mezcla ascendente  $\alpha_{LU}$ ; determinar 340 los parámetros de mezcla  $\alpha_{LM}$ ; y emitir 350 la señal de mezcla descendente y los metadatos para el almacenamiento y/o transmisión conjunta, en donde los metadatos comprenden los parámetros de mezcla ascendente  $\alpha_{LU}$  y los parámetros de mezcla  $\alpha_{LM}$ .

La Fig. 7 es un diagrama de bloques generalizado de una sección de decodificación 700 para proporcionar una señal de salida de dos canales  $\widetilde{L}_1$ ,  $\widetilde{L}_2$  en base a una señal de mezcla descendente de dos canales  $L_1$ ,  $L_2$  y metadatos asociados, según un ejemplo de realización.

20  
25

En el presente ejemplo de realización, la señal de mezcla descendente  $L_1$ ,  $L_2$  es la señal de mezcla descendente  $L_1$ ,  $L_2$  emitida por la sección de codificación 100 descrita con referencia a la Fig. 1, y se asocia con ambos de los parámetros de mezcla ascendente  $\alpha_{LU}$  y los parámetros de mezcla  $\alpha_{LM}$  emitidos por la sección de codificación 100. Como se describe con referencia a las Fig. 1 y 4, los parámetros de mezcla ascendente  $\alpha_{LU}$  se adaptan para la reconstrucción paramétrica de la señal de audio de cinco canales L, LS, LB, TFL, TBL en base a la señal de mezcla descendente  $L_1$ ,  $L_2$ . No obstante, también se pueden contemplar realizaciones en las que los parámetros de mezcla ascendente  $\alpha_{LU}$  se adaptan para la reconstrucción paramétrica de una señal de audio de M canales, donde  $M = 4$ , o  $M \geq 6$ .

30

En el presente ejemplo de realización, el primer canal  $L_1$  de la señal de mezcla descendente corresponde a una combinación lineal (por ejemplo, una suma) del primer grupo 401 de canales de la señal de audio de cinco canales, y el segundo canal  $L_2$  de la señal de mezcla descendente corresponde a una combinación lineal (por ejemplo, una suma) del segundo grupo 402 de canales de la señal de audio de cinco canales. El primer y segundo grupos 401, 402 constituyen una división de los cinco canales L, LS, LB, TFL, TBL de la señal de audio de cinco canales.

35  
40

En el presente ejemplo de realización, la sección de decodificación 700 recibe la señal de mezcla descendente de dos canales  $L_1$ ,  $L_2$  y los parámetros de mezcla ascendente  $\alpha_{LU}$ , y proporciona la señal de salida de dos canales  $\widetilde{L}_1$ ,  $\widetilde{L}_2$  en base a la señal de mezcla descendente  $L_1$ ,  $L_2$  y los parámetros de mezcla ascendente  $\alpha_{LU}$ . La sección de decodificación 700 comprende una sección de descorrelación 710 y una sección de mezcla 720. La sección de descorrelación 710 recibe la señal de mezcla descendente  $L_1$ ,  $L_2$  y emite, en base a la misma y según los parámetros de mezcla ascendente (compárese con las ecuaciones (4) y (5)), una señal descorrelacionada de un único canal D. La sección de mezcla 720 determina un conjunto de coeficientes de mezcla en base a los parámetros de mezcla ascendente  $\alpha_{LU}$ , y forma la señal de salida  $\widetilde{L}_1$ ,  $\widetilde{L}_2$  como una combinación lineal de la señal de mezcla descendente  $L_1$ ,  $L_2$  y la señal descorrelacionada D según los coeficientes de mezcla. En otras palabras, la sección de mezcla 720 realiza una proyección de tres canales a dos canales.

45

En el presente ejemplo de realización, la sección de decodificación 700 se configura para proporcionar la señal de salida  $\widetilde{L}_1$ ,  $\widetilde{L}_2$  según el segundo formato de codificación  $F_2$  descrito con referencia a la Fig. 5, y por lo tanto forma la señal de salida  $\widetilde{L}_1$ ,  $\widetilde{L}_2$  según la ecuación (5). En otras palabras, los coeficientes de mezcla corresponden a los elementos en la matriz de más a la izquierda de la ecuación (5), y se puede determinar por la sección de mezcla en base a los parámetros de mezcla ascendente  $\alpha_{LU}$ .

50

De ahí que la sección de mezcla 720 determine los coeficientes de mezcla de manera que un primer canal  $\widetilde{L}_1$  de la señal de salida se aproxima a una combinación lineal (por ejemplo, una suma) del tercer grupo 501 de canales de la señal de audio de cinco canales L, LS, LB, TFL, TBL, y de manera que un segundo canal  $\widetilde{L}_2$  de la señal de salida se aproxima a una combinación lineal (por ejemplo, una suma) del cuarto grupo de canales de la señal de audio de cinco canales L, LS, LB, TFL, TBL. Como se describe con referencia a la Fig. 5, el tercer y cuarto grupos 501, 502 constituyen una división de la señal de cinco canales L, LS, LB, TFL, TBL de la señal de audio de cinco canales, y ambos del tercer y cuarto grupos 501, 502 comprenden al menos un canal del primer grupo 401 de canales.

En algunos ejemplos de realizaciones, los coeficientes empleados para la reconstrucción paramétrica de la señal de audio de cinco canales  $L$ ,  $LS$ ,  $LB$ ,  $TFL$ ,  $TBL$  de la señal de mezcla descendente  $L_1$ ,  $L_2$  y de una señal descorrelacionada se puede representar por los parámetros de mezcla ascendente  $\alpha_{LU}$  de una forma compacta que incluye menos parámetros que el número de coeficientes reales empleados para la reconstrucción paramétrica. En tales realizaciones, los coeficientes reales se pueden derivar en el lado del decodificador en base al conocimiento de la forma compacta particular empleada.

La Fig. 8 es un diagrama de bloques generalizado de un sistema de decodificación de audio 800 que comprende la sección de decodificación 700 descrita con referencia a la Fig. 7, según un ejemplo de realización.

Una sección de recepción 801, por ejemplo, que incluye un demultiplexor, recibe el flujo de bits  $B$  transmitido desde el sistema de codificación de audio 200 descrito con referencia a la Fig. 2, y extrae la señal de mezcla descendente  $L_1$ ,  $L_2$  y los parámetros de mezcla ascendente  $\alpha_{LU}$  asociados, la señal de mezcla descendente  $R_1$ ,  $R_2$  adicional y los parámetros de mezcla ascendente  $\alpha_{RU}$  adicionales asociados, así como los canales  $C$  y  $LFE$ , del flujo de bits  $B$ .

Aunque los parámetros de mezcla  $\alpha_{LM}$  y los parámetros de mezcla  $\alpha_{RM}$  adicionales pueden estar disponibles en el flujo de bits  $B$ , estos parámetros no se emplean por el sistema de decodificación de audio 800 en el presente ejemplo de realización. En otras palabras, el sistema de decodificación de audio 800 del presente ejemplo de realización es compatible con flujos de bits de los que no se pueden extraer tales parámetros de mezcla. Una sección de decodificación que emplea los parámetros de mezcla  $\alpha_{LM}$  se describirá más adelante con referencia a la Fig. 9.

En caso de que la señal de mezcla descendente  $L_1$ ,  $L_2$ , la señal de mezcla descendente  $R_1$ ,  $R_2$  adicional y/o los canales  $C$  y  $LFE$  se codifiquen en el flujo de bits  $B$  usando un códec de audio perceptivo tal como Dolby Digital, MPEG AAC, o desarrollos de los mismos, el sistema de decodificación de audio 800 puede comprender un decodificador central (no mostrado en la Fig. 8) configurado para decodificar las señales y canales respectivos cuando se extraen del flujo de bits  $B$ .

Una sección de transformación 802 transforma la señal de mezcla descendente  $L_1$ ,  $L_2$  realizando una MDCT inversa y una sección de análisis de QMF 803 transforma la señal de mezcla descendente  $L_1$ ,  $L_2$  en un dominio de QMF para procesamiento por la sección de decodificación 700 de la señal de mezcla descendente  $L_1$ ,  $L_2$  en forma de cuadros de tiempo/frecuencia. Una sección de descuantificación 804 descuantifica los parámetros de mezcla ascendente  $\alpha_{LU}$ , por ejemplo, a partir de un formato codificado por entropía, antes de suministrarlos a la sección de decodificación 700. Como se describe con referencia a la Fig. 2, la cuantificación se puede haber realizado con uno de dos tamaños de pasos diferentes, por ejemplo, 0.1 o 0.2. El tamaño de paso real empleado puede estar predefinido, o se puede señalar al sistema de decodificación de audio 800 desde el lado del codificador, por ejemplo, a través del flujo de bits  $B$ .

En el presente ejemplo de realización, el sistema de decodificación de audio 800 comprende una sección de decodificación 805 adicional análoga a la sección de decodificación 700. La sección de decodificación 805 adicional está configurada para recibir la señal de mezcla descendente de dos canales  $R_1$ ,  $R_2$  adicional descrita con referencia a las Fig. 2 y 4, y los metadatos adicionales que incluyen parámetros de mezcla ascendente  $\alpha_{RU}$  adicionales para la reconstrucción paramétrica de las señales de audio de cinco canales  $R$ ,  $RS$ ,  $RB$ ,  $TFR$ ,  $TBR$  adicionales en base a la señal de mezcla descendente  $R_1$ ,  $R_2$  adicional. La sección de codificación 805 adicional se configura para proporcionar una señal de salida de dos canales  $\widetilde{R}_1$ ,  $\widetilde{R}_2$  adicional en base a la señal de mezcla descendente y los parámetros de mezcla ascendente  $\alpha_{RU}$  adicionales. La señal de salida  $\widetilde{R}_1$ ,  $\widetilde{R}_2$  adicional proporciona una representación de la señal de audio de cinco canales  $R$ ,  $RS$ ,  $RB$ ,  $TFR$ ,  $TBR$  adicional conforme al segundo formato de codificación  $F_2$  descrito con referencia a la Fig. 5.

Una sección de transformación 806 transforma la señal de mezcla descendente  $R_1$ ,  $R_2$  adicional realizando una MDCT inversa y una sección de análisis de QMF 807 transforma la señal de mezcla descendente  $R_1$ ,  $R_2$  adicional en un dominio de QMF para su procesamiento por la sección de decodificación 805 adicional de la señal de mezcla descendente  $R_1$ ,  $R_2$  adicional en forma de cuadros de tiempo/frecuencia. Una sección de descuantificación 808 descuantifica los parámetros de mezcla ascendente  $\alpha_{RU}$  adicionales, por ejemplo, de un formato codificado de entropía, antes de suministrarlos a la sección de decodificación 805 adicional.

En ejemplos de realizaciones donde se ha aplicado una ganancia de recorte a la señal de mezcla descendente  $L_1$ ,  $L_2$ , la señal de mezcla descendente  $R_1$ ,  $R_2$  adicional, y el canal  $C$  en el lado del codificador, una ganancia correspondiente, por ejemplo, correspondiente a 8.7 dB, se puede aplicar a estas señales en el sistema de decodificación de audio 800 para compensar la ganancia de recorte.

En el ejemplo de realización descrita con referencia a la Fig. 8, la señal de salida  $\widetilde{L}_1$ ,  $\widetilde{L}_2$  y la señal de salida  $\widetilde{R}_1$ ,  $\widetilde{R}_2$  adicional emitidas por la sección de decodificación 700 y la sección de decodificación 805 adicional, respectivamente, se transforman de nuevo desde el dominio de QMF por una sección de síntesis de QMF 811 antes de ser proporcionadas junto con los canales  $C$  y  $LFE$  como salida del sistema de decodificación de audio 800 para la

reproducción en un sistema de altavoces múltiples 812 que incluye, por ejemplo, cinco altavoces y un altavoz de graves. Las secciones de transformación 809, 810 transforman los canales C y LFE en el dominio del tiempo realizando una MDCT inversa antes de que estos canales se incluyan en la salida del sistema de decodificación de audio 800.

