(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 709 661**

(51) Int. CI.:

G10L 19/008 (2013.01)
G10L 19/22 (2013.01)
H04S 3/00 (2006.01)
H04S 7/00 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **29.10.2015 PCT/EP2015/075115**

(87) Fecha y número de publicación internacional: **06.05.2016 WO16066743**

(96) Fecha de presentación y número de la solicitud europea: **29.10.2015 E 15801335 (9)**

(97) Fecha y número de publicación de la concesión europea: **12.12.2018 EP 3213323**

(54) Título: **Codificación y decodificación paramétrica de señales de audio multicanal**

(30) Prioridad:

31.10.2014 US 201462073642 P
04.03.2015 US 201562128425 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
17.04.2019

(73) Titular/es:

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

(72) Inventor/es:

PURNHAGEN, HEIKO;
LEHTONEN, HEIDI-MARIA y
KLEJSA, JANUSZ

(74) Agente/Representante:

ELZABURU, S.L.P

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

Codificación y decodificación paramétrica de señales de audio multicanal

Referencia cruzada a solicitudes relacionadas

Esta solicitud reivindica prioridad a la Solicitud de Patente Provisional de EE.UU. Nº62/073,642, presentada el 31 de octubre de 2014 y la Solicitud de Patente Provisional de EE.UU. Nº 62/128,425 presentada el 4 de marzo de 2015.

Campo técnico

La invención descrita en la presente memoria se refiere en general a la codificación y decodificación paramétrica de señales de audio, y en particular a la codificación y decodificación paramétrica de señales de audio basadas en canales.

Antecedentes

Los sistemas de reproducción de audio que comprenden múltiples altavoces se utilizan frecuentemente para reproducir una escena de audio representada por una señal de audio multicanal, en donde los respectivos canales de la señal de audio multicanal se reproducen en los respectivos altavoces. La señal de audio multicanal puede por ejemplo haber sido grabada a través de una pluralidad de transductores acústicos o puede haber sido generada por equipo 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 tamaño de almacenamiento. En un lado de 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 (uno canal) o una estéreo (dos canales), y extraen información lateral que describe las propiedades de los canales por medio de parámetros como diferencias de nivel y correlación cruzada. La mezcla descendente y la información lateral se codifican luego y se envían a un lado de decodificador. En el lado de decodificador, se reconstruye la señal de audio multicanal, i.e., aproximada, a partir de la mezcla descendente bajo control de los parámetros de la información lateral.

En vista el amplio intervalo de diferentes tipos de dispositivos y sistemas disponibles para la reproducción de contenido de audio multicanal, incluyendo un segmento emergente dirigido a usuarios finales en sus hogares, hay una necesidad de maneras nuevas y alternativas de codificar eficientemente el contenido de audio multicanal, para reducir los requerimientos de ancho de banda y/o el tamaño de memoria requerido para el almacenamiento, facilitar la reconstrucción de la señal de audio multicanal en un lado de decodificador, y/o aumentar la fidelidad de la señal de audio multicanal reconstruida en un lado de decodificador.

El estándar internacional ISO/IEC FDIS 23003-1: 2006: E que describe MPEG Surround se refiere entre otras cosas al procesamiento de audio espacial de diferentes configuraciones de canales. Por ejemplo, se refiere a la mezcla descendente de diferentes programas de 7 canales en dos canales.

La solicitud de patente de EE.UU con número de publicación 2008/0244856 A1 se refiere a un codificador de audio que tiene una estructura de codificación jerárquica y que genera un flujo de datos que comprende uno o más canales de audio, datos de codificación de audio paramétrica, y una estructura jerárquica de árbol de decodificador.

Breve descripción de los dibujos

En lo que sigue, se describirán realizaciones ejemplares en mayor detalle y con referencia a los dibujos adjuntos, en los que:

Las Figuras 1 y 2 son diagramas de bloques generalizados de secciones de codificación para codificar señales de audio de M canales como señales de mezcla descendente de dos canales y los parámetros de mezcla ascendente asociados, según las realizaciones ejemplares;

La Figura 3 es un diagrama de bloques generalizado de un sistema de codificación de audio que comprende la sección de codificación representada en la Figura 1, según una realización ejemplar;

Las Figuras 4 y 5 son diagramas de flujo de métodos de codificación de audio para codificar señales de audio de M canales como señales de mezcla descendente de dos canales y los parámetros de mezcla ascendente asociados, según las realizaciones ejemplares;

Las Figuras 6-8 ilustran maneras alternativas de dividir una señal de audio de 11.1 canales (o 7.1+4 canales o 7.1.4 canales) en grupos de canales representados por los respectivos canales de mezcla descendente, según las realizaciones ejemplares;

La Figura 9 es un diagrama de bloques generalizado de una sección de decodificación para reconstruir una señal de audio de M canales en base a una señal de mezcla descendente de dos canales y los parámetros de mezcla ascendente asociados, según una realización ejemplar;

La Figura 10 es un diagrama de bloques generalizado de un sistema de decodificación de audio que comprende la sección de decodificación representada en la Figura 9, según una realización ejemplar;

La Figura 11 es un diagrama de bloques generalizado de una sección de mezcla comprendida en la sección de decodificación representada en la Figura 9, según una realización ejemplar;

La Figura 12 es un diagrama de flujo de un método de decodificación de audio para reconstruir una señal de audio de M canales en base a una señal de mezcla descendente de dos canales y los parámetros de mezcla ascendente asociados, según una realización ejemplar; y

La Figura 13 es un diagrama de bloques generalizado de una sección de decodificación para reconstruir una señal de audio de 13.1 canales en base a una señal de 5.1 canales y los parámetros de mezcla ascendente asociados, según una realización ejemplar;

La Figura 14 es un diagrama de bloques generalizado de una sección de decodificación configurada para determinar un formato de codificación adecuado para ser utilizado para codificar una señal de audio de M canales (y posibles canales adicionales) y, para el formato elegido, representar la señal de audio de M canales como una señal de mezcla descendente de dos canales y los parámetros de mezcla ascendente asociados;

La Figura 15 es un detalle de una sección de mezcla descendente de modo dual en la sección de codificación mostrada en la Figura 14;

La Figura 16 es un detalle de una sección de análisis de modo dual en la sección de codificación mostrada en la Figura 14; y

La Figura 17 es un diagrama de flujo de un método de codificación de audio que puede ser realizado por los componentes mostrados en las Figuras 14 a 16.

Todas las figuras son esquemáticas y en general solo muestran las partes que son necesarias para dilucidar la invención, mientras que otras partes pueden ser omitidas o simplemente sugeridas.

Descripción de las realizaciones ejemplares

Como se emplea en la presente memoria, una señal de audio puede ser una señal de audio independiente, una parte de audio de una señal audiovisual o señal multimedia o cualquiera de estas en combinación con metadatos. Como se emplea 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”.

I. Visión general – Lado de decodificador

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

Según las realizaciones ejemplares, se proporciona un método de decodificación de audio que comprende recibir una señal de mezcla descendente de dos canales y 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 $M \geq 4$. El método de decodificación de audio comprende recibir señalización que indica uno seleccionado de al menos dos formatos de codificación de la señal de audio de M canales, donde los formatos de codificación corresponden a respectivas particiones diferentes de los canales de la señal de audio de M canales en respectivos primer y segundo grupos de uno o más canales. En el formato de codificación indicado, un primer canal de la señal de mezcla descendente corresponde a una combinación lineal del 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 del segundo grupo de uno o más canales de la señal de audio de M canales. El método de decodificación de audio comprende además: determinar un conjunto de coeficientes de descorrelación previa en base al formato de codificación indicado; calcular una señal de entrada de descorrelación como un mapeo lineal de la señal de mezcla descendente, en donde el conjunto de coeficientes de descorrelación previa se aplica a la señal de mezcla descendente; generar una señal descorrelacionada en base a la señal de entrada de descorrelación; determinar conjuntos de coeficientes de mezcla ascendente de un primer tipo, denominados en la presente memoria como coeficientes de mezcla ascendente húmedos, y de un segundo tipo, denominados en la presente memoria como coeficientes de mezcla ascendente secos, en base a los parámetros de mezcla ascendente recibidos y el formato de codificación indicado; calcular una señal de mezcla ascendente de un primer tipo, denominada en la presente memoria como una señal de mezcla ascendente seca, como un mapeo lineal de la señal de mezcla descendente, en donde el conjunto de coeficientes de mezcla

ascendente secos se aplica a la señal de mezcla descendente; calcular una señal de mezcla ascendente de un segundo tipo, denominada en la presente memoria como una señal de mezcla ascendente húmeda, como un mapeo lineal de la señal descorrelacionada, en donde el conjunto de coeficientes de mezcla ascendente húmedos se aplica a la señal descorrelacionada; y combinar las señales de mezcla ascendente secas y húmedas para obtener una señal reconstruida multidimensional correspondiente a la señal de audio de M canales a reconstruir.

Dependiendo del contenido de audio de la señal de audio de M canales, diferentes particiones de los canales de la señal de audio de M canales en primer y segundo grupos, en donde cada grupo contribuye a un canal de la señal de mezcla descendente, pueden ser adecuadas para p.ej., facilitar la reconstrucción de la señal de audio de M canales a partir de la señal de mezcla descendente, mejorar la fidelidad (percibida) de la señal de audio de M canales reconstruida a partir de la señal de mezcla descendente, y/o mejorar la eficiencia de codificación de la señal de mezcla descendente. La capacidad del método de decodificación de audio para recibir señalización que indica uno seleccionado de los formatos de codificación, y adaptar la determinación de los coeficientes de descorrelación previa así como los coeficientes de mezcla ascendente húmedos y secos al formato de codificación indicado, permite que se seleccione un formato de codificación en un lado de codificador, p.ej., en base al contenido de audio de la señal de audio de M canales, para aprovechar las ventajas comparativas de emplear ese formato de codificación particular para representar la señal de audio de M canales.

En particular, determinar los coeficientes de descorrelación previa en base al formato de codificación indicado puede permitir que el canal, o canales, de la señal de mezcla descendente, a partir de la que se genera la señal descorrelacionada, sean seleccionados y/o ponderados, en base al formato de codificación indicado, antes de generar la señal descorrelacionada. La capacidad del método de decodificación de audio para determinar los coeficientes de descorrelación previa de manera diferente para diferentes formatos de codificación puede por lo tanto permitir mejorar la fidelidad de la señal de audio de M canales reconstruida.

El primer canal de la señal de mezcla descendente se puede por ejemplo haber formado p.ej., en un lado de codificador, como una combinación lineal del primer grupo de uno o más canales, de acuerdo con el formato de codificación indicado. Similarmente, el segundo canal de la señal de mezcla descendente se puede por ejemplo haber formado, en un lado de codificador, como una combinación lineal del segundo grupo de uno o más canales, de acuerdo con el formato de codificación indicado.

Los canales de la señal de audio de M canales pueden por ejemplo formar un subconjunto de un número mayor de canales que representan juntos un campo de sonido.

La señal descorrelacionada sirve para aumentar la dimensionalidad del contenido de audio de la señal de mezcla descendente, tal como es percibido por oyente. Generar la señal descorrelacionada puede por ejemplo incluir aplicar un filtro lineal a la señal de entrada de descorrelación.

El que la señal de entrada de descorrelación se calcule como un mapeo lineal de la señal de mezcla descendente significa que la señal de entrada de descorrelación se obtiene aplicando una primera transformación lineal a la señal de mezcla descendente. Esta primera transformación lineal toma los dos canales de la señal de mezcla descendente como entrada y proporciona los canales de la señal de entrada de descorrelación como salida, y los coeficientes de descorrelación previa son coeficientes que definen las propiedades cuantitativas de esta primera transformación lineal.

El que la señal de mezcla ascendente seca se calcule como un mapeo lineal de la señal de mezcla descendente significa que la señal de mezcla descendente seca se obtiene aplicando una segunda transformación lineal a la señal de mezcla descendente. Esta segunda transformación lineal toma los dos canales de la señal de mezcla descendente como entrada y proporciona M canales como salida, y los coeficientes de mezcla ascendente secos son coeficientes que definen las propiedades cuantitativas de esta segunda transformación lineal.