5 Los canales C y LFE se pueden extraer, por ejemplo, del flujo de bits B de una forma codificada de manera discreta y el sistema de decodificación 800 puede comprender, por ejemplo, secciones de decodificación de un único canal (no mostradas en la Fig. 8) configuradas para decodificar los canales codificados de manera discreta respectivos. La sección de decodificación de un único canal puede incluir, por ejemplo, decodificadores centrales para decodificar el contenido de audio codificado usando un códec de audio perceptivo tal como Dolby Digital, MPEG AAC, o desarrollos de los mismos.

10 La Fig. 9 es un diagrama de bloques generalizado de una sección de decodificación alternativa 900, según un ejemplo de realización. La sección de decodificación 900 es similar a la sección de decodificación 700 descrita con referencia a la Fig. 7, excepto que la sección de decodificación 900 emplea los parámetros de mezcla  $\alpha_{LM}$  proporcionados por la sección de codificación 100, descritos con referencia a la Fig. 1, en lugar de emplear los parámetros de mezcla ascendente  $\alpha_{LU}$  también proporcionados por la sección de codificación 100.

15 De manera similar a la sección de decodificación 700, la sección de decodificación 900 comprende una sección de descorrelación 910 y una sección de mezcla 920. La sección descorrelación 910 se configura para recibir la señal de mezcla descendente  $L_1, L_2$ , proporcionada por la sección de codificación 100 descrita con referencia a la Fig. 1, y para emitir, en base a la señal de mezcla descendente  $L_1, L_2$ , una señal descorrelacionada de un único canal D. La sección de mezcla 920 determina un conjunto de coeficientes de mezcla en base a los parámetros de mezcla  $\alpha_{LM}$ , y forma una señal de salida  $\widetilde{L}_1, \widetilde{L}_2$  como combinación lineal de la señal de mezcla descendente  $L_1, L_2$  y la señal descorrelacionada D, según los coeficientes de mezcla. La sección de mezcla 920 determina los parámetros de mezcla independientemente de los parámetros de mezcla ascendente  $\alpha_{LU}$  y forma la señal de salida  $\widetilde{L}_1, \widetilde{L}_2$  realizando una proyección de tres a dos canales.

25 En el presente ejemplo de realización, la sección de decodificación 900 se configura para proporcionar la señal de salida  $\widetilde{L}_1, \widetilde{L}_2$  según el segundo formato de codificación  $F_2$ , descrito con referencia a la Fig. 5 y, por lo tanto, forma la señal de salida  $\widetilde{L}_1, \widetilde{L}_2$  según la ecuación (9). En otras palabras, los parámetros de mezcla  $\alpha_{LM}$  recibidos pueden incluir los parámetros  $c_L, d_L, \gamma_L$  en la matriz de más a la izquierda de la ecuación (9), y los parámetros de mezcla  $\alpha_{LM}$  pueden haber sido determinados en el lado del codificador como se describe en relación con la ecuación (9). De ahí que la sección de mezcla 920 determine los coeficientes de mezcla de manera que un primer canal  $\widetilde{L}_1$  de la señal de salida se aproxime a una combinación lineal (por ejemplo, una suma) del tercer grupo 501 de canales de la señal de audio de cinco canales L, LS, LB, TFL, TBL descrita con referencia a las Fig. 4 a 6, y de manera que un segundo canal  $\widetilde{L}_2$  de la señal de salida se aproxime a una combinación lineal (por ejemplo, una suma) del cuarto grupo 502 de canales de la señal de audio de cinco canales L, LS, LB, TFL, TBL.

35 La señal de mezcla descendente  $L_1, L_2$  y los parámetros de mezcla  $\alpha_{LM}$  se pueden extraer, por ejemplo, a partir del flujo de bits B emitido por el sistema de codificación de audio 200 descrito con referencia a la Fig. 2. Los parámetros de mezcla ascendente  $\alpha_{LU}$  también codificados en el flujo de bits B no se pueden emplear por la sección de decodificación 900 del presente ejemplo de realización, y por lo tanto no necesitan ser extraídos del flujo de bits B.

40 La Fig. 10 es un diagrama de flujo de un método de decodificación de audio 1000 para proporcionar una señal de salida de dos canales en base a una señal de mezcla descendente de dos canales y parámetros de mezcla ascendente asociados, según un ejemplo de realización. El método de decodificación 1000 se puede realizar, por ejemplo, por el sistema de decodificación de audio 800 descrito con referencia a la Fig. 8.

45 El método de decodificación 1000 comprende recibir 1010 una señal de mezcla descendente de dos canales que se asocia con metadatos que comprenden parámetros de mezcla ascendente para la reconstrucción paramétrica de la señal de audio de cinco canales L, LS, LB, TFL, TBL, descrita con referencia a las Fig. 4 a 6, en base a la señal de mezcla descendente. La señal de mezcla descendente puede ser, por ejemplo, la señal de mezcla descendente  $L_1, L_2$  descrita con referencia a la Fig. 1, y puede ser conforme al primer formato de codificación  $F_1$ , descrito con respecto a la Fig. 4. El método de decodificación 1000 comprende además recibir 1020 al menos algunos de los metadatos. Los metadatos recibidos pueden incluir, por ejemplo, los parámetros de mezcla ascendente  $\alpha_{LU}$  y/o los parámetros de mezcla  $\alpha_{LM}$  descritos con referencia a la Fig. 1. El método de decodificación 1000 comprende además: generar 1040 una señal descorrelacionada en base a al menos un canal de la señal de mezcla descendente; determinar 1050 un conjunto de coeficientes de mezcla en base a los metadatos recibidos; y formar 1060 una señal de salida de dos canales como una combinación lineal de la señal de mezcla descendente y la señal descorrelacionada, según los coeficientes de mezcla. La señal de salida de dos canales puede ser, por ejemplo, la señal de salida de dos canales  $\widetilde{L}_1, \widetilde{L}_2$ , descrita con referencia a las Fig. 7 y 8, y puede ser conforme al segundo

formato de codificación  $F_2$  descrito con referencia a la Fig. 5. En otras palabras, los coeficientes de mezcla se pueden determinar de manera que: un primer canal  $\widetilde{L}_1$  de la señal de salida se aproxime a una combinación lineal del tercer grupo 501 de canales, y un segundo canal  $\widetilde{L}_2$  de la señal de salida se aproxima a una combinación lineal del cuarto grupo 502 de canales.

- 5 El método de decodificación 1000 puede comprender opcionalmente: recibir 1030 señalización que indica que la señal de mezcla descendente recibida  $L_1, L_2$  es conforme a uno del primer formato de codificación  $F_1$  y del segundo formato de codificación  $F_2$ , descritos con referencia a las Fig. 4 y 5, respectivamente. El tercer y cuarto grupos 501, 502 pueden estar predefinidos, y los coeficientes de mezcla se pueden determinar de manera que una única división de la señal de audio de cinco canales L, LS, LB, TFL, TBL en el tercer y cuarto grupos 501, 502 de canales,
- 10 aproximada por los canales de la señal de salida  $\widetilde{L}_1, \widetilde{L}_2$ , se mantenga para ambos formatos de codificación  $F_1, F_2$  posibles de la señal de mezcla descendente recibida. El método de decodificación 1000 puede comprender opcionalmente pasar 1070 la señal de mezcla descendente  $L_1, L_2$  a través como la señal de salida  $\widetilde{L}_1, \widetilde{L}_2$ , (y/o suprimir la contribución de la señal descorrelacionada a la señal de salida) en respuesta a la señalización que indica que la señal de mezcla descendente recibida es conforme al segundo formato de codificación  $F_2$ , dado que entonces el formato de codificación de la señal de mezcla descendente  $L_1, L_2$  recibida coincide con el formato de codificación
- 15 a ser proporcionado en la señal de salida  $\widetilde{L}_1, \widetilde{L}_2$ .

La Fig. 11 ilustra esquemáticamente un medio legible por ordenador 1100, según un ejemplo de realización. El medio legible por ordenador 1100 representa: la señal de mezcla descendente de dos canales  $L_1, L_2$  descrita con referencia a las Fig. 1 y 4; los parámetros de mezcla ascendente  $\alpha_{LU}$ , descritos con referencia a la Fig. 1, que

20 permiten la reconstrucción paramétrica de la señal de audio de cinco canales L, LS, LB, TFL, TBL en base a la señal de mezcla descendente  $L_1, L_2$ ; y los parámetros de mezcla  $\alpha_{LM}$ , descritos con referencia a la Fig. 1.

Se apreciará que aunque la sección de codificación 100 descrita con referencia a la Fig. 1 se configura para codificar la señal de audio de canales 11.1 según el primer formato de codificación  $F_1$ , y para proporcionar parámetros de mezcla  $\alpha_{LM}$  para proporcionar una señal de salida conforme al segundo formato de codificación  $F_2$ , se pueden proporcionar secciones de codificación similares que se configuran para codificar la señal de audio de canales 11.1 según uno cualquiera de los formatos de codificación  $F_1, F_2, F_3$ , y para proporcionar parámetros de mezcla para proporcionar una señal de salida conforme a uno cualquiera del primer formato  $F_1, F_2, F_3$ .

25

También se apreciará que aunque las secciones de decodificación 700, 900, descritas con referencia a las Fig. 7 y 9, se configuran para proporcionar una señal de salida conforme al segundo formato de codificación  $F_2$  en base a una señal de mezcla descendente conforme al primer formato de codificación  $F_1$ , se pueden proporcionar secciones de decodificación similares que se configuran para proporcionar una señal de salida conforme a uno cualquiera de los formatos de codificación  $F_1, F_2, F_3$  en base a una señal de mezcla descendente conforme a uno cualquiera de los formatos de codificación  $F_1, F_2, F_3$ .

30

Dado que el sexto grupo 602 de canales, descrito con referencia a la Fig. 6, incluye cuatro canales, se apreciará que proporcionando una señal de salida conforme al primer o segundo formatos de codificación  $F_1, F_2$  en base a una señal de mezcla descendente conforme al tercer formato de codificación  $F_3$ , puede incluir, por ejemplo: emplear más de un canal descorrelacionado; y/o emplear no más de uno de los canales de la señal de mezcla descendente como entrada a la sección de descorrelación.

35

Se apreciará que aunque los ejemplos descritos anteriormente se han formulado en términos de la señal de audio de canales 11.1 descrita con referencia a las Fig. 4 a 6, se pueden contemplar sistemas de codificación y sistemas de decodificación que incluyen cualquier número de secciones de codificación o secciones de decodificación, respectivamente, y que se pueden configurar para procesar señales de audio que comprenden cualquier número de señales de audio de M canales.

40

La Fig. 12 es un diagrama de bloques generalizado de una sección de decodificación 1200 para proporcionar una señal de salida de K canales  $\widetilde{L}_1, \dots, \widetilde{L}_K$  en base a una señal de mezcla descendente de dos canales  $L_1, L_2$  y metadatos asociados, según un ejemplo de realización. La sección de decodificación 1200 es similar a la de la sección de decodificación 700, descrita con referencia a la Fig. 7, excepto que la sección de decodificación 1200 proporciona una señal de salida de K canales  $\widetilde{L}_1, \dots, \widetilde{L}_K$ , donde  $2 \leq K < M$ , en lugar de una señal de salida de 2 canales  $\widetilde{L}_1, \widetilde{L}_2$ .

45

Más específicamente, la sección de decodificación 1200 se configura para recibir una señal de mezcla descendente de dos canales  $L_1, L_2$  que está asociada con metadatos, los metadatos que comprenden los parámetros de mezcla ascendente  $\alpha_{LU}$  para la reconstrucción paramétrica de una señal de audio de M canales en base a la señal de mezcla descendente  $L_1, L_2$ , donde  $M \geq 4$ . Un primer canal  $L_1$  de la señal de mezcla descendente  $L_1, L_2$  corresponde

50

a una combinación lineal (o suma) de un primer grupo de uno o más canales de la señal de audio de M canales (por ejemplo, el primer grupo 401 descrito con referencia a la Fig. 4). Un segundo canal  $L_2$  de la señal de mezcla descendente  $L_1, L_2$  corresponde a una combinación lineal (o suma) de un segundo grupo (por ejemplo, el segundo grupo 402, descrito con referencia a la Fig. 4) de uno o más canales de la señal de audio de M canales. El primer y segundo grupos constituyen una división de los M canales de la señal de audio de M canales. En otras palabras, el primer y segundo grupo son disjuntos y juntos incluyen todos los canales de la señal de audio de M canales.