El que la señal de mezcla ascendente húmeda se calcule como un mapeo lineal de la señal descorrelacionada significa que la señal de mezcla ascendente húmeda se obtiene aplicando una tercera transformación lineal a la señal descorrelacionada. Esta tercera transformación lineal toma los canales de la señal descorrelacionada como entrada y proporciona M canales como salida, y los coeficientes de mezcla ascendente húmedos son coeficientes que definen las propiedades cuantitativas de esta tercera transformación lineal.

Combinar las señales de mezcla ascendente secas y húmedas puede incluir añadir contenido de audio de los respectivos canales de la señal de mezcla ascendente seca al contenido de audio de los respectivos canales correspondientes de la señal de mezcla ascendente húmeda, p.ej., emplear mezcla aditiva sobre una base por muestra o por coeficiente de transformación.

La señalización se puede por ejemplo recibir junto con la señal de mezcla descendente y/o los parámetros de mezcla ascendente. La señal de mezcla descendente, los parámetros de mezcla ascendente y la señalización se pueden por ejemplo extraer de un flujo de bits.

En una realización ejemplar, se puede sostener que M = 5, i.e., la señal de audio de M canales puede ser una señal de audio de cinco canales. El método de decodificación de audio de la presente realización ejemplar se puede por

ejemplo emplear para reconstruir los cinco canales regulares en uno de los formatos de audio 5.1 actualmente establecidos a partir de una mezcla descendente de dos canales de esos cinco canales, o para reconstruir cinco canales en el lado izquierdo, o en el lado derecho, en una señal de audio multicanal 11.1, a partir de una mezcla descendente de dos canales de esos cinco canales.

Alternativamente, se puede sostener que M = 4, o M ≥ 6.

En una realización ejemplar, la señal de entrada de descorrelación y la señal descorrelacionada pueden cada una comprender M – 2 canales. En la presente realización ejemplar, un canal de la señal descorrelacionada se puede generar en base a no más de un canal de la señal de entrada de descorrelación. Por ejemplo, cada canal de la señal descorrelacionada se puede generar en base a no más de un canal de la señal de entrada de descorrelación, pero diferentes canales de la señal descorrelacionada se pueden por ejemplo generar en base a diferentes canales de la señal de entrada de descorrelación.

En la presente realización ejemplar, los coeficientes de descorrelación previa se pueden determinar de tal manera que, en cada uno de los formatos de codificación, un canal de la señal de entrada de descorrelación reciba contribución de no más de un canal de la señal de mezcla descendente. Por ejemplo, los coeficientes de descorrelación previa se pueden determinar de tal manera que, en cada uno de los formatos de codificación, cada canal de la señal de entrada de descorrelación coincide con un canal de la señal de mezcla descendente. Sin embargo, se apreciará que al menos algunos de los canales de la señal de entrada descorrelacionada pueden por ejemplo coincidir con diferentes canales de la señal de mezcla descendente en un formato de codificación dado y/o en los diferentes formatos de codificación.

Dado que, en cada formato de codificación dado, los dos canales de la señal de mezcla descendente representan el primer y segundo grupos disjuntos de uno o más canales, el primer grupo se puede reconstruir a partir del primer canal de la señal de mezcla descendente, p.ej., empleando uno o más canales de la señal descorrelacionada generada en base al primer canal de la señal de mezcla descendente, mientras que el segundo grupo se puede reconstruir a partir del segundo canal de la señal de mezcla descendente, p.ej., empleando uno o más canales de la señal descorrelacionada generados en base al segundo canal de la señal de mezcla descendente. En la presente realización ejemplar, la contribución del segundo grupo de uno o más canales, a una versión reconstruida del primer grupo de uno o más canales, a través de la señal descorrelacionada, se puede evitar en cada formato de codificación. Similarmente, la contribución del primer grupo de uno o más canales, a una versión reconstruida del segundo grupo de uno o más canales, a través de la señal descorrelacionada, se puede evitar en cada formato de codificación. La presente realización ejemplar puede por lo tanto permitir aumentar la fidelidad de la señal de audio de M canales reconstruida.

En una realización ejemplar, los coeficientes de descorrelación previa se pueden determinar de tal manera que un primer canal de la señal de audio de M canales contribuye, a través de la señal de mezcla descendente, a un primer canal fijo de la señal de entrada de descorrelación en al menos dos de los formatos de codificación. Es decir, el primer canal de la señal de audio de M canales puede contribuir, a través de la señal de mezcla descendente, al mismo canal de la señal de entrada de descorrelación en ambos de estos formatos de codificación. Se apreciará que en la presente realización ejemplar, el primer canal de la señal de audio de M canales puede por ejemplo contribuir, a través de la señal de mezcla descendente, a múltiples canales de la señal de entrada de descorrelación en un formato de codificación dado.

En la presente realización ejemplar, si el formato de codificación indicado conmuta entre los dos formatos de codificación, entonces al menos una porción del primer canal fijo de la señal de entrada de descorrelación permanece durante la conmutación. Esto puede permitir una transición más suave y/o menos abrupta entre los formatos de codificación, tal como es percibido por un oyente durante la reproducción de la señal de audio de M canales reconstruida. En particular, los inventores se han dado cuenta de que dado que la señal descorrelacionada se puede por ejemplo generar en base a una sección de la señal de mezcla descendente correspondiente a varias tramas de tiempo, durante las que se puede producir una conmutación entre los formatos de codificación en la señal de mezcla descendente, se pueden potencialmente generar artefactos audibles en la señal descorrelacionada como resultado de la conmutación entre los formatos de codificación. Incluso si se interpolan los coeficientes de mezcla ascendente húmedos y secos en respuesta a una conmutación entre los formatos de codificación, los artefactos generados en la señal descorrelacionada todavía pueden persistir en la señal de audio de M canales reconstruida. Proporcionar una señal de entrada de descorrelación de acuerdo con la presente realización ejemplar permite suprimir tales artefactos en la señal descorrelacionada que son provocados por la conmutación entre los formatos de codificación, y puede mejorar la calidad de reproducción de la señal de audio de M canales reconstruida.

En una realización ejemplar, los coeficientes de descorrelación previa se pueden determinar de tal manera que, adicionalmente, un segundo canal de la señal de audio de M canales contribuye, a través de la señal de mezcla descendente, a un segundo canal fijo de la señal de entrada de descorrelación en al menos dos de los formatos de codificación. Es decir, el segundo canal de la señal de audio de M canales contribuye, a través de la señal de mezcla descendente, al mismo canal de la señal de entrada de descorrelación en ambos formatos de codificación. En la presente realización ejemplar, si el formato de codificación indicado conmuta entre los dos formatos de codificación, entonces al menos una porción de la segunda señal de entrada de descorrelación fija permanece durante la conmutación. Como tal, una sola alimentación de descorrelador se ve afectada por una transición entre los formatos

de codificación. Esto puede permitir una transición más suave y/o menos abrupta entre los formatos de codificación, tal como es percibido por un oyente durante la reproducción de la señal de audio de M canales reconstruida.

El primer y segundo canales de la señal de audio de M canales pueden por ejemplo ser distintos entre sí. El primer y segundo canales fijos de la señal de entrada de descorrelación pueden por ejemplo ser distintos entre sí.

En una realización ejemplar, la señalización recibida puede indicar uno seleccionado de al menos tres formatos de codificación, y los coeficientes de descorrelación previa se pueden determinar de tal manera que el primer canal de la señal de audio de M canales contribuye, a través de la señal de mezcla descendente al primer canal fijo de la señal de entrada de descorrelación en al menos tres de los formatos de codificación. Es decir, el primer canal de la señal de audio de M canales contribuye, a través de la señal de mezcla descendente, al mismo canal de la señal de entrada de descorrelación en estos tres formatos de codificación. En la presente realización ejemplar, si el formato de codificación indicado cambia entre cualquiera de los tres formatos de codificación, entonces al menos una porción del primer canal fijo de la señal de entrada de descorrelación permanece durante la conmutación, lo que permite una transición más suave y/o menos abrupta entre los formatos de codificación, tal como es percibido por un oyente durante la reproducción de la señal de audio de M canales reconstruida.

En una realización ejemplar, los coeficientes de descorrelación previa se pueden determinar de tal manera que un par de canales de la señal de audio de M canales contribuye, a través de la señal de mezcla descendente, a un tercer canal fijo de la señal de entrada de descorrelación en al menos dos de los formatos de codificación. Es decir, el par de canales de la señal de audio de M canales contribuye, a través de la señal de mezcla descendente, al mismo canal de la señal de entrada de descorrelación en ambos formatos de codificación. En la presente realización ejemplar, si el formato de codificación indicado conmuta entre los dos formatos de codificación, entonces al menos una porción del tercer canal fijo de la señal de entrada de descorrelación permanece durante la conmutación, lo que permite una transición más suave y menos abrupta entre los formatos de codificación, tal como es percibido por un oyente durante la reproducción de la señal de audio de M canales reconstruida.

El par de canales puede por ejemplo ser distinto del primer y segundo canales de la señal de audio de M canales. El tercer canal fijo de la señal de entrada de descorrelación puede por ejemplo ser distinto del primer y segundo canales fijos de la señal de entrada de descorrelación.

En una realización ejemplar, el método de decodificación de audio puede además comprender: en respuesta a la detección de una conmutación del formato de codificación indicado de un primer formato de codificación a un segundo formato de codificación, realizar una transición gradual de valores de los coeficientes de descorrelación previa asociados con el primer formato de codificación a valores de los coeficientes de descorrelación previa asociados con el segundo formato de codificación. Emplear una transición gradual entre los coeficientes de descorrelación previa durante la conmutación entre formatos de codificación permite una transición más suave y/o menos abrupta entre los formatos de codificación, tal como es percibido por un oyente durante la reproducción de la señal de audio de M canales reconstruida. En particular, los inventores se han dado cuenta de que dado que la señal descorrelacionada se puede por ejemplo generar en base a una sección de la señal de mezcla descendente correspondiente a varias tramas de tiempo, durante las que se puede producir una conmutación entre los formatos de codificación en la señal de mezcla descendente, se pueden potencialmente generar artefactos audibles en la señal descorrelacionada como resultado de la conmutación entre formatos de codificación. Incluso si se interpolan los coeficientes de mezcla ascendente húmedos y secos en respuesta a una conmutación entre los formatos de codificación, los artefactos generados en la señal descorrelacionada todavía pueden persistir en la señal de audio de M canales reconstruida. Proporcionar una señal de entrada de descorrelación de acuerdo con la presente realización ejemplar permite suprimir tales artefactos en la señal descorrelacionada que son provocados por la conmutación entre los formatos de codificación, y puede mejorar la calidad de reproducción de la señal de audio de M canales reconstruida.

La transición gradual se puede por ejemplo realizar a través de interpolación lineal o continua. La transición gradual se puede por ejemplo realizar a través de interpolación con una tasa de cambio limitada.

En una realización ejemplar, el método de decodificación de audio puede además comprender: en respuesta a la detección de una conmutación del formato de codificación indicado de un primer formato de codificación a un segundo formato de codificación, realizar interpolación de valores de los coeficientes de mezcla ascendente húmedos y secos, incluyendo los coeficientes de valor cero, asociados con el primer formato de codificación a valores de los coeficientes de mezcla ascendente húmedos y secos, de nuevo incluyendo los coeficientes de valor cero, asociados con el segundo formato de codificación. Se recuerda que los canales de mezcla descendente corresponden a diferentes combinaciones de canales de la señal de audio de M canales originalmente codificada, de modo que un coeficiente de mezcla ascendente que es de valor cero en el primer formato de codificación no necesita ser de valor cero en el segundo formato de codificación también, y viceversa. Preferiblemente, la interpolación actúa sobre los coeficientes de mezcla ascendente en lugar de una representación compacta de los coeficientes, p.ej., la representación discutida a continuación.

La interpolación lineal o continua entre valores de los coeficientes de mezcla ascendente se puede por ejemplo emplear para proporcionar una transición más suave entre los formatos de codificación, tal como es percibido por un oyente durante la reproducción de la señal de audio de M canales reconstruida.

La interpolación pronunciada, en la que nuevos valores de los coeficientes de mezcla ascendente reemplazan a los antiguos valores de los coeficientes de mezcla ascendente en un cierto punto en el tiempo asociado con la conmutación entre los formatos de codificación, puede por ejemplo permitir una mayor fidelidad de la señal de audio de M canales reconstruida, p.ej., en casos donde el contenido de audio de la señal de audio de M canales cambia rápidamente y donde el formato de codificación se conmuta en un lado de codificador, en respuesta a estos cambios, para aumentar la fidelidad de la señal de audio de M canales reconstruida.

En una realización ejemplar, el método de decodificación de audio puede además comprender recibir señalización que indica uno de una pluralidad de esquemas de interpolación a ser empleados para la interpolación de parámetros de mezcla ascendente húmedos y secos dentro de un formato de codificación (i.e., cuando se asignan nuevos valores a los coeficientes de mezcla ascendente en un periodo de tiempo donde no se produce cambio de formato de codificación), y emplear el esquema de interpolación indicado. La señalización que indica uno de una pluralidad de esquemas de interpolación se puede por ejemplo recibir junto con la señal de mezcla descendente y/o los parámetros de mezcla ascendente. Preferiblemente, el esquema de interpolación indicado por la señalización se puede además emplear para la transición entre formatos de codificación.

En un lado de codificador, donde está disponible la señal de audio de M canales original, se pueden por ejemplo seleccionar esquemas de interpolación que son particularmente adecuados para el contenido de audio real de la señal de audio de M canales. Por ejemplo, se puede emplear interpolación lineal o continua donde la conmutación suave es importante para la impresión global de la señal de audio de M canales reconstruida, mientras que se puede emplear interpolación pronunciada, i.e., en la que nuevos valores de los coeficientes de mezcla ascendente reemplazan a los antiguos valores de los coeficientes de mezcla ascendente en un cierto punto en el tiempo asociado con la transición entre los formatos de codificación, cuando la conmutación rápida es importante para la impresión global de la señal de audio de M canales reconstruida.

En una realización ejemplar, los al menos dos formatos de codificación pueden incluir un primer formato de codificación y un segundo formato de codificación. Hay una ganancia que controla una contribución, en cada formato de codificación, de 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. En la presente realización ejemplar, una ganancia en el primer formato de codificación puede coincidir con una ganancia en el segundo formato de codificación que controla una contribución del mismo canal de la señal de audio de M canales.

Emplear las mismas ganancias en el primer y segundo formatos de codificación puede por ejemplo aumentar la similitud entre el contenido de audio combinado de los canales de la señal de mezcla descendente en el primer formato de codificación y el contenido de audio combinado de los canales de la señal de mezcla descendente en el segundo formato de codificación. Puesto que los canales de la señal de mezcla descendente se utilizan para reconstruir la señal de mezcla descendente de M canales, esto puede contribuir a transiciones más suaves entre estos dos formatos de codificación, tal como es percibido por un oyente.

Emplear las mismas ganancias en el primer y segundo formatos de codificación puede por ejemplo permitir que el contenido de audio del primer y segundo canales, respectivamente, de la señal de mezcla descendente en el primer formato de codificación sea más similar al contenido de audio del primer y segundo canales, respectivamente, de la señal de mezcla descendente en el segundo formato de codificación. Esto puede contribuir a transiciones más suaves entre estos dos formatos de codificación, tal como es percibido por un oyente.

En la presente realización ejemplar, se pueden por ejemplo emplear diferentes ganancias para diferentes canales de la señal de audio de M canales. En un primer ejemplo, todas las ganancias en el primer y segundo formatos de codificación tienen 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, tanto en el primer como en el segundo formato de codificación. En un segundo ejemplo, al menos algunas de las ganancias pueden tener valores diferentes a 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.

En una realización ejemplar, 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 verticalmente separadas 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 la reproducción por fuentes de audio situadas a sustancialmente la misma altura que un oyente (o el oído de un oyente) y/o que se propagan sustancialmente horizontalmente, y dos canales destinados a la reproducción por fuentes de audio situadas a otras alturas y/o que se propagan (sustancialmente) de forma no horizontal. Los dos canales pueden por ejemplo representar direcciones elevadas.

En una realización ejemplar, en un primer formato de codificación, el segundo grupo de canales puede comprender los dos canales que representan direcciones verticalmente separadas de las de los tres canales en el entorno de reproducción. Tener estos dos canales en el segundo grupo, y emplear el mismo canal de la señal de mezcla descendente para representar estos dos canales, puede por ejemplo mejorar la fidelidad de la señal de audio de M

canales reconstruida en casos donde una dimensión vertical en el entorno de reproducción es importante para la impresión global de la señal de audio de M canales.

En una realización ejemplar, en un primer formato de codificación, el primer grupo de uno o más canales puede comprender los tres canales que representan diferentes direcciones horizontales en un entorno de reproducción de la señal de audio de M canales, y el segundo grupo de uno o más canales puede comprender los dos canales que representan direcciones verticalmente separadas de las de los tres canales en el entono de reproducción. En la presente realización ejemplar, el primer formato de codificación permite que el primer canal de la señal de mezcla descendente represente los tres canales y el segundo canal de la señal de mezcla descendente represente los dos canales, lo que puede por ejemplo mejorar la fidelidad de la señal de audio de M canales reconstruida en casos donde una dimensión vertical en el entorno de reproducción es importante para la impresión global de la señal de audio de M canales.

En una realización ejemplar, en un segundo formato de codificación, cada uno del primer y segundo grupos puede comprender uno de los dos canales que representan direcciones verticalmente separadas de las de los tres canales en un entorno de reproducción de la señal de audio de M canales. Tener estos tres canales en diferentes grupos, y emplear los diferentes canales de la señal de mezcla descendente para representar estos dos canales, puede por ejemplo mejorar la fidelidad de la señal de audio de M canales reconstruida en casos donde una dimensión vertical en el entorno de reproducción no es tan importante para la impresión global de la señal de audio de M canales.

En una realización ejemplar, en un formato de codificación, denominado en la presente memoria como un formato de codificación particular, el primer grupo de uno o más canales puede constar de N canales, donde $N \geq 3$. En la presente realización ejemplar, en respuesta a que el formato de codificación indicado sea el formato de codificación particular: los coeficientes de descorrelación previa se pueden determinar de tal manera que se generan $N - 1$ canales de la señal descorrelacionada en base al primer canal de la señal de mezcla descendente; y los coeficientes de mezcla ascendente secos y húmedos se pueden determinar de tal manera que el primer grupo de uno o más canales se reconstruye como un mapeo lineal del primer canal de la señal de mezcla descendente y los $N - 1$ canales de la señal descorrelacionada, en donde un subconjunto de los coeficientes de mezcla ascendente secos se aplica al primer canal de la señal de mezcla descendente y un subconjunto de los coeficientes de mezcla ascendente húmedos se aplica a los $N - 1$ canales de la señal descorrelacionada.

Los coeficientes de descorrelación previa se pueden por ejemplo determinar de tal manera que $N - 1$ canales de la señal de entrada de descorrelación coinciden con el primer canal de la señal de mezcla descendente. Los $N - 1$ canales de la señal descorrelacionada se pueden por ejemplo generar procesando estos $N - 1$ canales de la señal de entrada de descorrelación.

El que el primer grupo de uno o más canales se reconstruya como un mapeo lineal del primer canal de la señal de mezcla descendente y los $N - 1$ canales de la señal descorrelacionada significa que una versión reconstruida del primer grupo de uno o más canales se obtiene aplicando una transformación lineal al primer canal de la señal de mezcla descendente y los $N - 1$ canales de la señal descorrelacionada. Esta transformación lineal toma N canales como entrada y proporciona N canales como salida, donde el subconjunto de los coeficientes de mezcla ascendente secos y el subconjunto de los coeficientes de mezcla ascendente húmedos juntos constan de coeficientes que definen las propiedades cuantitativas de esta transformación lineal.

En una realización ejemplar, los parámetros de mezcla ascendente recibidos pueden incluir parámetros de mezcla ascendente de un primer tipo, denominados en la presente memoria como parámetros de mezcla ascendente húmedos, y parámetros de mezcla ascendente de un segundo tipo, denominados en la presente memoria como parámetros de mezcla ascendente secos. En la presente realización ejemplar, determinar los conjuntos de coeficientes de mezcla ascendente húmedos y secos, en el formato de codificación particular, puede comprender: determinar, en base a los parámetros de mezcla ascendente secos, el subconjunto de los coeficientes de mezcla ascendente secos; rellenar una matriz intermedia que tiene más elementos que el número de parámetros de mezcla ascendente húmedos recibidos, en base a los parámetros de mezcla ascendente húmedos recibidos y sabiendo que la matriz intermedia pertenece a una clase de matrices predefinidas; y obtener el subconjunto de los coeficientes de mezcla ascendente húmedos multiplicando la matriz intermedia por una matriz predefinida, en donde el subconjunto de los coeficientes de mezcla ascendente húmedos corresponde a la matriz que resulta de la multiplicación e incluye más coeficientes que el número de elementos en la matriz intermedia.

En la presente realización ejemplar, el número de coeficientes de mezcla ascendente húmedos en el subconjunto de coeficientes de mezcla ascendente húmedos es mayor que el número de parámetros de mezcla ascendente húmedos recibidos. Aprovechando el conocimiento de la matriz predefinida y la clase de matrices predefinidas para obtener el subconjunto de coeficientes de mezcla ascendente húmedos a partir de los parámetros de mezcla ascendente húmedos recibidos, se puede reducir la cantidad de información necesaria para la reconstrucción paramétrica del primer grupo de uno o más canales, permitiendo una reducción de la cantidad de metadatos transmitidos junto con la señal de mezcla descendente desde un lado de 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 clase de matrices predefinidas se puede asociar con propiedades conocidas de al menos algunos elementos de la matriz que son válidas para todas las matrices de la clase, tal como ciertas relaciones entre algunos de los elementos de la matriz, o que algunos elementos de la matriz son cero. El conocimiento de estas propiedades permite rellenar la matriz intermedia en base a menos parámetros de mezcla ascendente húmedos que el número total de elementos de la matriz en la matriz intermedia. El lado de 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 en base al menor número parámetros de mezcla ascendente húmedos.

Cómo determinar y emplear la matriz predefinida y la clase de matrices predefinidas se describe en más detalle en la página 16, línea 15 a página 20, línea 2 en la solicitud de patente provisional de EE.UU Nº 61/974,544; 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 una realización ejemplar, los parámetros de mezcla ascendente recibidos pueden incluir N(N – 1)/2 parámetros de mezcla ascendente húmedos. En la presente realización ejemplar, rellenar 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 húmedos recibidos y saber que la matriz intermedia pertenece a la clase de matrices predefinidas. Esto puede incluir insertar los valores de los parámetros de mezcla ascendente húmedos inmediatamente como elementos de la matriz, o procesar los parámetros de mezcla ascendente húmedos de manera adecuada para obtener los valores para los elementos de la matriz. En la presente realización ejemplar, la matriz predefinida puede incluir N(N – 1) elementos, y el subconjunto de los coeficientes de mezcla ascendente húmedos puede incluir N(N – 1) coeficientes. Por ejemplo, los parámetros de mezcla ascendente recibidos pueden incluir no más de N(N – 1)/2 parámetros de mezcla ascendente húmedos asignables independientemente y/o el número de parámetros de mezcla ascendente húmedos puede ser no más de la mitad del número de coeficientes de mezcla ascendente húmedos en el subconjunto de coeficientes de mezcla ascendente húmedos.

En una realización ejemplar, los parámetros de mezcla ascendente recibidos pueden incluir (N – 1) parámetros de mezcla ascendente secos. En la presente realización ejemplar, el subconjunto de los coeficientes de mezcla ascendente secos puede incluir N coeficientes, y el subconjunto de los coeficientes de mezcla ascendente secos se puede determinar en base a los (N – 1) parámetros de mezcla ascendente secos recibidos y en base a una relación predefinida entre los coeficientes en el subconjunto de los coeficientes de mezcla ascendente secos. Por ejemplo, los parámetros de mezcla ascendente recibidos pueden incluir no más de (N – 1) parámetros de mezcla ascendente secos asignables independientemente.