La sección de decodificación 1200 se configura para recibir al menos una parte de los metadatos (por ejemplo, incluyendo los parámetros de mezcla ascendente  $\alpha_{LU}$ ), y para proporcionar la señal de salida de K canales  $\widetilde{L}_1, \dots,$

$\widetilde{L}_K$  en base a la señal de mezcla descendente  $L_1, L_2$  y los metadatos recibidos. La sección de decodificación 1200 comprende una sección de descorrelación 1210 configurada para recibir al menos un canal de la señal de mezcla descendente  $L_1, L_2$  y para emitir, en base al mismo, una señal descorrelacionada D. La sección de decodificación 1200 comprende además una sección de mezcla 1220 configurada para determinar un conjunto de coeficientes de

mezcla en base a los metadatos recibidos, y para formar la señal de salida  $\widetilde{L}_1, \dots, \widetilde{L}_K$  como una combinación lineal de la señal de mezcla descendente  $L_1, L_2$  y la señal descorrelacionada según los coeficientes de mezcla. La sección de mezcla 1220 se configura para determinar los coeficientes de mezcla de manera que cada uno de los K canales

de la señal de salida  $\widetilde{L}_1, \dots, \widetilde{L}_K$  se aproxime a una combinación lineal de un grupo de uno o más canales de la señal de audio de M canales. Los coeficientes de mezcla se determinan de manera que los grupos correspondientes

a los canales respectivos de la señal de salida  $\widetilde{L}_1, \dots, \widetilde{L}_K$  constituyan una división de los M canales de la señal de audio de M canales en K grupos de uno o más canales, y de manera que al menos dos de estos K grupos comprendan al menos un canal del primer grupo de canales de la señal de M canales (es decir, el grupo correspondiente al primer canal  $L_1$  de la señal de mezcla descendente).

La señal descorrelacionada D puede ser, por ejemplo, una señal de un único canal. Como se indica en la Fig. 12, la señal descorrelacionada D puede ser, por ejemplo, una señal de dos canales. En algunas realizaciones, la señal descorrelacionada D puede comprender más de dos canales.

La señal de M canales puede ser, por ejemplo, la señal de cinco canales L, LS, LB, TFL, TBL, descrita con referencia a la Fig. 4, y la señal de mezcla descendente  $L_1, L_2$  puede ser, por ejemplo, una representación de dos canales de la señal de cinco canales L, LS, LB, TFL, TBL según cualquiera de los formatos de codificación  $F_1, F_2, F_3$  descritos con referencia a las Fig. 4-6.

El sistema de decodificación de audio 800, descrito con referencia a la Fig. 8, puede comprender, por ejemplo, una o más secciones de decodificación 1200 del tipo descrito con referencia a la Fig. 12, en lugar de las secciones de decodificación 700 y 805, y el sistema de múltiples altavoces 812 puede incluir, por ejemplo, más de los cinco altavoces y un altavoz de graves descrito con referencia a la Fig. 8.

El sistema de decodificación de audio 800 se puede adaptar, por ejemplo, para realizar un método de decodificación de audio similar al método de decodificación de audio 1000, descrito con referencia a la Fig. 10, excepto que se proporciona una señal de salida de K canales en lugar de una señal de salida de dos canales.

Ejemplos de implementaciones de la sección de decodificación 1200 y del sistema de decodificación de audio 800 se describirán a continuación con referencia a las Fig. 12-16.

De manera similar a las Fig. 4-6, las Fig. 12-13 ilustran formas alternativas para dividir una señal de audio de canales 11.1 en grupos de uno o más canales.

Con el fin de representar la señal de audio de canales 11.1 (o canales 7.1 + 4, o canales 7.1.4) como una señal de audio de canales 7.1 (o canales 5.1 + 2 o canales 5.1.2), la colección de canales L, LS, LB, TFL, TBL, R, RS, RB, TFR, TBR, C, y LFE se puede dividir en grupos de canales representados por los canales respectivos. La señal de audio de cinco canales L, LS, LB, TFL, TBL se puede representar por una señal de tres canales  $L_1, L_2, L_3$ , mientras que la señal de audio de cinco canales R, RS, RB, TFR, TBR adicional se puede representar por una señal de tres canales  $R_1, R_2, R_3$  adicional. Los canales C y LFE se pueden mantener como canales separados también en la representación de canales 7.1 de la señal de audio de canales 11.1.

La Fig. 13 ilustra un cuarto formato de codificación  $F_4$  que proporciona una representación de canales 7.1 de la señal de audio de canales 11.1. En el cuarto formato de codificación  $F_4$ , la señal de audio de cinco canales L, LS, LB, TFL, TBL se divide en un primer grupo 1301 de canales que incluye solamente el canal L, un segundo grupo 1302 de canales que incluye los canales LS, LB y un tercer grupo 1303 de canales que incluye los canales TFL, TBL. Los canales  $L_1, L_2, L_3$  de la señal de tres canales  $L_1, L_2, L_3$  corresponden a combinaciones lineales (por ejemplo, sumas ponderadas o no ponderadas) de los grupos 1301, 1302, 1303 respectivos de canales. De manera similar, la señal de audio de cinco canales R, RS, RB, TFR, TBR adicional se divide en un primer grupo 1304 adicional que incluye el canal R, un segundo grupo 1305 adicional que incluye los canales RS, RB, y un tercer grupo 1306 adicional que

incluye los canales TFR, TBR. Los canales  $R_1, R_2, R_3$  de la señal de tres canales  $R_1, R_2, R_3$  adicional corresponden a combinaciones lineales (por ejemplo, sumas ponderadas o no ponderadas) de los grupos 1304, 1305, 1306 respectivos de canales adicionales.

5 Los inventores se han dado cuenta de que los metadatos asociados con una representación de canales 5.1 de la señal de audio de canales 11.1 según uno del primer, segundo y tercer formatos de codificación  $F_1, F_2, F_3$  se pueden emplear para generar una representación de canales 7.1 según el cuarto formato de codificación  $F_4$  sin reconstruir primero la señal original de canales 11.1. La señal de cinco canales L, LS, LB, TFL, TBL representa el plano de la mitad izquierda de la señal de audio de canales 11.1, y la señal de cinco canales R, RS, RB, TFR, TBR adicional representa el plano de la mitad derecha, y se pueden tratar de manera análoga.

10 Recuerde que dos canales  $x_4$  y  $x_5$  se pueden reconstruir a partir de la suma  $m_2 = x_4 + x_5$  usando la ecuación (3).

Si el segundo formato de codificación  $F_2$  se emplea para proporcionar una representación paramétrica de la señal de canales 11.1, y el cuarto formato de codificación  $F_4$  se desea en el lado del decodificador para representación de canales 7.1 del contenido de audio, entonces la aproximación dada por la ecuación (1) se puede aplicar una vez con

$$x_1 = TBL, x_2 = LS, x_3 = LB,$$

15 y una vez con

$$x_1 = TBR, x_2 = RS, x_3 = RB,$$

y la aproximación dada por la ecuación (3) se puede aplicar una vez con

$$x_4 = L, x_5 = TFL,$$

y una vez con

20  $x_4 = R, x_5 = TFR.$

Indicando la naturaleza aproximada de algunas de las cantidades del lado izquierdo (seis canales de la señal de salida) mediante tildes, tal aplicación de las ecuaciones (1) y (3) produce

$$\begin{bmatrix} \tilde{L}_1 \\ \tilde{R}_1 \\ C \\ \tilde{L}_2 \\ \tilde{R}_2 \\ \tilde{L}_3 \\ \tilde{R}_3 \end{bmatrix} = A \begin{bmatrix} L_1 \\ R_1 \\ C \\ L_2 \\ R_2 \\ D(L_1) \\ D(L_2) \\ D(R_1) \\ D(R_2) \end{bmatrix}, \quad (10)$$

donde

$$A = \begin{bmatrix} d_{1,L} & 0 & 0 & 0 & 0 & q_{1,L} & 0 & 0 & 0 \\ 0 & d_{1,R} & 0 & 0 & 0 & 0 & 0 & q_{1,R} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 - c_{1,L} & 0 & 0 & -p_{1,L} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 - c_{1,R} & 0 & 0 & 0 & -p_{1,R} \\ 1 - d_{1,L} & 0 & 0 & c_{1,L} & 0 & -q_{1,L} & p_{1,L} & 0 & 0 \\ 0 & 1 - d_{1,R} & 0 & 0 & c_{1,R} & 0 & 0 & -q_{1,R} & p_{1,R} \end{bmatrix}$$

25

y donde, según el cuarto formato de codificación  $F_4$ ,

$$\begin{aligned} \tilde{L}_1 &\approx L, & \tilde{L}_2 &\approx LS + LB, & \tilde{L}_3 &\approx TFL + TBL, \\ \tilde{R}_1 &\approx R & \tilde{R}_2 &\approx RS + RB, & \tilde{R}_3 &\approx TFR + TBR. \end{aligned}$$

30 En la matriz A anterior, los parámetros  $c_{1,L}, p_{1,L}$  y  $c_{1,R}, p_{1,R}$  son versiones del canal izquierdo y del canal derecho, respectivamente, de los parámetros de mezcla ascendente  $c_1, p_1$  de la ecuación (1), los parámetros  $d_{1,L}, q_{1,L}$  y  $d_{1,R}, q_{1,R}$  son versiones del canal izquierdo y del canal derecho, respectivamente, de los parámetros de mezcla  $d_1, q_1$  de la ecuación (3), y D denota un operador de descorrelación. De ahí que se pueda obtener una aproximación del cuarto formato de codificación  $F_4$  a partir del segundo formato de codificación  $F_2$  en base a parámetros de mezcla ascendente (por ejemplo, los parámetros de mezcla ascendente  $\alpha_{LU}, \alpha_{RU}$  descritos con referencia a las Fig. 1 y 2)

para la reconstrucción paramétrica de la señal de audio de canales 11.1 sin tener que reconstruir realmente la señal de audio de canales 11.1.

Dos casos de la sección de decodificación 1200, descrita con referencia a la Fig. 12 (con  $K = 3$ ,  $M = 5$  y una señal

descorrelacionada de dos canales  $D$ ), pueden proporcionar las señales de salida de tres canales  $\widetilde{L}_1, \widetilde{L}_2, \widetilde{L}_3$  y

5  $\widetilde{R}_1, \widetilde{R}_2, \widetilde{R}_3$  aproximando las señales de tres canales  $L_1, L_2, L_3$  y  $R_1, R_2, R_3$  del cuarto formato de codificación  $F_4$ .

Más específicamente, las secciones de mezcla 1220 de las secciones de decodificación 1200 pueden determinar los coeficientes de mezcla en base a los parámetros de mezcla ascendente según la matriz  $A$  de la ecuación (10). Un sistema de decodificación de audio similar al sistema de decodificación de audio 800, descrito con referencia a la Fig. 8, puede emplear dos de tales secciones de decodificación 1200 para proporcionar una representación de

10 canales 7.1 de la señal de audio 11.1 para la reproducción de canales 7.1.