En una realización ejemplar, la clase de matrices predefinidas puede ser una de: matrices triangulares inferiores o superiores, en donde las propiedades conocidas de todas las matrices de la clase incluyen que algunos elementos de la matriz predefinida son cero; matrices simétricas, en donde las propiedades conocidas de todas las matrices de la clase incluyen que algunos elementos de la matriz predefinida (a cada lado de la diagonal principal) son iguales; y productos de una matriz ortogonal y una matriz diagonal, en donde las propiedades conocidas de todas las matrices de la clase incluyen relaciones conocidas entre los elementos de la matriz predefinida. En otras palabras, la clase de matrices predefinidas 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 una realización ejemplar, la matriz predefinida y/o la clase de matrices predefinidas se puede asociar con el formato de codificación indicado, p.ej., permitiendo que el método de decodificación ajuste la determinación del conjunto de coeficientes de mezcla ascendente húmedos en consecuencia.

Según las realizaciones ejemplares, se proporciona un método de decodificación de audio que comprende: recibir señalización que indica una de al menos dos configuraciones de canales predefinidas; en respuesta a la detección de la señalización recibida que indica una primera configuración de canales predefinida, realizar cualquiera de los métodos de decodificación de audio del primer aspecto. El método de decodificación de audio puede comprender, en respuesta a la detección de la señalización recibida que indica una segunda configuración de canales predefinida: recibir una señal de mezcla descendente de dos canales y los parámetros de mezcla ascendente asociados; realizar la reconstrucción paramétrica de una primera señal de audio de tres canales en base a un primer canal de la señal de mezcla descendente y al menos algunos de los parámetros de mezcla ascendente; y realizar la reconstrucción paramétrica de una segunda señal de audio de tres canales en base a un segundo canal de la señal de mezcla descendente y al menos algunos de los parámetros de mezcla ascendente.

La primera configuración de canales predefinida puede corresponder a la señal de audio de M canales representada por la señal de mezcla descendente de dos canales recibida y los parámetros de mezcla ascendente asociados. La segunda configuración de canales predefinida puede corresponder a la primera y segunda señales de audio de tres canales representadas por el primer y segundo canales de la señal de mezcla descendente recibida, respectivamente, y por los parámetros de mezcla ascendente asociados.

La capacidad de recibir señalización que indica una de al menos dos configuraciones de canales predefinidas, y de realizar la reconstrucción paramétrica en base a la configuración de canales indicada, puede permitir que se emplee un formato común para un medio legible por ordenador que porta una representación paramétrica de la señal de audio de M canales o las dos señales de audio de tres canales, desde un lado de codificador a un lado de decodificador.

Según las realizaciones ejemplares, se proporciona un sistema de decodificación de audio que comprende una sección de decodificación configurada para reconstruir una señal de audio de M canales en base a una señal de mezcla descendente de dos canales y los parámetros de mezcla ascendente asociados, donde M ≥ 4. El sistema de decodificación de audio comprende una sección de control configurada para 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 corresponden a respectivas particiones diferentes de los canales de la señal de audio de M canales en respectivos primer y segundo grupos de uno o más canales. En el formato de codificación indicado, un primer canal de la señal de mezcla descendente corresponde a una combinación lineal del 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 del segundo grupo de uno o más canales de la señal de audio de M canales. La sección de decodificación comprende: una sección de descorrelación previa configurada para determinar un conjunto de coeficientes de descorrelación previa en base al formato de codificación indicado, y calcular una señal de entrada de descorrelación como un mapeo lineal de la señal de mezcla descendente, en donde el conjunto de coeficientes de descorrelación previa se aplica a la señal de mezcla descendente; y una sección de descorrelación configurada para generar una señal descorrelacionada en base a la señal de entrada de descorrelación. La sección de decodificación comprende una sección de mezcla configurada para: determinar conjuntos de coeficientes de mezcla ascendente húmedos y secos en base a los parámetros de mezcla ascendente recibidos y el formato de codificación indicado; calcular una señal de mezcla ascendente seca como un mapeo lineal de la señal de mezcla descendente, en donde el conjunto de coeficientes de mezcla ascendente secos se aplica a la señal de mezcla descendente; calcular una señal de mezcla ascendente húmeda como un mapeo lineal de la señal descorrelacionada, en donde el conjunto de coeficientes de mezcla ascendente húmedos se aplica a la señal descorrelacionada; y combinar las señales de mezcla ascendente secas y húmedas para obtener una señal reconstruida multidimensional correspondiente a la señal de audio de M canales a reconstruir.

En una realización ejemplar, el sistema de decodificación de audio puede comprender además una sección de decodificación adicional configurada para reconstruir una señal de audio de M canales adicional en base a una señal de mezcla descendente de dos canales adicional y los parámetros de mezcla ascendente adicionales asociados. La sección de control se puede configurar para recibir señalización que indica uno seleccionado de al menos dos formatos de codificación de la señal de audio de M canales adicional. Los formatos de codificación de la señal de audio de M canales adicional puede corresponder a respectivas particiones diferentes de los canales de la señal de audio de M canales adicional en respectivos primer y segundo grupos de uno o más canales. En el formato de codificación indicado de la señal de audio de M canales adicional, un primer canal de la señal de mezcla descendente adicional puede corresponder a una combinación lineal del 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 del segundo grupo de uno o más canales de la señal de audio de M canales adicional. La sección de decodificación adicional puede comprender: una sección de descorrelación previa adicional configurada para determinar un conjunto de coeficientes de descorrelación previa adicional en base al formato de codificación indicado de la señal de audio de M canales adicional, y calcular una señal de entrada de descorrelación adicional como un mapeo lineal de la señal de mezcla descendente adicional, en donde el conjunto de coeficientes de descorrelación previa adicional se aplica a la señal de mezcla descendente adicional; y una sección de descorrelación adicional configurada para generar una señal descorrelacionada adicional en base a la señal de entrada de descorrelación adicional. La sección de decodificación adicional puede comprender además una sección de mezcla adicional configurada para : determinar conjuntos de coeficientes de mezcla ascendente húmedos y secos adicionales en base a los parámetros de mezcla ascendente adicionales recibidos y el formato de codificación indicado de la señal de audio de M canales adicional; calcular una señal de mezcla ascendente seca adicional como un mapeo lineal de la señal de mezcla descendente adicional, en donde el conjunto de coeficientes de mezcla ascendente secos adicional se aplica a la señal de mezcla descendente adicional; calcular una señal de mezcla ascendente húmeda adicional como un mapeo lineal de la señal descorrelacionada adicional, en donde el conjunto de coeficientes de mezcla ascendente húmedos adicional se aplica a la señal descorrelacionada adicional; y combinar las señales de mezcla ascendente secas y húmedas adicionales para obtener una señal reconstruida multidimensional adicional correspondiente a la señal de audio de M canales adicional a reconstruir.

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

En la presente realización ejemplar, la sección de decodificación adicional, la sección de descorrelación previa adicional, la sección de descorrelación adicional y la sección de mezcla adicional pueden por ejemplo ser funcionalmente equivalentes a (o estar análogamente configuradas como) la sección de decodificación, la sección de descorrelación previa, la sección de descorrelación y la sección de mezcla, respectivamente. Alternativamente, al

menos una de la sección de decodificación adicional, la sección de descorrelación previa adicional, la sección de descorrelación adicional y la sección de mezcla adicional se pueden por ejemplo configurar para realizar al menos un tipo diferente de interpolación que la realizada por la sección correspondiente de la sección de decodificación, la sección de descorrelación previa, la sección de descorrelación y la sección de mezcla.

Por ejemplo, la señalización recibida puede indicar diferentes formatos de codificación para la señal de audio de M canales y la señal de audio de M canales adicional. Alternativamente, los formatos de codificación de las dos señales de audio de M canales pueden por ejemplo coincidir siempre, y la señalización recibida puede indicar uno seleccionado de al menos dos formatos de codificación comunes para las dos señales de audio de M canales.

Los esquemas de interpolación empleados para las transiciones graduales entre los coeficientes de descorrelación previa, en respuesta a la conmutación entre formatos de codificación de la señal de audio de M canales, pueden coincidir con, o ser diferentes de los esquemas de interpolación empleados las para transiciones graduales entre los coeficientes de descorrelación previa adicionales, en respuesta a la conmutación entre formatos de codificación de la señal de audio de M canales adicional.

Similarmente, los esquemas de interpolación empleados para la interpolación de valores de los coeficientes de mezcla ascendente húmedos y secos, en respuesta a la conmutación entre formatos de codificación de la señal de audio de M canales, pueden coincidir con, o ser diferentes de los esquemas de interpolación empleados para la interpolación de valores de los coeficientes de mezcla ascendente húmedos y secos adicionales, en respuesta a la conmutación entre formatos de codificación de la señal de audio de M canales adicional.

En una realización ejemplar, el sistema de decodificación de audio puede comprender además un demultiplexor configurado para extraer, de un flujo de bis, la señal de mezcla descendente, los parámetros de mezcla ascendente asociados con la señal de mezcla descendente, y un canal de audio codificado de manera discreta. El sistema de decodificación puede comprender además una sección de decodificación de un solo canal operable para decodificar el canal de audio codificado de manera discreta. El canal de audio codificado de manera discreta se puede por ejemplo codificar en el flujo de bits utilizando un códec de audio perceptual tal como Dolby Digital, MPEG AAC, o desarrollos de los mismos, y la sección de decodificación de un solo canal puede por ejemplo comprender un decodificador central para decodificar el canal de audio codificado de manera discreta. La sección de decodificación de un solo canal puede por ejemplo ser operable para decodificar el canal de audio codificado de manera discreta independientemente de la sección de decodificación.

Según las realizaciones ejemplares, 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.

II. Visión general – Lado de codificador

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

Según las realizaciones ejemplares, se proporciona un método de codificación de audio que comprende: recibir una señal de audio de M canales, para la que $M \geq 4$. El método de codificación de audio comprende seleccionar repetidamente uno de al menos dos formatos de codificación en base a cualquier criterio de selección adecuado, p.ej., propiedades de la señal, carga del sistema, preferencia del usuario, condiciones de la red. La selección se puede repetir una vez para cada trama de tiempo de la señal de audio o una vez para cada trama de tiempo enésima, posiblemente conduciendo a la selección de un formato diferente del elegido inicialmente; alternativamente, la selección puede ser accionada por evento. Los formatos de codificación corresponden a respectivas particiones diferentes de los canales de la señal de audio de M canales en respectivos primer y segundo grupos de uno o más canales. En cada uno de los formatos de codificación, una señal de mezcla descendente de dos canales incluye un primer canal formado como una combinación lineal del primer grupo de uno o más canales de la señal de audio de M canales, y un segundo canal formado como una combinación lineal del segundo grupo de uno o más canales de la señal de audio de M canales. Para el formato de codificación seleccionado, se calcula el canal de mezcla descendente en base a la señal de audio de M canales. Una vez calculado, la señal de mezcla descendente del formato de codificación actualmente seleccionado es emitida, como lo es la señalización que indica el formato de codificación actualmente seleccionado e información lateral que permite la reconstrucción paramétrica de la señal de audio de M canales. Si la selección da como resultado un cambio de un primer formato de codificación seleccionado a un segundo formato de codificación seleccionado distinto, se puede iniciar una transición, por lo que se emite un desvanecimiento cruzado de la señal de mezcla descendente según el primer formato de codificación seleccionado y la señal de mezcla descendente según el segundo formato de codificación seleccionado. En este contexto, un desvanecimiento cruzado puede ser una interpolación de tiempo lineal o no lineal de dos señales. Como ejemplo,

$y(t) = tx_1(t) + (1- t)\, x_2(t),\ t \in [0,1]$

proporciona un desvanecimiento cruzado *y* de la función x_2 a la función x_1 linealmente a lo largo del tiempo, en donde x_1, x_2 pueden ser funciones de valor vectorial de tiempo que representan las señales de mezcla descendente según los respectivos formatos de codificación. Para simplificar la notación, el intervalo de tiempo, sobre el que se lleva a cabo el desvanecimiento cruzado, se ha reajustado a [0,1], en donde t= 0 representa el comienzo del desvanecimiento cruzado y t= 1 representa el punto en el tiempo en el que se ha completado el desvanecimiento cruzado.

La ubicación de los puntos t= 0 y t = 1 en unidades físicas puede ser importante para la calidad de salida percibida del audio reconstruido. Como una posible directriz para ubicar el desvanecimiento cruzado, el comienzo se puede producir tan pronto como sea posible después de que se haya determinado la necesidad de un formato diferente, y/o el desvanecimiento cruzado se puede completar en el menor tiempo posible que sea perceptualmente imperceptible. Como tal, para las implementaciones donde la selección de un formato de codificación se repite cada trama, algunas realizaciones ejemplares proporcionan que el desvanecimiento cruzado comience (t=0) al comienzo de la trama, y tenga su punto final (t=1) tan cerca como sea posible pero lo suficientemente distante como para que un oyente promedio sea incapaz de notar los artefactos o degradaciones debidas a una transición entre dos reconstrucciones de una señal de audio de M canales común (con contenido típico) en base a dos formatos de codificación distintos. En una realización ejemplar, la señal de mezcla descendente emitida por el método de codificación de audio es segmentada en tramas de tiempo y un desvanecimiento cruzado puede ocupar una trama. En otra realización ejemplar, la señal de mezcla descendente emitida por el método de codificación de audio es segmentada en tramas de tiempo que se solapan y la duración de un desvanecimiento cruzado corresponde al paso de una trama de tiempo a la siguiente.

En las realizaciones ejemplares, la señalización que indica el formato de codificación actualmente seleccionado se puede codificar en una base de una trama por trama. Alternativamente, la señalización puede ser diferencial de tiempo en el sentido de que tal señalización se puede omitir en una o más tramas consecutivas si no hay ningún cambio en el formato de codificación seleccionado. En el lado de decodificador, tal secuencia de tramas se puede interpretar que significa que el formato de codificación señalizado más recientemente permanece seleccionado.

Dependiendo del contenido de audio de la señal de audio de M canales, pueden ser adecuadas diferentes particiones de los canales de la señal de audio de M canales en primer y segundo grupos, representadas por los respectivos canales de la señal de mezcla descendente, para capturar y codificar eficientemente la señal de audio de M canales, y para preservar la fidelidad cuando se reconstruye esta señal a partir de la señal de mezcla descendente y los parámetros de mezcla ascendente asociados. La fidelidad de la señal de audio de M canales reconstruida se puede por lo tanto aumentar seleccionando un formato de codificación apropiado, concretamente el más adecuado de una serie de formatos de codificación predefinidos.

En una realización ejemplar, la información lateral incluye coeficientes de mezcla ascendente secos y húmedos, en el mismo sentido que se han utilizado estos términos anteriormente en esta descripción. Salvo por razones de implementación específicas, es generalmente suficiente calcular la información lateral (en particular, los coeficientes de mezcla ascendente secos y húmedos) para el formato de codificación actualmente seleccionado. En particular, el conjunto de coeficientes de mezcla ascendente secos (que se puede representar como una matriz de dimensiones M x 2) puede definir un mapeo lineal de la respectiva señal de mezcla descendente que se aproxima a la señal de audio de M canales. El conjunto de coeficientes de mezcla ascendente húmedos (que se puede representar como una matriz de dimensiones M x P, donde P, el número de descorreladores, se puede establecer a P = M – 2) define un mapeo lineal de la señal descorrelacionada de tal manera que una covarianza de la señal obtenida por dicho mapeo lineal de la señal descorrelacionada suplementa una covarianza de la señal de audio de M canales aproximada por el mapeo lineal de la señal de mezcla descendente del formato de codificación seleccionado. El mapeo de la señal descorrelacionada que define el conjunto de coeficientes de mezcla ascendente húmedos suplementará la covarianza de la señal de audio de M canales (aproximada) en el sentido de que la covarianza de la suma de la señal de audio de M canales y el mapeo de la señal descorrelacionada es típicamente más próxima a la covarianza de la señal de audio de M canales recibida. Un efecto de añadir la covarianza suplementaria puede ser una fidelidad mejorada de una señal reconstruida en el lado de decodificador.

El mapeo lineal de la señal de mezcla descendente proporciona una aproximación de la señal de audio de M canales. Al reconstruir la señal de audio de M canales en un lado de decodificador, la señal descorrelacionada se emplea para aumentar la dimensionalidad del contenido de audio de la señal de mezcla descendente, y la señal obtenida por el mapeo lineal de la señal descorrelacionada se combina con la señal obtenida por el mapeo lineal de la señal de mezcla descendente para mejorar la fidelidad de la aproximación de la señal de audio de M canales. Dado que la señal descorrelacionada se determina en base a al menos un canal de la señal de mezcla descendente, y no comprende ningún contenido de audio de la señal de audio de M canales que no esté ya disponible en la señal de mezcla descendente, la diferencia entre la covarianza de la señal de audio de M canales recibida y la covarianza de la señal de audio de M canales aproximada por el mapeo lineal de la señal de mezcla descendente, puede ser indicativa no solo de una fidelidad de la señal de audio de M canales aproximada por el mapeo lineal de la señal de mezcla descendente, sino también de una fidelidad de la señal de audio de M canales reconstruida utilizando tanto la señal de mezcla descendente como la señal descorrelacionada. En particular, una diferencia reducida entre la covarianza

de la señal de audio de M canales recibida y la covarianza de la señal de audio de M canales aproximada por el mapeo lineal de la señal de mezcla descendente puede ser indicativa de una fidelidad mejorada de la señal de audio de M canales reconstruida. El mapeo de la señal descorrelacionada que define el conjunto de coeficientes de mezcla ascendente húmedos suplementa la covarianza de la señal de audio de M canales (obtenida a partir de la señal de mezcla descendente) en el sentido de que la covarianza de la suma de la señal de audio de M canales y el mapeo de la señal descorrelacionada es más próxima a la covarianza de la señal de audio de M canales recibida. Seleccionar uno de los formatos de codificación en base a las respectivas diferencias calculadas permite por lo tanto mejorar la fidelidad de la señal de audio de M canales reconstruida.

Se apreciará que se puede seleccionar el formato de codificación p.ej., directamente en base a las diferencias calculadas, o en base a los coeficientes y/o valores determinados en base a las diferencias calculadas.

También se apreciará que se puede seleccionar el formato de codificación en base a p.ej., los respectivos parámetros de mezcla ascendente secos calculados además de las respectivas diferencias calculadas.

El conjunto de coeficientes de mezcla ascendente secos se puede por ejemplo determinar a través de una aproximación del error cuadrático medio mínimo bajo la suposición de que solo está disponible para la reconstrucción la señal de mezcla descendente, i.e., bajo la suposición de que la señal descorrelacionada no se emplea para la reconstrucción.

Las diferencias calculadas pueden por ejemplo ser diferencias entre una matriz de covarianza de la señal de audio de M canales recibida y las matrices de covarianza de la señal de audio de M canales aproximada por los respectivos mapeos lineales de la señal de mezcla descendente de los diferentes formatos de codificación. Seleccionar uno de los formatos de codificación puede por ejemplo incluir calcular las normas de la matriz para las respectivas diferencias entre las matrices de covarianza, y seleccionar uno de los formatos de codificación en base a las normas de la matriz calculadas, p.ej., seleccionar un formato de codificación asociado con una mínima de las normas de la matriz calculadas.

La señal descorrelacionada puede por ejemplo incluir al menos un canal y como máximo M – 2 canales.

El que el conjunto de coeficientes de mezcla ascendente secos defina un mapeo lineal de la señal de mezcla descendente que se aproxima a la señal de mezcla descendente de M canales significa que se obtiene una aproximación de la señal de mezcla descendente de M canales aplicando una transformación lineal a la señal de mezcla descendente. Esta transformación lineal toma los dos canales de la señal de mezcla descendente como entrada y proporciona M canales como salida, y los coeficientes de mezcla ascendente secos son coeficientes que definen las propiedades cuantitativas de esta transformación lineal.

Similarmente, los parámetros de mezcla ascendente húmedos definen las propiedades cuantitativas de una transformación lineal que toma el(los) canal(es) de la señal descorrelacionada como entrada, y proporciona M canales como salida.

En una realización ejemplar, los parámetros de mezcla ascendente húmedos se pueden determinar de tal manera que una covarianza de la señal obtenida por el mapeo lineal (que definen los parámetros de mezcla ascendente húmedos) de la señal descorrelacionada se aproxima a una diferencia entre la covarianza de la señal de audio de M canales recibida y una covarianza de la señal de audio de M canales aproximada por el mapeo lineal de la señal de mezcla descendente del formato de codificación seleccionado. Puesto de manera diferente, la covarianza de una suma de un primer mapeo lineal (definido por los parámetros de mezcla ascendente secos) de la señal de mezcla descendente y un segundo mapeo lineal (definido por los parámetros de mezcla ascendente húmedos, determinados de acuerdo con esta realización ejemplar) de la señal descorrelacionada será próxima a la covarianza de la señal de audio de M canales que constituye la entrada al método de codificación de audio descrito anteriormente en la presente memoria. Determinar los coeficientes de mezcla ascendente húmedos de acuerdo con la presente realización ejemplar puede mejorar la fidelidad de la señal de audio de M canales reconstruida.

Alternativamente, los parámetros de mezcla ascendente húmedos se pueden determinar de tal manera que una covarianza de la señal obtenida por el mapeo lineal de la señal descorrelacionada se aproxima a una porción de una diferencia entre la covarianza de la señal de audio de M canales recibida y una covarianza de la señal de audio de M canales aproximada por el mapeo lineal de la señal de mezcla descendente del formato de codificación seleccionado. Si, por ejemplo, un número limitado de descorreladores están disponibles en un lado de decodificador, puede no ser posible reestablecer completamente la covarianza de la señal de audio de M canales recibida. En tal ejemplo, los parámetros de mezcla ascendente húmedos adecuados para la reconstrucción parcial de la covarianza de la señal de audio de M canales, empleando un número reducido de descorreladores, se pueden determinar en el lado de codificador.

En una realización ejemplar, el método de codificación de audio puede comprender además, para cada uno de los al menos dos formatos de codificación: determinar un conjunto de coeficientes de mezcla ascendente húmedos que junto con los coeficientes de mezcla ascendente secos (de ese formato de codificación) permita la reconstrucción paramétrica de la señal de audio de M canales a partir de la señal de mezcla descendente (de ese formato de

codificación) y a partir de una señal descorrelacionada determinada en base a la señal de mezcla descendente (de ese formato), en donde el conjunto de coeficientes de mezcla ascendente húmedos define un mapeo lineal de la señal descorrelacionada de tal manera que una covarianza de una señal obtenida por el mapeo lineal de la señal descorrelacionada se aproxima a una diferencia entre la covarianza de la señal de audio de M calanes recibida y una covarianza de la señal de audio de M canales aproximada por el mapeo lineal de la señal de mezcla descendente (en ese formato). En la presente realización ejemplar, el formato de codificación seleccionado se puede seleccionar en base a los valores de los respectivos conjuntos determinados de coeficientes de mezcla ascendente húmedos.

Una indicación de la fidelidad de la señal de audio de M canales reconstruida se puede por ejemplo obtener en base a los coeficientes de mezcla ascendente húmedos determinados. La selección de un formato de codificación se puede por ejemplo basar en sumas ponderadas o no ponderadas de los coeficientes de mezcla ascendente húmedos determinados, en sumas ponderadas o no ponderadas de magnitudes de los coeficientes de mezcla ascendente húmedos determinados, y/o en sumas ponderadas o no ponderadas de cuadrados de los coeficientes de mezcla ascendente húmedos determinados, p.ej., también en base a las sumas correspondientes de los respectivos coeficientes de mezcla ascendente secos calculados.

Los parámetros de mezcla ascendente húmedos se pueden por ejemplo calcular para una pluralidad de bandas de frecuencia de la señal de M canales, y la selección de un formato de codificación se puede por ejemplo basar en los valores de los respectivos conjuntos determinados de coeficientes de mezcla ascendente húmedos en las respectivas bandas de frecuencia.