Si el primer formato de codificación  $F_1$  se emplea para proporcionar una representación paramétrica de la señal de canales 11.1 y el cuarto formato de codificación  $F_4$  se desea en el lado del decodificador para la representación del contenido de audio, entonces la aproximación dada por la ecuación (1) se puede aplicar una vez con

$$x_1 = L, x_2 = LS, x_3 = LB,$$

15 y una vez con

$$x_1 = R, x_2 = RS, x_3 = RB.$$

Indicando la naturaleza aproximada de algunas de las cantidades del lado izquierdo (seis canales de la señal de salida) mediante tildes, tal aplicación de la ecuación (1) produce

$$\begin{bmatrix} \widetilde{L}_1 \\ \widetilde{R}_1 \\ C \\ \widetilde{L}_2 \\ \widetilde{R}_2 \\ \widetilde{L}_3 \\ \widetilde{R}_3 \end{bmatrix} = \begin{bmatrix} c_{1,L} & 0 & 0 & 0 & 0 & p_{1,L} & 0 & 0 & 0 \\ 0 & c_{1,R} & 0 & 0 & 0 & 0 & 0 & p_{1,R} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 - c_{1,L} & 0 & 0 & 0 & 0 & -p_{1,L} & 0 & 0 & 0 \\ 0 & 1 - c_{1,R} & 0 & 0 & 0 & 0 & 0 & -p_{1,R} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} L_1 \\ R_1 \\ C \\ L_2 \\ R_2 \\ D(L_1) \\ D(L_2) \\ D(R_1) \\ D(R_2) \end{bmatrix} \quad (11)$$

20 donde, según el cuarto formato de codificación  $F_4$ ,

$$\widetilde{L}_1 \approx L, \quad \widetilde{L}_2 \approx LS + LB, \quad \widetilde{L}_3 = TFL + TBL \quad (\text{no aproximado}),$$

$$\widetilde{R}_1 \approx R, \quad \widetilde{R}_2 \approx RS + RB, \quad \widetilde{R}_3 = TFR + TBR \quad (\text{no aproximado}).$$

En la ecuación anterior (11), los parámetros  $c_{1,L}$ ,  $p_{1,L}$  y  $c_{1,R}$ ,  $p_{1,R}$  son versiones del canal izquierdo y del canal derecho, respectivamente, de los parámetros  $c_1$ ,  $p_1$  de la ecuación (1), y  $D$  denota un operador de descorrelación. De ahí que se pueda obtener una aproximación del cuarto formato de codificación  $F_4$  a partir del primer formato de codificación  $F_1$  en base a parámetros de mezcla ascendente para la reconstrucción paramétrica de la señal de audio de canales 11.1, sin tener que reconstruir realmente la señal de audio de canales 11.1.

25

Dos casos de la sección de decodificación 1200, descrita con referencia a la Fig. 12 (con  $K = 3$ ,  $M = 5$ ), pueden

proporcionar las señales de salida de tres canales  $\widetilde{L}_1, \widetilde{L}_2, \widetilde{L}_3$  y  $\widetilde{R}_1, \widetilde{R}_2, \widetilde{R}_3$  aproximando las señales de tres

30 canales  $L_1, L_2, L_3$  y  $R_1, R_2, R_3$  del cuarto formato de codificación  $F_4$ . Más específicamente, las secciones de mezcla 1220 de las secciones de decodificación pueden determinar los coeficientes de mezcla en base a los parámetros de mezcla ascendente según la ecuación (11). Un sistema de decodificación de audio similar al sistema de decodificación de audio 800, descrito con referencia a la Fig. 8, puede emplear dos de tales secciones de decodificación 1200 para proporcionar una representación de canales 7.1 de la señal de audio 11.1 para la

35 reproducción de canales 7.1.

Como se puede ver en la ecuación (11), solamente se necesitan realmente dos canales descorrelacionados. Aunque los canales descorrelacionados  $D(L_2)$  y  $D(R_2)$  no son necesarios para proporcionar el cuarto formato de codificación  $F_4$  a partir del primer formato de codificación  $F_1$ , tales descorrelacionadores se pueden mantener, por ejemplo, ejecutándose (o ser mantenidos activos) de todas formas, de modo que los almacenadores temporales/memorias de los descorrelacionadores se mantengan actualizados y disponibles en caso de que el formato de codificación de la señal de mezcla descendente cambie a, por ejemplo, el segundo formato de codificación  $F_2$ . Recordemos que se

40

emplean cuatro canales descorrelacionados cuando se proporciona el cuarto formato de codificación  $F_4$  a partir del segundo formato de codificación  $F_2$  (véase la ecuación (10) y la matriz A asociada).

5 Si el tercer formato de codificación  $F_3$  se emplea para proporcionar una representación paramétrica de la señal de audio de canales 11.1, y el cuarto formato de codificación  $F_4$  se desea en el lado del decodificador para la representación del contenido de audio, relaciones similares a las presentadas en las ecuaciones (10) y (11) se pueden derivar usando las mismas ideas. Un sistema de decodificación de audio similar al sistema de decodificación de audio 800, descrito con referencia a la Fig. 8, puede emplear dos secciones de decodificación 1200 para proporcionar una representación de 7.1 canales de la señal de audio 11.1 según el cuarto formato de codificación  $F_4$ .

10 Con el fin de representar la señal de audio de canales 11.1 como una señal de audio de canales 9.1 (o canales 5.1 + 4, o canales 5.1.4), la colección de canales L, LS, LB, TFL, TBL, R, RS, RB, TFR, TBR, C, y LFE se puede dividir en grupos de canales representados por los canales respectivos. La señal de audio de cinco canales L, LS, LB, TFL, TBL se puede representar por una señal de cuatro canales  $L_1, L_2, L_3, L_4$ , mientras que la señal de audio de cinco canales R, RS, RB, TFR, TBR adicional se puede representar por una señal de cuatro canales  $R_1, R_2, R_3, R_4$  adicional. Los canales C y LFE se pueden mantener como canales separados también en la representación de canales 9.1 de la señal de audio de canales 11.1.

15 La Fig. 14 ilustra un quinto formato de codificación  $F_5$  que proporciona una representación de canales 9.1 de una señal de audio de canales 11.1. En el quinto formato de codificación, la señal de audio de cinco canales L, LS, LB, TFL, TBL se divide en un primer grupo 1401 de canales que incluyen solamente el canal L, un segundo grupo 1402 de canales que incluyen los canales LS, LB, un tercer grupo 1403 de canales que incluyen solamente el canal TFL y un cuarto grupo 1404 de canales que incluye solamente el canal TBL. Los canales  $L_1, L_2, L_3, L_4$  de la señal de cuatro canales  $L_1, L_2, L_3, L_4$  corresponden a combinaciones lineales (por ejemplo, sumas ponderadas o no ponderadas) de los grupos 1401, 1402, 1403, 1404 respectivos de uno o más canales. De manera similar, la señal de audio de cinco canales R, RS, RB, TFR, TBR adicional se divide en un primer grupo 1405 adicional que incluye el canal R, un segundo grupo 1406 adicional que incluye los canales RS, RB, un tercer grupo 1407 adicional que incluye el canal TFR, y un cuarto grupo 1408 adicional que incluye el canal TBR. Los canales  $R_1, R_2, R_3, R_4$  de la señal de cuatro canales  $R_1, R_2, R_3, R_4$  adicional corresponden a combinaciones lineales (por ejemplo, sumas ponderadas o no ponderadas) de los grupos 1405, 1406, 1407, 1408 adicionales respectivos de uno o más canales.

20 Los inventores se han dado cuenta de que los metadatos asociados con una representación de canales 5.1 de la señal de audio de canales 11.1 según uno de los formatos de codificación  $F_1, F_2, F_3$  se pueden emplear para generar una representación de canales 9.1 según el quinto formato de codificación  $F_5$  sin reconstruir primero la señal de canales 11.1 original. La señal de cinco canales L, LS, LB, TFL, TBL que representa el plano de la mitad izquierda de la señal de audio de canales 11.1, y la señal de cinco canales R, RS, RB, TFR, TBR adicional que representa el plano de la mitad derecha, se pueden tratar de manera análoga.

25 Si el segundo formato de codificación  $F_2$  se emplea para proporcionar una representación paramétrica de la señal de canales 11.1, y el quinto formato de codificación  $F_5$  se desea en el lado del decodificador para representar el contenido de audio, entonces la aproximación proporcionada por la ecuación (1) se puede aplicar una vez con

$$x_1 = TBL, x_2 = LS, x_3 = LB,$$

y una vez con

$$x_1 = TBR, x_2 = RS, x_3 = RB,$$

40 y la aproximación de la ecuación (3) se puede aplicar una vez con

$$x_4 = L, x_5 = TFL,$$

y una vez con

$$x_4 = R, x_5 = TFR.$$

45 Indicando la naturaleza aproximada de algunas de las cantidades del lado izquierdo (ocho canales de la señal de salida) mediante tildes, tal aplicación de las ecuaciones (1) y (3) produce

$$\begin{bmatrix} \tilde{L}_1 \\ \tilde{R}_1 \\ C \\ \tilde{L}_2 \\ \tilde{R}_2 \\ \tilde{L}_3 \\ \tilde{R}_3 \\ \tilde{L}_4 \\ \tilde{R}_4 \end{bmatrix} = A \begin{bmatrix} L_1 \\ R_1 \\ C \\ L_2 \\ R_2 \\ D(L_1) \\ D(L_2) \\ D(R_1) \\ D(R_2) \end{bmatrix}, \quad (12)$$

donde

$$A = \begin{bmatrix} d_{1,L} & 0 & 0 & 0 & 0 & q_{1,L} & 0 & 0 & 0 & 0 \\ 0 & d_{1,R} & 0 & 0 & 0 & 0 & 0 & q_{1,R} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 - c_{1,L} & 0 & 0 & -p_{1,L} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 - c_{1,R} & 0 & 0 & 0 & 0 & -p_{1,R} \\ 1 - d_{1,L} & 0 & 0 & 0 & 0 & -q_{1,L} & 0 & 0 & 0 & 0 \\ 0 & 1 - d_{1,R} & 0 & 0 & 0 & 0 & 0 & -q_{1,R} & 0 & 0 \\ 0 & 0 & 0 & c_{1,L} & 0 & 0 & p_{1,L} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{1,R} & 0 & 0 & 0 & 0 & p_{1,R} \end{bmatrix},$$

y donde, según el quinto formato de codificación  $F_5$ ,

$$\begin{aligned} 5 \quad \tilde{L}_1 &\approx L, & \tilde{L}_2 &\approx LS + LB, & \tilde{L}_3 &\approx TFL, & \tilde{L}_4 &\approx TBL \\ \tilde{R}_1 &\approx R & \tilde{R}_2 &\approx RS + RB, & \tilde{R}_3 &\approx TFR, & \tilde{R}_4 &\approx TBR. \end{aligned}$$

En la matriz A anterior, los parámetros  $c_{1,L}$ ,  $p_{1,L}$  y  $c_{1,R}$ ,  $p_{1,R}$  son versiones del canal izquierdo y del canal derecho, respectivamente, de los parámetros de mezcla ascendente  $c_1$ ,  $p_1$  de la ecuación (1),  $d_{1,L}$ ,  $q_{1,L}$  y  $d_{1,R}$ ,  $q_{1,R}$  son versiones del canal izquierdo y del canal derecho, respectivamente, de los parámetros de mezcla ascendente  $d_1$ ,  $q_1$  de la ecuación (3), y D denota un operador de descorrelación. De ahí que se pueda obtener una aproximación del quinto formato de codificación  $F_5$  a partir del segundo formato de codificación  $F_2$  en base a parámetros de mezcla ascendente para la reconstrucción paramétrica de la señal de audio de canales 11.1 sin tener que reconstruir realmente la señal de audio de canales 11.1.

Dos casos de la sección de decodificación 1200, descrita con referencia a la Fig. 12 (con  $K = 4$ ,  $M = 5$  y una señal descorrelacionada de dos canales D), pueden proporcionar las señales de salida de cuatro canales  $\tilde{L}_1, \tilde{L}_2, \tilde{L}_3, \tilde{L}_4$  y  $\tilde{R}_1, \tilde{R}_2, \tilde{R}_3, \tilde{R}_4$  aproximando las señales de cuatro canales  $L_1, L_2, L_3, L_4$  y  $R_1, R_2, R_3, R_4$  del quinto formato de codificación  $F_5$ . Más específicamente, las secciones de mezcla 1220 de las secciones de decodificación 1200 pueden determinar los coeficientes de mezcla en base a los parámetros de mezcla ascendente según la ecuación (12). Un sistema de decodificación de audio similar al sistema de decodificación de audio 800, descrito con referencia a la Fig. 8, puede emplear dos de tales secciones de decodificación 1200 para proporcionar una representación de canales 9.1 de la señal de audio 11.1 para la reproducción de canales 9.1.