En una realización ejemplar, una transición entre un primer y un segundo formato de codificación incluye emitir valores discretos de los coeficientes de mezcla ascendente secos y húmedos del primer formato de codificación en una trama de tiempo y del segundo formato de codificación en una trama de tiempo posterior. Las funcionalidades en un decodificador que finalmente reconstruye la señal de M canales pueden incluir interpolación de los coeficientes de mezcla ascendente entre los valores discretos de salida. En virtud de tales funcionalidades del lado de decodificador, resultará de manera efectiva un desvanecimiento cruzado del primer al segundo formato de codificación. Como el desvanecimiento cruzado aplicado a la señal de mezcla descendente, como se describió anteriormente, tal desvanecimiento cruzado puede conducir a una transición menos perceptible entre los formatos de codificación cuando se reconstruye la señal de audio de M canales.

Se entiende que los coeficientes empleados para calcular la señal de mezcla descendente en base a la señal de audio de M canales se pueden interpolar, i.e., a partir de valores asociados con una trama donde la señal de mezcla descendente se calcula según un primer formato de codificación, a valores asociados con una trama donde la señal de mezcla descendente se calcula según el segundo formato de codificación. Al menos si la mezcla descendente tiene lugar en el dominio del tiempo, un desvanecimiento cruzado de mezcla descendente resultante de la interpolación de coeficientes del tipo indicado será equivalente a un desvanecimiento cruzado resultante de la interpolación realizada directamente sobre las respectivas señales de mezcla descendente. Se recuerda que los valores de los coeficientes empleados para calcular la señal de mezcla descendente típicamente no son dependientes de la señal sino que pueden ser predefinidos para cada uno de los formatos de codificación disponibles.

Volviendo al desvanecimiento cruzado de la señal de mezcla descendente y los coeficientes de mezcla ascendente, se considera ventajoso asegurar la sincronicidad entre los dos desvanecimientos cruzados. Preferiblemente, los respectivos periodos de transición para la señal de mezcla descendente y los coeficientes de mezcla ascendente pueden coincidir. En particular, las entidades responsables de los respectivos desvanecimientos cruzados pueden ser controladas por un flujo común de datos de control. Tales datos de control pueden incluir puntos de inicio y puntos de finalización del desvanecimiento cruzado, y opcionalmente una forma de onda de desvanecimiento cruzado, tal como lineal, no lineal, etc. En el caso de los coeficientes de mezcla ascendente, la forma de onda de desvanecimiento cruzado puede ser dada por una regla de interpolación predeterminada que rige el comportamiento de un dispositivo de decodificación; los puntos de inicio y finalización de los desvanecimientos cruzados pueden sin embargo ser controlados implícitamente por las posiciones en las que se definen y/o emiten los valores discretos de los coeficientes de mezcla ascendente. La similitud en la dependencia del tiempo de los dos procesos de desvanecimiento cruzado asegura una buena correspondencia entre la señal de mezcla descendente y los parámetros proporcionados para su reconstrucción, lo que puede conducir a una reducción de artefactos en el lado de decodificador.

En una realización ejemplar, la selección de un formato de codificación se basa en comparar la diferencia, en términos de covarianza, de la señal de M canales recibida y la señal de M canales reconstruida en base a la señal de mezcla descendente. En particular, la reconstrucción puede ser igual a un mapeo lineal de la señal de mezcla descendente definida por los coeficientes de mezcla ascendente secos solamente, es decir, sin una contribución de una señal que ha sido determinada utilizando descorrelación (p.ej., para aumentar la dimensionalidad del contenido de audio de la señal de mezcla descendente). En particular, no se debe considerar en la comparación ninguna contribución del mapeo lineal definido por cualquier conjunto de coeficientes de mezcla ascendente húmedos. Dicho de otra manera, la comparación se hace como si no hubiera estado disponible ninguna señal descorrelacionada. Esta base para la selección puede favorecer un formato de codificación que actualmente permite una reproducción más fiel. Opcionalmente, después de que se haya realizado esta comparación y se haya tomado una decisión en cuanto a la selección de un formato de codificación, se determina un conjunto de coeficientes de mezcla ascendente húmedos.

Una ventaja asociada con este proceso es que no hay determinación duplicada de los coeficientes de mezcla ascendente húmedos para una sección dada de la señal de audio de M canales recibida.

En una variación de la realización ejemplar descrita en el párrafo anterior, se calculan los coeficientes de mezcla ascendente secos y húmedos para todos los formatos de codificación y se utiliza una medida cuantitativa de los coeficientes de mezcla ascendente húmedos como base para la selección de un formato de codificación. De hecho, una cantidad calculada en base a los coeficientes de mezcla ascendente húmedos determinados puede proporcionar una indicación (inversa) de la fidelidad de la señal de audio de M canales reconstruida. La selección de un formato de codificación se puede por ejemplo basar en sumas ponderadas o no ponderadas de los coeficientes de mezcla ascendente húmedos determinados, en sumas ponderadas o no ponderadas de magnitudes de los coeficientes de mezcla ascendente húmedos determinados, y/o en sumas ponderadas o no ponderadas de cuadrados de los coeficientes de mezcla ascendente húmedos determinados. Cada una de estas opciones se puede combinar con las sumas correspondientes de los respectivos coeficientes de mezcla ascendente secos calculados. Los parámetros de mezcla ascendente húmedos se pueden por ejemplo calcular para una pluralidad de bandas de frecuencia de la señal de M canales, y la selección de un formato de codificación se puede por ejemplo basar en valores de los respectivos conjuntos determinados de coeficientes de mezcla ascendente húmedos en las respectivas bandas de frecuencia.

En una realización ejemplar, el método de decodificación de audio puede comprender además: para cada uno de los al menos dos formatos de codificación, calcular una suma de cuadrados de los coeficientes de mezcla ascendente húmedos correspondientes y una suma de cuadrados de los coeficientes de mezcla ascendente secos correspondientes. En la presente realización ejemplar, el formato de codificación seleccionado se puede seleccionar en base a las sumas de cuadrados calculadas. Lon inventores se han dado cuenta de que las sumas calculadas de cuadrados pueden proporcionar una indicación particularmente buena de la pérdida de fidelidad, tal como es percibida por un oyente, que se produce cuando se reconstruye la señal de audio de M canales en base a la mezcla de las contribuciones húmedas y secas.

Por ejemplo, se puede formar una relación para cada formato de codificación, en base a las sumas de cuadrados calculadas para el respectivo formato de codificación, y el formato de codificación seleccionado se puede asociar con una mínima o máxima de las relaciones formadas. Formar una relación puede por ejemplo incluir dividir, por un lado, una suma de cuadrados de los coeficientes de mezcla ascendente húmedos por, por otro lado, una suma de una suma de cuadrados de los coeficientes de mezcla ascendente secos y una suma de cuadrados de los coeficientes de mezcla ascendente húmedos. Alternativamente, se puede formar la relación dividiendo una suma de cuadrados de los coeficientes de mezcla ascendente húmedos por una suma de cuadrados de los coeficientes de mezcla ascendente secos.

En una realización ejemplar, el método proporciona codificación de una señal de audio de M canales y al menos una señal de audio (de M_2 canales) asociada. Las señales de audio se pueden asociar en el sentido de que describen una escena de audio común, p.ej., por haber sido grabadas simultáneamente o generadas en un proceso de creación común. Las señales de audio no necesitan ser codificadas por medio de una señal de mezcla descendente común, sino que se pueden codificar en procesos separados. En tal configuración, la selección de uno de los formatos de codificación tiene en cuenta adicionalmente los datos relativos a dicho al menos un canal de audio adicional, y el formato de codificación así seleccionado se debe utilizar para codificar tanto la señal de audio de M canales como la señal de audio (de M_2 canales) asociada.

En una realización ejemplar, la señal de mezcla descendente emitida por el método de codificación de audio se puede segmentar en tramas de tiempo, la selección de un formato de codificación se puede realizar una vez por trama, y el formato de codificación seleccionado se puede mantener durante al menos un número predefinido de tramas de tiempo antes de que se seleccione un formato de codificación diferente. La selección de un formato de codificación para una trama se puede realizar por cualquiera de los métodos indicados anteriormente, p.ej., teniendo en cuenta las diferencias entre las covarianzas, teniendo en cuenta los valores de los coeficientes de mezcla ascendente húmedos para los formatos de codificación disponibles, y similares. Manteniendo el formato de codificación seleccionado para un número mínimo de tramas de tiempo, se pueden por ejemplo evitar saltos repetidos hacia atrás y hacia adelante entre formatos de codificación. La presente realización ejemplar puede por ejemplo mejorar la calidad de reproducción, tal como es percibida por un oyente, de la señal de audio de M canales reconstruida.

El número mínimo de tramas de tiempo puede ser por ejemplo 10.

La señal de audio de M canales recibida se puede por ejemplo almacenar en memoria intermedia para el número mínimo de tramas de tiempo, y la selección de un formato de codificación se puede por ejemplo realizar en base a una decisión por mayoría sobre una ventana móvil que comprende un número de tramas de tiempo elegidas en vista de dicho número mínimo de tramas que un formato de codificación seleccionado se debe mantener. Una implementación de tal funcionalidad estabilizadora puede incluir uno de los diversos filtros de suavizado, en particular filtros de suavizado de respuesta finita al impulso que son conocidos en el procesamiento de señales digitales. Como alternativa a este enfoque, el formato de codificación se puede conmutar a un nuevo formato de codificación cuando se descubre que se ha seleccionado el nuevo formato de codificación para dicho número mínimo de tramas en la secuencia. Para hacer cumplir este criterio, se puede aplicar una ventana de tiempo móvil con el número mínimo de tramas consecutivas a las selecciones de formato de codificación anteriores, p.ej., para las tramas almacenadas en

memoria intermedia. Si, después de una secuencia de tramas de un primer formato de codificación, ha permanecido seleccionado un segundo formato de codificación para cada trama dentro de la ventana móvil, la transición al segundo formato de codificación se confirma y tiene efecto desde el comienzo de la ventana móvil hacia adelante. Una implementación de la funcionalidad estabilizadora anterior puede incluir una máquina de estado.

En una realización ejemplar, se proporciona una representación compacta de los parámetros de mezcla ascendente secos y húmedos, que entre otras cosas incluye generar una matriz intermedia que en virtud de pertenecer a una clase de matrices predefinidas está determinada de manera única por un número más pequeño de parámetros que los elementos en la matriz. Los aspectos de esta representación compacta se han descrito en secciones anteriores de esta descripción, y con particular referencia a la Solicitud de Patente Provisional de EE.UU. Nº 61/974,544, primer inventor nombrado: Lars Villemoes; fecha de presentación: 3 de abril de 2014.

En una realización ejemplar, en el formato de codificación seleccionado, el primer grupo de uno o más canales de la señal de audio de M canales puede constar de N canales, donde $N \geq 3$. El primer grupo de uno o más canales puede ser reconstruible a partir del primer canal de la señal de mezcla descendente y N – 1 canales de la señal descorrelacionada aplicando al menos algunos de los coeficientes de mezcla ascendente húmedos y secos.

En la presente realización ejemplar, determinar el conjunto de coeficientes de mezcla ascendente secos del formato de codificación seleccionado puede incluir determinar un subconjunto de los coeficientes de mezcla ascendente secos del formato de codificación seleccionado para definir un mapeo lineal del primer canal de la señal de mezcla descendente del formato de codificación seleccionado que aproxime el primer grupo de uno o más canales del formato de codificación seleccionado.

En la presente realización ejemplar, determinar el conjunto de coeficientes de mezcla ascendente húmedos del formato de codificación seleccionado puede incluir: determinar una matriz intermedia en base a una diferencia entre una covarianza del primer grupo de uno o más canales del formato de codificación seleccionado recibido, y una covarianza del primer grupo de uno o más canales del formato de codificación seleccionado aproximado por el mapeo lineal del primer canal de la señal de mezcla descendente del formato de codificación seleccionado. Cuando se multiplica por una matriz predefinida, la matriz intermedia puede corresponder a un subconjunto de los coeficientes de mezcla ascendente húmedos del formato de codificación seleccionado que define un mapeo lineal de los N – 1 canales de la señal descorrelacionada como parte de la reconstrucción paramétrica del primer grupo de uno o más canales del formato de codificación seleccionado. El subconjunto de los coeficientes de mezcla ascendente húmedos del formato de codificación seleccionado puede incluir más coeficientes que el número de elementos en la matriz intermedia.