Si el primer  $F_1$  o el tercer  $F_3$  formato de codificación se emplea para proporcionar una representación paramétrica de la señal de audio de canales 11.1, y el quinto formato de codificación  $F_5$  se desea en el lado del decodificador para representar el contenido de audio, relaciones similares a la relación presentada en la ecuación (12) se pueden derivar usando las mismas ideas.

Las Fig. 15-16 ilustran formas alternativas para dividir una señal de audio de canales 13.1 (o canales 9.1 + 4 o canales 9.1.4) en grupos de canales para representar la señal de audio de canales 13.1 como una señal de audio de canales 5.1, y una señal de canales 7.1, respectivamente.

La señal de audio de canales 13.1 comprende los canales LW (ancho izquierdo), LSCRN (pantalla izquierda), LS (lado izquierdo), LB (parte posterior izquierda), TFL (parte superior delantera izquierda), TBL (parte superior posterior izquierda), RW (ancho derecho), RSCRN (pantalla derecha), RS (lado derecho), RB (parte posterior derecha), TFR (parte superior delantera derecha), TBR (parte superior posterior derecha), C (centro) y LFE (efectos de baja frecuencia). Los seis canales LW, LSCRN, LS, LB, TFL y TBL forman una señal de audio de seis canales que representa un espacio de la mitad izquierda en un entorno de reproducción de la señal de audio de canales 13.1. Los cuatro canales LW, LSCRN, LS y LB representan diferentes direcciones horizontales en el entorno de reproducción y los dos canales TFL y TBL representan direcciones separadas verticalmente de las de los cuatro

canales LW, LSCRN, LS y LB. Los dos canales TFL y TBL, por ejemplo, se pueden destinar a la reproducción en altavoces de techo. De manera similar, los seis canales RW, RSCRN, RS, RB, TFR y TBR forman una señal de audio de seis canales adicional que representa un espacio de la mitad derecha del entorno de reproducción, los cuatro canales RW, RSCRN, RS y RB que representan diferentes direcciones horizontales en el entorno de reproducción y los dos canales TFR y TBR que representan direcciones separadas verticalmente de las de los cuatro canales RW, RSCRN, RS y RB.

La Fig. 15 ilustra un sexto formato de codificación  $F_6$ , en el que la señal de audio de seis canales LW, LSCRN, LS, LB, TFL, TBL se divide en un primer grupo 1501 de canales LW, LSCRN, TFL y un segundo grupo 1502 de canales LS, LB, TBL, y en el que la señal de audio de seis canales RW, RSCRN, RS, RB, TFR, TBR adicional se divide en un primer grupo 1503 adicional de canales RW, RSCRN, TFR y un segundo grupo 1504 adicional de canales RS, RB, TBR. Los canales  $L_1, L_2$  de una señal de mezcla descendente de dos canales  $L_1, L_2$  corresponden a combinaciones lineales (por ejemplo, sumas ponderadas o no ponderadas) de los grupos 1501, 1502 respectivos de canales. De manera similar, los canales  $R_1, R_2$  de una señal de mezcla descendente de dos canales  $R_1, R_2$  corresponden a combinaciones lineales (por ejemplo, sumas ponderadas o no ponderadas) de los grupos 1503, 1504 de canales adicionales respectivos.

La Fig. 16 ilustra un séptimo formato de codificación  $F_7$ , en el que la señal de audio de seis canales LW, LSCRN, LS, LB, TFL, TBL se divide en un primer grupo 1601 de canales LW, LSCRN, en un segundo grupo 1602 de canales LS, LB y en un tercer grupo 1603 de canales TFL, TBL, y en el que la señal de audio de seis canales RW, RSCRN, RS, RB, TFR, TBR adicional se divide en un primer grupo 1604 adicional de canales RW, RSCRN, un segundo grupo 1605 adicional de canales RS, RB y un tercer grupo 1606 adicional de canales TFR, TBR. Tres canales  $L_1, L_2, L_3$  corresponden a combinaciones lineales (por ejemplo, sumas ponderadas o no ponderadas) de los grupos 1601, 1602, 1603 respectivos de canales. De manera similar, tres canales  $R_1, R_2, R_3$  adicionales corresponden a combinaciones lineales (por ejemplo, sumas ponderadas o no ponderadas) de los grupos 1604, 1605, 1606 de canales adicionales respectivos.

Los inventores se han dado cuenta de que metadatos asociados con una representación de canales 5.1 de la señal de audio de canales 13.1 según el sexto formato de codificación  $F_6$  se pueden emplear para generar una representación de canales 7.1 según el séptimo formato de codificación  $F_7$  sin reconstruir primero la señal de canales 13.1 original. La señal de seis canales LW, LSCRN, LS, LB, TFL, TBL que representa el plano de la mitad izquierda de la señal de audio de canales 13.1 y la señal de seis canales RW, RSCRN, RS, RB, TFR, TBR adicional que representa el plano de la mitad derecha, se pueden tratar de forma análoga.

Recordemos que dos canales  $x_4$  y  $x_5$  se pueden reconstruir a partir de la suma  $m_2 = x_4 + x_5$  usando la ecuación (3).

Si el sexto formato de codificación  $F_6$  se emplea para proporcionar una representación paramétrica de la señal de canales 13.1, y el séptimo formato de codificación  $F_7$  se desea en el lado del decodificador para canales 7.1 (o canales 5.1 + 2 o canales 5.1.2) que representa el contenido de audio, entonces la aproximación dada por la ecuación (1) se puede aplicar cuatro veces, una vez con

$$x_1 = TBL, x_2 = LS, x_3 = LB,$$

y una vez con

$$x_1 = TBR, x_2 = RS, x_3 = RB,$$

y una vez con

$$x_1 = TFL, x_2 = LW, x_3 = LSCRN,$$

y una vez con

$$x_1 = TFR, x_2 = RW, x_3 = RSCRN,$$

Indicando la naturaleza aproximada de algunas de las cantidades del lado izquierdo (seis canales de la señal de salida) mediante tildes, tal aplicación de la ecuación (1) produce

$$\begin{bmatrix} \tilde{L}_1 \\ \tilde{R}_1 \\ C \\ \tilde{L}_2 \\ \tilde{R}_2 \\ \tilde{L}_3 \\ \tilde{R}_3 \end{bmatrix} = A \begin{bmatrix} L_1 \\ R_1 \\ C \\ L_2 \\ R_2 \\ D(L_1) \\ D(L_2) \\ D(R_1) \\ D(R_2) \end{bmatrix}, \quad (13)$$

donde

$$A = \begin{bmatrix} 1 - c_{1,L} & 0 & 0 & 0 & 0 & 0 & -p_{1,L} & 0 & 0 & 0 \\ 0 & 1 - c_{1,R} & 0 & 0 & 0 & 0 & 0 & 0 & -p_{1,R} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 - c'_{1,L} & 0 & 0 & 0 & -p'_{1,L} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 - c'_{1,R} & 0 & 0 & 0 & 0 & -p'_{1,R} \\ c_{1,L} & 0 & 0 & c'_{1,L} & 0 & p_{1,L} & p'_{1,L} & 0 & 0 & 0 \\ 0 & c_{1,R} & 0 & 0 & c'_{1,R} & 0 & 0 & p_{1,R} & p'_{1,R} & 0 \end{bmatrix}$$

y donde, según el séptimo formato de codificación F<sub>7</sub>,

$$\widetilde{L}_1 \approx LW + LSCRN, \quad \widetilde{L}_2 \approx LS + LB, \quad \widetilde{L}_3 \approx TFL + TBL,$$

$$\widetilde{R}_1 \approx RW + RSCN \quad \widetilde{R}_2 \approx RS + RB, \quad \widetilde{R}_3 \approx TFR + TBR.$$

En la matriz A anterior, los parámetros  $c_{1,L}$ ,  $p_{1,L}$  y  $c'_{1,L}$ ,  $p'_{1,L}$  son dos casos diferentes de los parámetros de mezcla ascendente  $c_1$ ,  $p_1$  de la ecuación (1) para el lado izquierdo,  $c_{1,R}$ ,  $p_{1,R}$  y  $c'_{1,R}$ ,  $p'_{1,R}$  son dos casos diferentes de los parámetros de mezcla ascendente  $c_1$ ,  $p_1$  de la ecuación (1) para el lado derecho, y D denota un operador de descorrelación. De ahí que se pueda obtener una aproximación del séptimo formato de codificación F<sub>7</sub> a partir del sexto formato de codificación F<sub>6</sub> en base a los parámetros de mezcla ascendente para la reconstrucción paramétrica de la señal de audio de canales 13.1 sin tener que reconstruir realmente la señal de audio de canales 13.1.

Dos casos de la sección de decodificación 1200, descrita con referencia a la Fig. 12 (con K = 3, M = 6, y una señal descorrelacionada de dos canales D), pueden proporcionar las señales de salida de tres canales  $\widetilde{L}_1, \widetilde{L}_2, \widetilde{L}_3$  y  $\widetilde{R}_1, \widetilde{R}_2, \widetilde{R}_3$  aproximando las señales de tres canales L<sub>1</sub>, L<sub>2</sub>, L<sub>3</sub> y R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub> del séptimo formato de codificación F<sub>7</sub> en base a las señales de mezcla descendente de dos canales generados en un lado del codificador según el sexto formato de codificación F<sub>6</sub>. Más específicamente, las secciones de mezcla 1220 de las secciones de decodificación 1200 pueden determinar coeficientes de mezcla en base a los parámetros de mezcla ascendente según la matriz A de la ecuación (13). Un sistema de decodificación de audio similar al sistema de decodificación de audio 800, descrito con referencia a la Fig. 8, puede emplear dos de tales secciones de decodificación 1200 para proporcionar una representación de canales 7.1 de la señal de audio 13.1 para la reproducción de canales 7.1.

Como se puede ver en las ecuaciones (10)-(13) (y las matrices A asociadas), si dos canales de la señal de salida (por ejemplo, los canales  $\widetilde{L}_1$  y  $\widetilde{L}_2$  en la ecuación (11)) reciben contribuciones del mismo canal descorrelacionado (por ejemplo, D(L<sub>1</sub>) en la ecuación (11)), entonces estas dos contribuciones tienen igual magnitud, pero de signos opuestos (por ejemplo, indicados por los coeficientes de mezcla  $p_{1,L}$  y  $-p_{1,L}$  en la ecuación (11)).

Como se puede ver en las ecuaciones (10)-(13) (y las matrices A asociadas), si dos canales de la señal de salida (por ejemplo, los canales  $\widetilde{L}_1$  y  $\widetilde{L}_2$  en la ecuación (11)) reciben contribuciones del mismo canal de mezcla descendente (por ejemplo, el canal L<sub>1</sub> en la ecuación (11)), entonces la suma de los dos coeficientes de mezcla que controlan estas dos contribuciones (por ejemplo, los coeficientes de mezcla  $c_{1,L}$  y  $1 - c_{1,L}$  en la ecuación (11)) tiene el valor 1.

Como se ha descrito anteriormente con referencia a las Fig. 12-16, la sección de decodificación 1200 puede proporcionar una señal de salida de K canales  $\widetilde{L}_1, \dots, \widetilde{L}_K$  en base a una señal de mezcla descendente de dos canales L<sub>1</sub>, L<sub>2</sub> y los parámetros de mezcla ascendente  $\alpha_{LU}$ . Los parámetros de mezcla ascendente  $\alpha_{LU}$  se pueden adaptar para la reconstrucción paramétrica de una señal de audio de M canales original, y la sección de mezcla 1220 de la sección de decodificación 1200 puede ser capaz de calcular parámetros de mezcla adecuados, en base a

los parámetros de mezcla ascendente  $\alpha_{LU}$ , para proporcionar la señal de salida de K canales  $\widetilde{L}_1, \dots, \widetilde{L}_K$  sin reconstruir la señal de audio de M canales.

En algunos ejemplos de realizaciones, los parámetros de mezcla  $\alpha_{LM}$  dedicados se pueden enviar desde el lado del codificador para facilitar la provisión de la señal de salida de K canales  $\widetilde{L}_1, \dots, \widetilde{L}_K$  en el lado del decodificador.

Por ejemplo, la sección de decodificación 1200 se puede configurar de manera similar a la sección de decodificación 900 descrita anteriormente con referencia a la Fig. 9.

Por ejemplo, la sección de decodificación 1200 puede recibir los parámetros de mezcla  $\alpha_{LM}$  en forma de los elementos (o coeficientes de mezcla) de una o más de las matrices de mezcla mostradas en las ecuaciones (10)-

(13) (es decir, las matrices denotadas A). En tal ejemplo, puede no ser necesario para la sección de decodificación 1200 calcular cualquiera de los elementos en las matrices de mezcla en las ecuaciones (10)-(13).

5 Se pueden contemplar ejemplos de realizaciones en las que la sección de análisis 120, descrita con referencia a la Fig. 1 (y de manera similar, la sección de análisis 203 adicional, descrita con referencia a la Fig. 2), determina los parámetros de mezcla  $\alpha_{LM}$  para obtener, en base a la señal de mezcla descendente  $L_1, L_2$ , una señal de salida de K canales, donde  $2 \leq K < M$ . Los parámetros de mezcla  $\alpha_{LM}$  se pueden proporcionar, por ejemplo, en forma de los elementos (o coeficientes de mezcla) de una o más de las matrices de mezcla de las ecuaciones (10)-(13) (es decir, las matrices denotadas A).

10 Se pueden proporcionar, por ejemplo, múltiples conjuntos de parámetros de mezcla  $\alpha_{LM}$ , donde los conjuntos respectivos de parámetros de mezcla  $\alpha_{LM}$  se destinan a diferentes tipos de representación en el lado del decodificador. Por ejemplo, el sistema de codificación de audio 200, descrito anteriormente con referencia a la Fig. 2, puede proporcionar un flujo de bits B en el que se proporciona una representación de mezcla descendente 5.1 de una señal de audio de canales 11.1 original, y en qué conjuntos de parámetros de mezcla  $\alpha_{LM}$  se pueden proporcionar para la representación de canales 5.1 (según el primer, segundo y/o tercer formatos de codificación  $F_1, F_2, F_3$ ), para la representación de canales 7.1 (según el cuarto formato de codificación  $F_4$ ) y/o para representación de canales 9.1 (según el quinto formato de codificación  $F_5$ ).

15 El método de codificación de audio 300, descrito con referencia a la Fig. 3 puede incluir, por ejemplo, determinar 340 parámetros de mezcla  $\alpha_{LM}$  para obtener, en base a la señal de mezcla descendente  $L_1, L_2$ , una señal de salida de K canales, donde  $2 \leq K < M$ .

20 Se pueden contemplar ejemplos de realizaciones en las que el medio legible por ordenador 1100, descrito con referencia a la Fig. 11, representa: una señal de mezcla descendente de dos canales (por ejemplo, la señal de mezcla descendente de dos canales  $L_1, L_2$  descrita con referencia a las Fig. 1 y 4); los parámetros de mezcla ascendente (por ejemplo, los parámetros de mezcla ascendente  $\alpha_{LU}$ , descritos con referencia a la Fig. 1) que permiten la reconstrucción paramétrica de una señal de audio de M canales (por ejemplo, la señal de audio de cinco canales L, LS, LB, TFL, TBL) en base a la señal de mezcla descendente; y los parámetros de mezcla  $\alpha_{LM}$  que permiten la provisión de una señal de salida de K canales en base a la señal de mezcla descendente. Como se ha descrito anteriormente,  $M \geq 4$  y  $2 \leq K < M$ .

25 Se apreciará que aunque los ejemplos descritos anteriormente se han formulado en términos de señales de audio originales con  $M = 5$  y  $M = 6$  canales, y señales de salida con  $K = 2, K = 3$  y  $K = 4$  canales, sistemas de codificación similares (y secciones de codificación) y los sistemas de decodificación (y las secciones de decodificación) se pueden contemplar para cualquier M y K que satisfaga  $M \geq 4$  y  $2 \leq K < M$ .

V. Equivalentes, extensiones, alternativas y varios.

35 Incluso aunque la presente descripción describe y representa ejemplos de realizaciones específicos, la invención no está restringida a estos ejemplos específicos. Se pueden hacer modificaciones y variaciones a los ejemplos de realizaciones anteriores sin apartarse del alcance de la invención, que se define solamente por las reivindicaciones que se acompañan.

40 En las reivindicaciones, la palabra "que comprende" no excluye otros elementos o pasos, y el artículo indefinido "un", "uno" o "una" no excluye una pluralidad. El mero hecho de que ciertas medidas se reciten en reivindicaciones dependientes mutuamente diferentes, no indica que una combinación de estas medidas no se pueda usar con ventaja. Cualquier signo de referencia que aparezca en las reivindicaciones no se ha de entender como limitante de su alcance.

45 Los dispositivos y métodos descritos anteriormente se pueden implementar como software, microprogramas, hardware o una combinación de los mismos. En una implementación de hardware, la división de tareas entre las unidades funcionales a las que se hace referencia en la descripción anterior no necesariamente corresponde a la división en unidades físicas; por el contrario, un componente físico puede tener múltiples funcionalidades, y una tarea se puede llevar a cabo de una forma distribuida, por varios componentes físicos en cooperación. Ciertos componentes o todos los componentes se pueden implementar como software ejecutado por un procesador digital, procesador de señales o microprocesador, o se pueden implementar como hardware o como un circuito integrado de aplicaciones específicas. Tal software se puede distribuir en medios legibles por ordenador, que pueden comprender medios de almacenamiento en ordenador (o medios no transitorios) y medios de comunicación (o medios transitorios). Como es bien sabido por un experto en la técnica, el término medios de almacenamiento en ordenador incluye tanto medios volátiles como no volátiles, extraíbles y no extraíbles implementados en cualquier método o tecnología para almacenamiento de información, tal como instrucciones legibles por ordenador, estructuras de datos, módulos de programas u otros datos. Los medios de almacenamiento en ordenador incluyen, pero no se limitan a, RAM, ROM, EEPROM, memoria rápida u otra tecnología de memoria, CD-ROM, discos versátiles digitales (DVD) u otro almacenamiento en disco óptico, casetes magnéticos, cinta magnética, almacenamiento en disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro medio que se pueda usar para almacenar la información deseada y a los que se puede acceder mediante un ordenador. Además, es bien sabido por los expertos

que los medios de comunicación típicamente incorporan instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos en una señal de datos modulada, tal como una onda portadora u otro mecanismo de transporte, e incluye cualquier medio de entrega de información.

**REIVINDICACIONES**

1. Un método de decodificación de audio (1000) que comprende:

5 recibir (1010) una señal de mezcla descendente de dos canales ( $L_1, L_2$ ), que está asociada con metadatos, los metadatos que comprenden parámetros de mezcla ascendente ( $\alpha_{LU}$ ) para la reconstrucción paramétrica de una señal de audio de M canales (L, LS, LB, TFL, TBL) en base a la señal de mezcla descendente, donde  $M \geq 4$ ;

recibir (1020) al menos una parte de dichos metadatos;

generar (1040) una señal descorrelacionada (D) en base a al menos un canal de la señal de mezcla descendente;

determinar (1050) un conjunto de coeficientes de mezcla en base a los metadatos recibidos; y

10 formar (1060) una señal de salida de K canales ( $\tilde{L}_1, \dots, \tilde{L}_K$ ) como una combinación lineal de la señal de mezcla descendente y la señal descorrelacionada según los coeficientes de mezcla, en donde  $2 \leq K < M$ ,

caracterizado por que

15 los coeficientes de mezcla se determinan de manera que una suma de un coeficiente de mezcla que controla una contribución del primer canal de la señal de mezcla descendente a un canal de la señal de salida, y un coeficiente de mezcla que controla una contribución del primer canal de la señal de mezcla descendente a otro canal de la señal de salida, tenga el valor 1,

en donde, si la señal de mezcla descendente representa la señal de audio de M canales según un primer formato de codificación ( $F_1$ ) en el que:

20 un primer canal ( $L_1$ ) de la señal de mezcla descendente corresponde a una cierta combinación lineal de un primer grupo (401) de uno o más canales de la señal de audio de M canales;

un segundo canal ( $L_2$ ) de la señal de mezcla descendente corresponde a una cierta combinación lineal de un segundo grupo (402) de uno o más canales de la señal de audio de M canales; y

el primer y segundo grupos constituyen una cierta división de los M canales de la señal de audio de M canales,

25 entonces, la señal de salida de K canales representa la señal de audio de M canales según un segundo formato de codificación ( $F_2, F_4$ ) en el que:

cada uno de los K canales de la señal de salida se aproxima a una combinación lineal de un grupo de uno o más canales de la señal de audio de M canales;

los grupos correspondientes a los canales respectivos de la señal de salida constituyen una división de los M canales de la señal de audio de M canales en K grupos (501-502, 1301- 1303) de uno o más canales; y

30 al menos dos de los K grupos comprenden al menos un canal de dicho primer grupo.

2. El método de decodificación de audio de la reivindicación 1, en donde  $K = 2, K = 3$  o  $K = 4$ , y/o en donde  $M = 5$  o  $M = 6$ .

35 3. El método de decodificación de audio de cualquiera de las reivindicaciones precedentes, en donde los metadatos recibidos incluyen los parámetros de mezcla ascendente y en donde los coeficientes de mezcla se determinan procesando los parámetros de mezcla ascendente.

4. El método de decodificación de audio de cualquiera de las reivindicaciones precedentes, en donde:

en el primer formato de codificación, cada uno de los canales de la señal de audio de M canales se asocia con una ganancia distinta de cero que controla una contribución de este canal a una de las combinaciones lineales a las que corresponden los canales de la señal de mezcla descendente;

40 en el segundo formato de codificación, cada uno de los canales de la señal de audio de M canales se asocia con una ganancia distinta de cero que controla una contribución de este canal a una de las combinaciones lineales aproximadas por los canales de la señal de salida; y

45 para cada uno de los canales de la señal de audio de M canales, la ganancia distinta de cero asociada con el canal en el primer formato de codificación coincide con la ganancia distinta de cero asociada con el canal en el segundo formato de codificación.