En la presente realización ejemplar, los parámetros de mezcla ascendente de salida pueden incluir un conjunto de parámetros de mezcla ascendente de un primer tipo, denominados en la presente memoria como parámetros de mezcla ascendente secos, a partir de los cuales se puede derivar el subconjunto de coeficientes de mezcla ascendente secos, y un conjunto de parámetros de mezcla ascendente de un segundo tipo, denominados en la presente memoria como parámetros de mezcla ascendente húmedos, que definen de manera única la matriz intermedia siempre que la matriz intermedia pertenezca a una clase de matrices predefinidas. La matriz intermedia puede tener más elementos que el número de elementos en el subconjunto de los parámetros de mezcla ascendente húmedos del formato de codificación seleccionado.

En la presente realización ejemplar, una copia de la reconstrucción paramétrica del primer grupo de uno o más canales en un lado de decodificador incluye, a modo de contribución, una señal de mezcla ascendente seca formada por el mapeo lineal del primer canal de la señal de mezcla descendente y, a modo de contribución adicional, una señal de mezcla ascendente húmeda formada por el mapeo lineal de los N – 1 canales de la señal descorrelacionada. El subconjunto de coeficientes de mezcla ascendente secos define el mapeo lineal del primer canal de la señal de mezcla descendente y el subconjunto de coeficientes de mezcla ascendente húmedos define el mapeo lineal de la señal descorrelacionada. Emitiendo los parámetros de mezcla ascendente húmedos con un número inferior que el número de coeficientes en el subconjunto de coeficientes de mezcla ascendente húmedos, y a partir de los cuales se puede derivar el subconjunto de coeficientes de mezcla ascendente húmedos en base a la matriz predefinida y la clase de matrices predefinidas, se puede reducir la cantidad de información enviada a un lado de decodificador para permitir 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 por ejemplo determinar de tal manera que una covarianza de la señal obtenida por el mapeo lineal de los N – 1 canales de la señal descorrelacionada suplementa la covarianza del primer grupo de uno o más canales aproximados por el mapeo lineal del primer canal de la matriz de mezcla descendente.

Cómo determinar y emplear la matriz predefinida y la clase de matrices predefinidas se describe en más detalle en la página 16, línea 15 a página 20, línea 2 en la Solicitud de Patente Provisional de EE.UU. Nº 61/974,544 anteriormente mencionada. Véase en particular la ecuación (9) en la misma para ejemplos de la matriz predefinida.

En una realización ejemplar, determinar la matriz intermedia puede incluir determinar la matriz intermedia de tal manera que una covarianza de la señal obtenida por el mapeo lineal de los N – 1 canales de la señal

descorrelacionada, definida por el subconjunto de coeficientes de mezcla ascendente húmedos, se aproxima, o coincide sustancialmente con, la diferencia entre la covarianza del primer grupo de uno o más canales recibidos y la covarianza del primer grupo de uno o más canales aproximados por el mapeo lineal del primer canal de la señal de mezcla descendente. En otras palabras, la matriz intermedia se puede determinar de tal manera que una copia de la reconstrucción del primer grupo de uno o más canales, obtenidos como una suma de una señal de mezcla ascendente seca formada por el mapeo lineal del primer canal de la señal de mezcla descendente y una señal de mezcla ascendente húmeda formada por el mapeo lineal de los N – 1 canales de la señal descorrelacionada reestablece completamente, o al menos aproximadamente, la covarianza del primer grupo de uno o más canales recibidos.

En una realización ejemplar, los parámetros de mezcla ascendente húmedos pueden incluir no más de N (N – 1)/2 parámetros de mezcla ascendente húmedos asignables independientemente. En la presente realización ejemplar, la matriz intermedia puede tener $(N – 1)^2$ elementos de matriz y puede estar definida de manera única por los parámetros de mezcla ascendente húmedos siempre que la matriz intermedia pertenezca a la clase de matrices predefinidas. En la presente realización ejemplar, el subconjunto de coeficientes de mezcla ascendente húmedos puede incluir N (N – 1) coeficientes.

En una realización ejemplar, el subconjunto de coeficientes de mezcla ascendente secos puede incluir N coeficientes. En la presente realización ejemplar, los parámetros de mezcla ascendente secos pueden incluir no más de N – 1 parámetros de mezcla ascendente secos, y el subconjunto de coeficientes de mezcla ascendente secos se puede derivar a partir de los N – 1 parámetros de mezcla ascendente secos utilizando una regla predefinida.

En una realización ejemplar, el subconjunto determinado de coeficientes de mezcla ascendente secos puede definir un mapeo lineal del primer canal de la señal de mezcla descendente correspondiente a una aproximación del error cuadrático medio mínimo del primer grupo de uno o más canales, i.e., entre el conjunto de mapeos lineales del primer canal de la señal de mezcla descendente, el conjunto determinado de coeficientes de mezcla ascendente secos puede definir el mapeo lineal que mejor se aproxima al primer grupo de uno o más canales en un sentido cuadrático medio mínimo.

En una realización ejemplar, 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 audio de dos canales y los parámetros de mezcla ascendente asociados, donde M ≥ 4. La sección de codificación comprende: una sección de mezcla descendente configurada para, para al menos uno de al menos dos formatos de codificación correspondientes a las respectivas particiones diferentes de los canales de la señal de audio de M canales en respectivos primer y segundo grupos de uno o más canales, calcular, de acuerdo con el formato de codificación, 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 del 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 del segundo grupo de uno o más canales de la señal de audio de M canales.

El sistema de codificación de audio comprende además una sección de control configurada para seleccionar uno de los formatos de codificación en base a cualquier criterio adecuado, p.ej., propiedades de la señal, carga del sistema, preferencia del usuario, condiciones de la red. El sistema de codificación de audio comprende además un interpolador de mezcla descendente, que realiza un desvanecimiento cruzado de la señal de mezcla descendente entre dos formatos de codificación cuando ha sido ordenada una transición por la sección de control. Durante tal transición, se pueden calcular las señales de mezcla descendente para ambos formatos de codificación. Además de la señal de mezcla descendente - o cuando sea aplicable un desvanecimiento cruzado de la misma – el sistema de codificación de audio al menos emite señalización que indica un formato de codificación actualmente seleccionado e información lateral que permite la reconstrucción paramétrica de la señal de audio de M canales en base a la señal de mezcla descendente. Si el sistema comprende múltiples secciones de codificación que operan en paralelo, p.ej., para codificar respectivos grupos de canales de audio, entonces la sección de control se puede implementar de forma autónoma de cada una de estas y siendo responsable de seleccionar un formato de codificación común a ser utilizado por cada una de las secciones de codificación.

En una realización ejemplar, se proporciona un producto de programa de ordenador que comprende un medio legible por ordenador con instrucciones para realizar cualquiera de los métodos descritos en esta sección.

III. Realizaciones ejemplares

Las Figuras 6-8 ilustran maneras alternativas de particionar una señal de audio de 11.1 canales en grupos de canales para la codificación paramétrica de la señal de audio de 11.1 canales como una señal de audio de 5.1 canales. La señal de audio de 11.1 canales comprende los canales L (izquierdo), LS (lateral izquierdo), LB (posterior izquierdo), TFL (izquierdo frontal superior), TBL (izquierdo posterior superior), R (derecho), RS (lateral derecho), RB (posterior derecho), TFR (derecho frontal superior), TBR (derecho posterior superior), C (central), 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 medio espacio izquierdo en un entorno de reproducción de la señal de audio de 11.1 canales. 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 verticalmente separadas de las de los tres canales L, LS y LB. Los dos canales TFL y TBL

pueden por ejemplo estar destinados a la reproducción en altavoces de techo. Similarmente, los cinco canales R, RS, RB, TFR y TBR forman una señal de audio de cinco canales adicional que representa un medio espacio derecho del entorno de reproducción, representando los tres canales R, RS y RB diferentes direcciones horizontales en el entorno de reproducción y representando los dos canales TFR y TBR direcciones verticalmente separadas de las de los tres canales R, RS y RB.

Para representar la señal de audio de 11.1 canales como una señal de audio de 5.1 canales, la colección de canales L, LS, LB, TFL, TBL, R, RS, RB, TFR, TBR, C y LFE se pueden particionar en grupos de canales representados por los respectivos canales de mezcla descendente y los parámetros de mezcla ascendente 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 parámetros de mezcla ascendente asociados, mientras que la señal de audio de cinco canales R, RS, RB, TFR, TBR adicional se puede representar por una señal de mezcla descendente de dos canales R_1, R_2 adicional y los parámetros de mezcla ascendente adicionales asociados. Los canales C y LFE se pueden mantener como canales separados también en la representación de 5.1 canales de la señal de audio de 11.1 canales.

La Figura 6 ilustra un primer formato de codificación F_1, en el que la señal de audio de cinco canales L, LS, LB, TFL, TBL se particiona en un primer grupo 601 de canales L, LS, LB y un segundo grupo 602 de canales TFL, TBL, y en el que la señal de audio de cinco canales R, RS, RB, TFR, TBR adicional se particiona en un primer grupo 603 de canales R, RS, RB adicional y un segundo grupo 604 de canales TFR, TBR adicional. En el primer formato de codificación F_1, el primer grupo de canales 601 se representa por un primer canal L_1 de la señal de mezcla descendente de dos canales, y el segundo grupo 602 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 601 de canales según $L_1 = L + LS + LB$, y el segundo canal L_2 de la señal de mezcla descendente puede corresponder a una suma del segundo grupo 602 de canales según $L_2 = TFL + TBL$.

En algunas realizaciones ejemplares, algunos o todos los canales se pueden reajustar antes de la suma, de modo que el primer canal L_1 de la señal de mezcla descendente puede corresponder a una combinación lineal del primer grupo 601 de canales según $L_1 = c_1L + c_2LS + c_3LB$, y el segundo anal L_2 de la señal de mezcla descendente puede corresponder a una combinación lineal del segundo grupo 602 de canales según $L_2= c_4TFL + c_5TBL$. Las ganancias c_2, c_3, c_4, c_5 pueden por ejemplo coincidir, mientras que la ganancia c_1 puede por ejemplo tener un valor diferente; p.ej., c_1 puede corresponder a no reajustar en absoluto. Por ejemplo, se pueden utilizar los valores $c_1 = 1$, y $c_2 = c_3 = c_4 = c_5 = 1/\sqrt{2}$. Si, por ejemplo, las ganancias c_1, ..., c_5 aplicadas a los respectivos canales L, LS, LB, TFL, TBL en el primer formato de codificación F_1 coinciden 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 Figuras 7 y 8, estas ganancias no afectan a cómo cambia la señal de mezcla descendente cuando se conmuta entre los diferentes formatos de codificación F_1, F_2, F_3, y los canales reajustados c_1L, c_2LS, C_3LB, c_4TFL, c_5TBL se pueden por lo tanto tratar como si fueran los canales originales L, LS, LB, TFL, TBL. Si, por otro lado, se emplean diferentes ganancias para el reajuste del mismo canal en diferentes formatos de codificación, la conmutación entre estos formatos de codificación puede por ejemplo provocar saltos entre versiones ajustadas de manera diferente de los canales L, LS, LB, TFL, TBL en la señal de mezcla descendente, lo que puede potencialmente provocar artefactos audibles en el lado de decodificador. Tales artefactos se pueden por ejemplo suprimir empleando la interpolación de coeficientes empleados para formar la señal de mezcla descendente antes de la conmutación de formato de codificación, a coeficientes empleados para formar la señal de mezcla descendente después de la conmutación de formato de codificación, y/o empleando la interpolación de coeficientes de descorrelación previa, como se describe a continuación en relación a las ecuaciones (3) y (4).

Similarmente, el primer grupo adicional de canales 603 se representa por un primer canal R_1 de la señal de mezcla descendente adicional, y el segundo grupo 604 de canales adicional se representa por un segundo canal R_2 de la señal de mezcla descendente adicional.