5. El método de decodificación de audio de cualquiera de las reivindicaciones precedentes, en donde la señal descorrelacionada es una señal de dos canales, y en donde dicha señal de salida se forma incluyendo no más de dos canales de señal descorrelacionados en dicha combinación lineal de la señal de mezcla descendente y la señal descorrelacionada.
- 5 6. El método de decodificación de audio de la reivindicación 5, en donde  $K = 3$ , y en donde formar la señal de salida equivale a una proyección de cuatro canales a tres canales.
7. El método de decodificación de audio de cualquiera de las reivindicaciones precedentes, en donde la señal de audio de  $M$  canales comprende o bien tres o bien cuatro canales (L, LS, LB o LSCRN, LW, LS, LB) que representan diferentes direcciones horizontales en un entorno de reproducción para la señal de audio de  $M$  canales, y dos canales (TFL, TBL) que representan direcciones separadas verticalmente de las de dichos tres o cuatro canales en dicho entorno de reproducción.
- 10 8. El método de decodificación de audio de la reivindicación 7, en donde dicho primer grupo consiste en dichos tres canales, y en donde dicho segundo grupo consiste en los dos canales que representan direcciones separadas verticalmente de las de dichos tres canales en dicho entorno de reproducción; o en donde uno de los  $K$  grupos comprende ambos de los dos canales que representan direcciones separadas verticalmente de las de dichos tres o cuatro canales en dicho entorno de reproducción.
- 15 9. El método de decodificación de audio de cualquiera de las reivindicaciones 1-8, en donde la señal descorrelacionada comprende dos canales, un primer canal de la señal descorrelacionada que se obtiene en base al primer canal de la señal de mezcla descendente y un segundo canal de la señal descorrelacionada que se obtiene en base al segundo canal de la señal de mezcla descendente.
- 20 10. El método de decodificación de audio de cualquiera de las reivindicaciones precedentes, que comprende además: recibir señalización (1030) que indica uno de al menos dos formatos de codificación ( $F_1, F_2, F_3$ ) de la señal de audio de  $M$  canales, los formatos de codificación correspondientes a diferentes divisiones respectivas de los canales de la señal de audio de  $M$  canales en los respectivos primer y segundo grupos asociados con los canales de la señal de mezcla descendente,
- 25 en donde los  $K$  grupos están predefinidos, y en donde los coeficientes de mezcla se determinan de manera que una única división de la señal de audio de  $M$  canales en los  $K$  grupos de canales, aproximada por los canales de la señal de salida, se mantenga para dichos al menos dos formatos de codificación.
11. El método de decodificación de audio de la reivindicación 10, en donde:
- 30 en un primer formato de codificación ( $F_1$ ) de dichos al menos dos formatos de codificación, dicho primer grupo consta de tres canales (L, LS, LB) que representan diferentes direcciones horizontales en un entorno de reproducción para la señal de audio de  $M$  canales, y dicho segundo grupo consta de dos canales (TFL, TBL) que representan direcciones separadas verticalmente de las de dichos tres canales en dicho entorno de reproducción; y
- 35 en un segundo formato de codificación ( $F_2$ ) de dichos al menos dos formatos de codificación, cada uno de dichos primer y segundo grupos comprende uno de dichos dos canales que representan direcciones separadas verticalmente de las de dichos tres canales en dicho entorno de reproducción.
12. Un sistema de decodificación de audio (800) que comprende una sección de decodificación (700, 1200) configurada para:
- 40 recibir una señal de mezcla descendente de dos canales ( $L_1, L_2$ ), que se asocia con metadatos, los metadatos que comprenden parámetros de mezcla ascendente ( $\alpha_{LU}$ ) para la reconstrucción paramétrica de una señal de audio de  $M$  canales (L, LS, LB, TFL, TBL) en base a la señal de mezcla descendente, donde  $M \geq 4$ ;
- recibir al menos una parte de dichos metadatos; y
- 45 proporcionar una señal de salida de  $K$  canales ( $\tilde{L}_1, \dots, \tilde{L}_K$ ) en base a la señal de mezcla descendente y los metadatos recibidos, en donde  $2 \leq K < M$ ,
- la sección de decodificación que comprende:
- una sección de descorrelación (710, 1210) configurada para recibir al menos un canal de la señal de mezcla descendente y para emitir, en base a la misma, una señal descorrelacionada (D); y
- una sección de mezcla (720, 1220) configurada para
- 50 determinar un conjunto de coeficientes de mezcla en base a los metadatos recibidos, y

formar la señal de salida como una combinación lineal de la señal de mezcla descendente y la señal descorrelacionada según los coeficientes de mezcla,

caracterizado por que

5 la sección de mezcla está configurada para determinar los coeficientes de mezcla de manera que una suma de un coeficiente de mezcla que controla una contribución del primer canal de la señal de mezcla descendente a un canal de la señal de salida, y un coeficiente de mezcla que controla una contribución del primer canal de la señal de mezcla descendente a otro canal de la señal de salida, tenga el valor 1,

en donde, si la señal de mezcla descendente representa la señal de audio de M canales según un primer formato de codificación ( $F_1$ ) en el que:

10 un primer canal ( $L_1$ ) de la señal de mezcla descendente corresponde a una cierta combinación lineal de un primer grupo (401) de uno o más canales de la señal de audio de M canales;

un segundo canal ( $L_2$ ) de la señal de mezcla descendente corresponde a una cierta combinación lineal de un segundo grupo (402) de uno o más canales de la señal de audio de M canales; y

el primer y segundo grupos constituyen una cierta división de los M canales de la señal de audio de M canales,

15 entonces, la señal de salida de K canales representa la señal de audio de M canales según un segundo formato de codificación ( $F_2, F_4$ ) en el que:

cada uno de los K canales de la señal de salida se aproxima a una combinación lineal de un grupo de uno o más canales de la señal de audio de M canales;

20 los grupos correspondientes a los canales respectivos de la señal de salida constituyen una división de los M canales de la señal de audio de M canales en K grupos (501-502, 1301-1303) de uno o más canales; y

al menos dos de los K grupos comprenden al menos un canal de dicho primer grupo.

13. El sistema de decodificación de audio de la reivindicación 12, que comprende además una sección de decodificación (805) adicional configurada para:

25 recibir una señal de mezcla descendente de dos canales ( $R_1, R_2$ ) adicional, que se asocia con metadatos adicionales, los metadatos adicionales que comprenden parámetros de mezcla ascendentes ( $\alpha_{RU}$ ) adicionales para la reconstrucción paramétrica de una señal de audio de M canales ( $R, RS, RB, TFR, TBR$ ) adicional en base a la señal de mezcla descendente adicional,

recibir al menos una parte de los metadatos adicionales; y

30 proporcionar una señal de salida de K canales ( $\widetilde{R}_1, \dots, \widetilde{R}_K$ ) adicional en base a la señal de mezcla descendente adicional y los metadatos recibidos adicionales,

la sección de decodificación adicional que comprende:

una sección de descorrelación adicional configurada para recibir al menos un canal de la señal de mezcla descendente adicional y para emitir, en base al mismo, una señal descorrelacionada adicional; y

una sección de mezcla adicional configurada para:

35 determinar un conjunto de coeficientes de mezcla adicionales en base a los metadatos adicionales recibidos, y

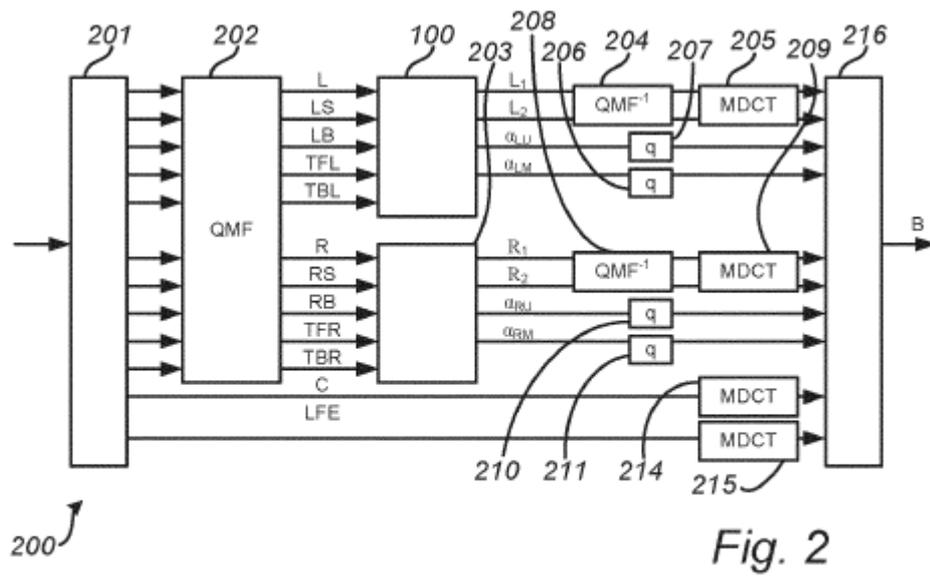
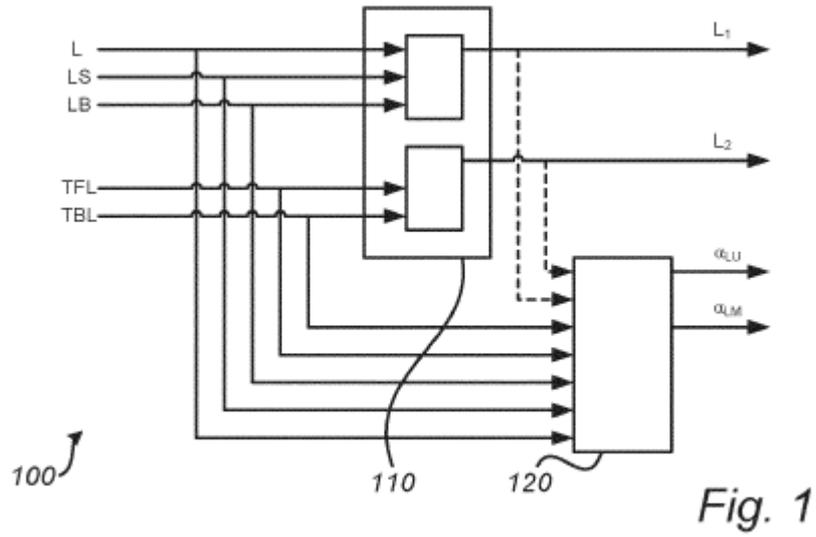
formar la señal de salida adicional como una combinación lineal de la señal de mezcla descendente adicional y la señal descorrelacionada adicional según los coeficientes de mezcla adicionales,

40 en donde la sección de mezcla adicional está configurada para determinar los coeficientes de mezcla adicionales de manera que una suma de un coeficiente de mezcla que controla una contribución del primer canal de la señal de mezcla descendente adicional a un canal de la señal de salida adicional, y un coeficiente de mezcla que controla una contribución del primer canal de la señal de mezcla descendente adicional a otro canal de la señal de salida adicional, tenga el valor 1,

en donde, si la señal de mezcla descendente adicional representa la señal de audio de M canales adicional según un tercer formato de codificación en el que:

45 un primer canal ( $R_1$ ) de la señal de mezcla descendente adicional corresponde a una combinación lineal de un primer grupo (403) de uno o más canales de la señal de audio de M canales adicional;

- un segundo canal ( $R_2$ ) de la señal de mezcla descendente adicional corresponde a una combinación lineal de un segundo grupo (404) de uno o más canales de la señal de audio de M canales adicional; y
- el primer y segundo grupos de canales de la señal de audio de M canales adicional constituyen una división de los M canales de la señal de audio de M canales adicional,
- 5 entonces, la señal de salida de K canales adicional representa la señal de audio de M canales adicional según un cuarto formato de codificación en el que:
- cada uno de los K canales de la señal de salida adicional se aproxima a una combinación lineal de un grupo de uno o más canales de la señal de audio de M canales;
- 10 los grupos correspondientes a los canales respectivos de la señal de salida adicional constituyen una división de los M canales de la señal de audio de M canales adicional en K grupos (503-504, 1304-1306) de uno o más canales; y
- al menos dos de los K grupos de uno o más canales de la señal de audio de M canales adicional comprenden al menos un canal de dicho primer grupo de canales de la señal de audio de M canales adicional.
14. El sistema de decodificación de cualquiera de las reivindicaciones 12-13, que comprende además:
- 15 un demultiplexor (801) configurado para extraer, de un flujo de bits (B), la señal de mezcla descendente, dichos metadatos recibidos, y un canal de audio (C) codificado de manera discreta; y
- una sección de decodificación de un único canal operable para decodificar dicho canal de audio codificado de manera discreta.
- 20 15. Un producto de programa de ordenador que comprende un medio legible por ordenador con instrucciones para realizar el método de cualquiera de las reivindicaciones 1-11, cuando dicho programa se ejecuta en un ordenador.



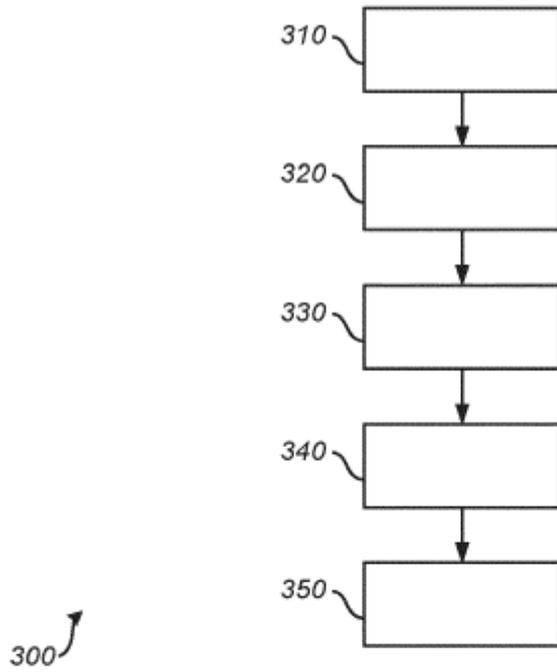


Fig. 3

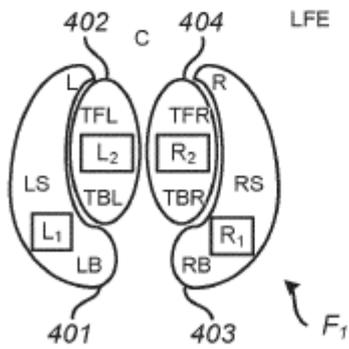


Fig. 4

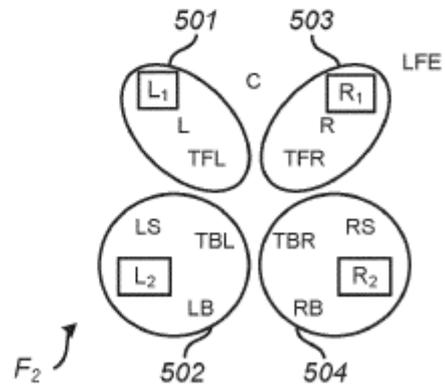


Fig. 5

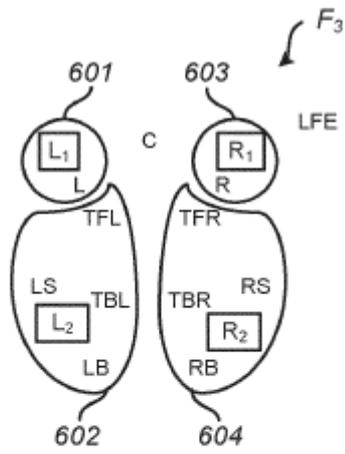


Fig. 6

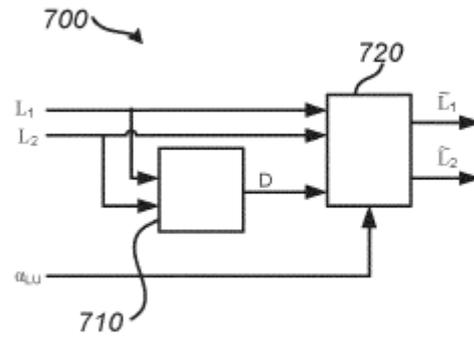


Fig. 7

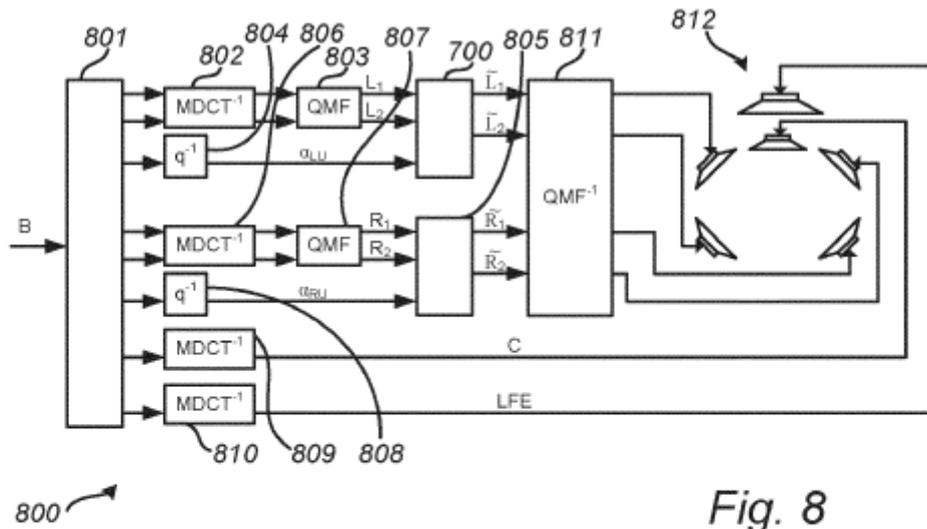


Fig. 8

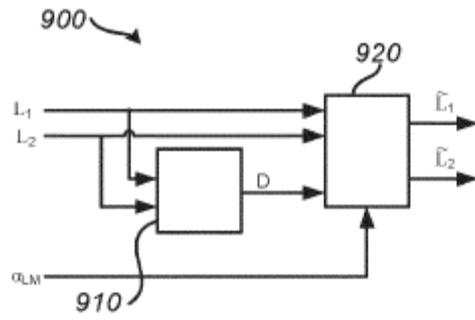


Fig. 9

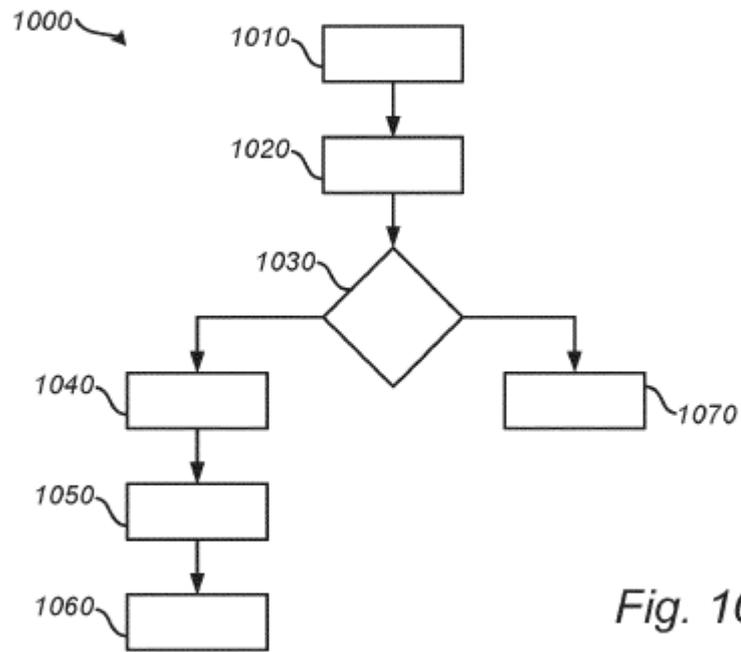


Fig. 10

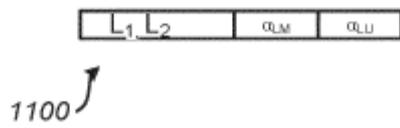
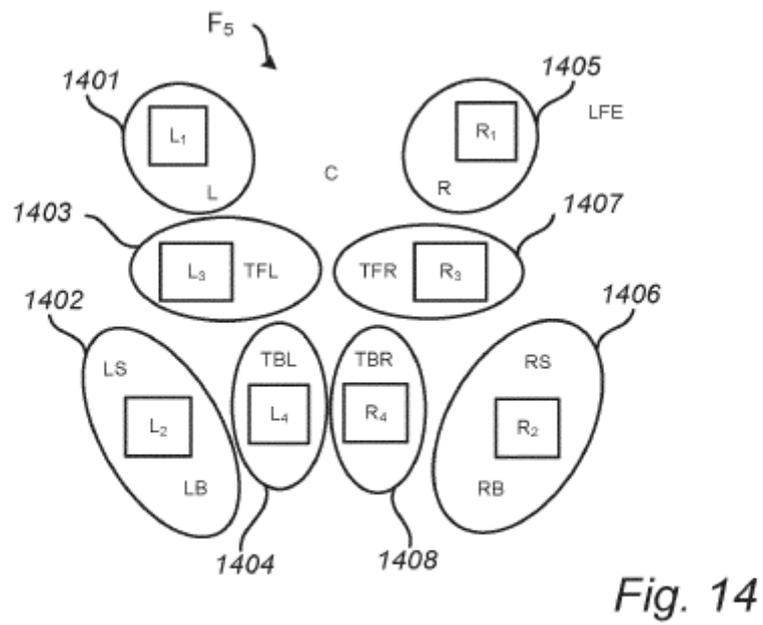
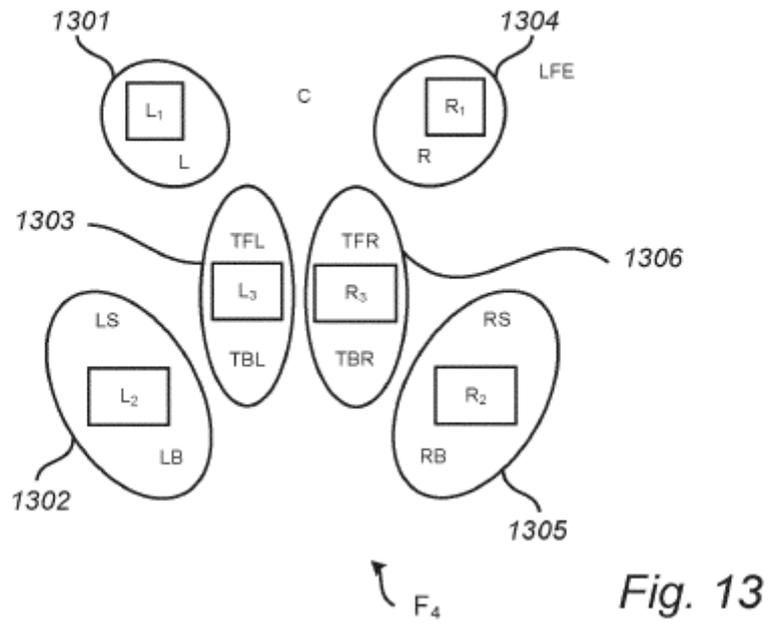


Fig. 11



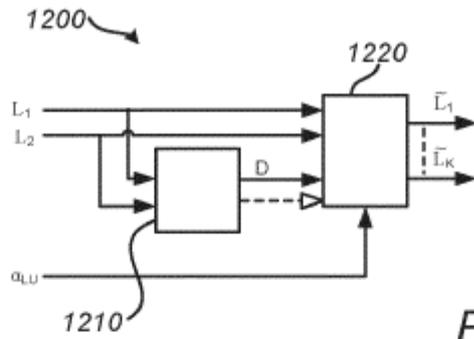


Fig. 12

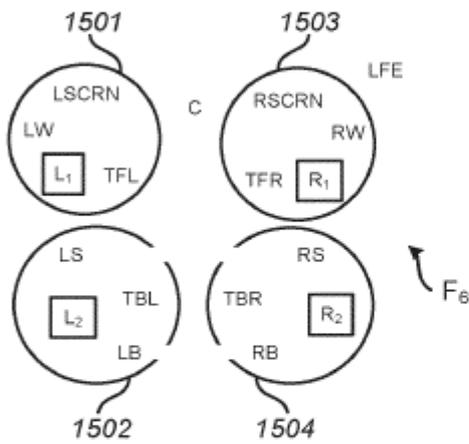


Fig. 15

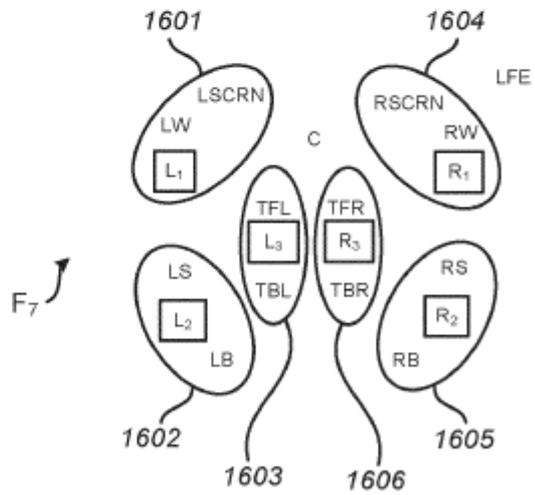


Fig. 16