El primer formato de codificación F1 proporciona canales L_2 y R_2 de mezcla descendente dedicados para representar los canales TFL, TBL, TFR y TBR de techo. El uso del primer formato de codificación F_1 puede por lo tanto permitir la reconstrucción paramétrica de la señal de audio de 11.1 canales con fidelidad relativamente alta en casos donde, p.ej., una dimensión vertical en el entorno de reproducción es importante para la impresión global de la señal de audio de 11.1 canales.

La Figura 7 ilustra un segundo formato de codificación F_2, en el que la señal de audio de cinco canales L, LS, LB, TFL, TBL se particiona en primer 701 y segundo 702 grupos de canales representados por los respectivos canales L_1, L_2 de una señal de mezcla descendente, donde los canales L_1 y L_2 corresponden a sumas de los respectivos grupos 701 y 702 de canales, o combinaciones lineales de los respectivos grupos 701 y 702 de canales que emplean las mismas ganancias c_1, ..., c_5 para reajustar los respectivos canales L, LS, LB, TFL, TBL como en el primer formato de codificación F_1. Similarmente, la señal de audio de cinco canales R, RS, RB, TFR, TBR adicional se particiona en primer 703 y segundo 704 grupos de canales adicionales representados por los respectivos canales R_1 y R_2.

El segundo formato de codificación F_2 no proporciona canales de mezcla descendente dedicados para representar los canales TFL, TBL, TFR y TBR de techo pero puede permitir la reconstrucción paramétrica de la señal de audio de

11.1 canales con fidelidad relativamente alta p.ej., en casos donde la dimensión vertical en el entorno de reproducción no es tan importante para la impresión global de la señal de audio de 11.1 canales.

La Figura 8 ilustra un tercer formato de codificación F_3, en el que la señal de audio de cinco canales L, LS, LB, TFL, TBL se particiona en primer 801 y segundo 802 grupos de uno o más canales representados por respectivos canales L_1 y L_2 de una señal de mezcla descendente, donde la señal de los canales L_1 y L_2 corresponde a las sumas de los respectivos grupos 801 y 802 de uno o más canales, o combinaciones lineales de los respectivos grupos 801 y 802 de uno o más canales que emplean los mismos coeficientes c_1, ..., c_5 para reajustar los respectivos canales L, LS, LB, TFL, TBL que en el primer formato de codificación F_1. Similarmente, la señal de cinco canales R, RS, RB, TFR, TBR adicional se particiona en primer 803 y segundo 804 grupos de canales representados por respectivos canales R_1 y R_2. En el tercer formato de codificación F_3, solo el canal L se representa por el primer canal L_1 de la señal de mezcla descendente, mientras que los cuatro canales LS, LB, TFL y TBL se representan por el segundo canal L_2 de la señal de mezcla descendente.

En un lado de decodificador, que se describirá con referencia a las Figuras 1 – 5, se calcula una señal de mezcla descendente de dos canales L_1, L_2 como un mapeo lineal de la señal de audio de cinco canales X = [L LS LB TFL TBL]^T según

$$\begin{bmatrix} L_1 \\ L_2 \end{bmatrix} = \begin{bmatrix} d_{1,1} & \cdots & d_{1,5} \\ d_{2,1} & \cdots & d_{2,5} \end{bmatrix} \begin{bmatrix} L \\ LS \\ LB \\ TFL \\ TBL \end{bmatrix} = DX,$$

donde $d_{n,m}$, n=1, 2, m=1 ...,5 son coeficientes de mezcla descendente representados por una matriz D de mezcla descendente. En un lado de decodificador, que se describirá con referencia a las Figuras 9 – 13, la reconstrucción paramétrica de la señal de audio de cinco canales [L LS LB TFL TBL]^T se realiza según

$$\hat{X} = \begin{bmatrix} c_{1,1} & c_{1,2} \\ \vdots & \vdots \\ c_{5,1} & c_{5,2} \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} + \begin{bmatrix} p_{11} & p_{1,2} & p_{1,3} \\ \vdots & \vdots & \vdots \\ p_{5,1} & p_{5,2} & p_{5,3} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \beta_L \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} + \gamma_L Z, \qquad (2)$$

donde $c_{n,m}$, n=1, ...5, m = 1, 2 son coeficientes de mezcla ascendente secos representados por una matriz β_L de mezcla ascendente seca, $p_{n,k}$, n = 1, ..., 5, k = 1, 2, 3 son coeficientes de mezcla ascendente húmedos representados por una matriz γ_L de mezcla ascendente húmeda, y z_k, k = 1, 2, 3 son los canales de una señal Z descorrelacionada de tres canales generada en base a la señal de mezcla descendente L_1, L_2.

La Figura 1 es un diagrama de bloques generalizado de una sección 100 de codificación para codificar una señal de audio de M canales como una señal de mezcla descendente de dos canales y los parámetros de mezcla ascendente asociados, según una realización ejemplar.

La señal de audio de M canales se ejemplifica en la presenta memoria por la señal L de audio de cinco canales, LS, LB, TFL y TBL descrita con referencia a las Figuras 6 – 8. También se pueden contemplar realizaciones ejemplares en las que la sección 100 de codificación calcula una señal de mezcla descendente de dos canales en base a una señal de audio de M canales, donde M = 4, o M ≥ 6.

La sección 100 de codificación comprende una sección 110 de mezcla descendente y una sección 120 de análisis. Para cada uno de los formatos de codificación F_1, F_2, F_3, descritos con referencia a las Figuras 6 – 8, la sección 110 de mezcla descendente calcula, de acuerdo con el formato de codificación, una señal de mezcla descendente de dos canales L_1, L_2 en base a la señal de audio de cinco canales L, LS, LB, TFL, TBL. En por ejemplo el primer formato de codificación F_1, el primer canal L_1 de la señal de mezcla descendente se forma como una combinación lineal (p.ej., una suma) del primer grupo 601 de canales de la señal de audio de cinco canales L, LS, LB, TFL, TBL, y el segundo canal L_2 de la señal de mezcla descendente se forma como una combinación lineal (p.ej., una suma) del segundo grupo 602 de canales de la señal de audio de cinco canales L, LS, LB, TFL, TBL. La operación realizada por la sección 110 de mezcla descendente se puede por ejemplo expresar como la ecuación (1).

Para cada uno de los formatos de codificación F_1, F_2, F_3, la sección 120 de análisis determina un conjunto de coeficientes β_L de mezcla ascendente secos que definen un mapeo lineal de la respectiva señal de mezcla descendente L_1, L_2 que se aproxima a la señal de audio de cinco canales L, LS, LB, TFL, TBL, y calcula una diferencia

entre una covarianza de la señal de audio de cinco canales L, LS, LB, TFL, TBL recibida y una covarianza de la señal de audio de cinco canales aproximada por el respectivo mapeo lineal de la respectiva señal de mezcla descendente L_1, L_2. La diferencia calculada se ejemplifica en la presente memoria por una diferencia entre la matriz de covarianza de la señal de audio de cinco canales L, LS, LB, TFL, TBL recibida y la matriz de covarianza de la señal de audio de cinco canales aproximada por el respectivo mapeo lineal de la respectiva señal de mezcla descendente L_1, L_2. Para cada uno de los formatos de codificación F_1, F_2, F_3, la sección 120 de análisis determina un conjunto de coeficientes γ_L de mezcla ascendente húmedos, en base a la respectiva diferencia calculada, que junto con los coeficientes β_L de mezcla ascendente secos permite la reconstrucción paramétrica según la ecuación (2) de la señal de audio de cinco canales L, LS, LB, TFL, TBL a partir de la señal de mezcla descendente L_1, L_2 y a partir de una señal descorrelacionada de tres canales determinada en un lado de decodificador en base a la señal de mezcla descendente L_1, L_2. El conjunto de coeficientes γ_L de mezcla descendente húmedos define un mapeo lineal de la señal descorrelacionada de tal manera que la matriz de covarianza de la señal obtenida por el mapeo lineal de la señal descorrelacionada se aproxima a la diferencia entre la matriz de covarianza de la señal de audio de cinco canales L, LS, LB, TFL, TBL recibida y la matriz de covarianza de la señal de audio de cinco canales aproximada por el mapeo lineal de la señal de mezcla descendente L_1, L_2.

La sección 110 de mezcla descendente puede por ejemplo calcular la señal de mezcla descendente L_1, L_2 en el dominio del tiempo, i.e., en base a una representación de dominio del tiempo de la señal de audio de cinco canales L, LS, LB, TFL, TBL, o en un dominio de la frecuencia, i.e., en base a una representación de dominio de la frecuencia de la señal de audio de cinco canales L, LS, LB, TFL, TBL.

La sección 120 de análisis puede por ejemplo determinar los coeficientes β_L de mezcla ascendente secos y los coeficientes γ_L de mezcla ascendente húmedos en base a un análisis de dominio de la frecuencia de la señal de audio de cinco canales L, LS, LB, TFL, TBL. La sección 120 de análisis puede por ejemplo recibir la señal de mezcla descendente L_1, L_2 calculada por la sección 110 de mezcla descendente, o puede calcular su propia versión de la señal de mezcla descendente L_1, L_2, para determinar los coeficientes β_L de mezcla ascendente secos y los coeficientes γ_L de mezcla ascendente húmedos.

La Figura 3 es un diagrama de bloques generalizado de un sistema 300 de codificación de audio que comprende la sección 100 de codificación descrita con referencia a la Figura 1, según una realización ejemplar. En la presente realización ejemplar, el contenido de audio, p.ej., grabado por uno o más transductores acústicos 301, o generado por un equipo 301 de creación de audio, se proporciona en forma de la señal de audio de 11.1 canales descrita con referencia a las Figuras 6 – 8. Una sección 302 (o banco de filtros) de análisis de filtro espejo en cuadratura (QMF) transforma la señal de audio de cinco canales L, LS, LB, TFL, TBL, segmento de tiempo a segmento de tiempo, a un dominio de QMF para su procesamiento por la sección 100 de codificación de la señal de audio de cinco canales L, LS, LB, TFL, TBL en forma de baldosas de tiempo/frecuencia. (Como se explicará adicionalmente a continuación, la sección 302 de análisis de QMF y su contraparte, una sección 305 de síntesis de QMF, son opcionales). El sistema 300 de codificación de audio comprende una sección 303 de codificación adicional análoga a la sección 100 de codificación y adaptada para codificar la señal de audio de cinco canales R, RS, RB, TRF y TBR adicional como la señal de mezcla descendente de dos canales R_1, R_2 adicional y los parámetros β_R de mezcla ascendente secos adicionales y los parámetros γ_R de mezcla ascendente húmedos adicionales. La sección 302 de análisis de QMF también transforma la señal de audio de cinco canales R, RS, RB, TRF y TBR adicional a un dominio de QMF para su procesamiento por la sección 303 de codificación adicional.

Una sección 304 de control selecciona uno de los formatos de codificación F_1, F_2, F_3 en base a los coeficientes γ_L, γ_R y β_L, β_R de mezcla ascendente húmedos y secos determinados por la sección 100 de codificación y la sección 303 de codificación adicional para los respectivos formatos de codificación F_1, F_2, F_3. Por ejemplo, para cada uno de los formatos de codificación F_1, F_2, F_3 , la sección 304 de control puede calcular una relación

$$E = \frac{E_{wet}}{E_{wet} + E_{dry}},$$

donde E_{wet} es una suma de cuadrados de los coeficientes γ_L y γ_R de mezcla ascendente húmedos, y E_{dry} es una suma de cuadrados de los coeficientes β_L, β_R de mezcla ascendente secos. El formato de codificación seleccionado se puede asociar con la mínima de las relaciones E de los formatos de codificación F_1, F_2, F_3, i.e., la sección 304 de control puede seleccionar el formato de codificación correspondiente a la relación E más pequeña. Los inventores se han dado cuenta de que un valor reducido para la relación E puede ser indicativo de una mayor fidelidad de la señal de audio de 11.1 canales reconstruida a partir del formato de codificación asociado.

En algunas realizaciones ejemplares, la suma de cuadrados E_{dry} de los coeficientes β_L, β_R de mezcla ascendente secos puede por ejemplo incluir un término adicional con el valor 1, correspondiente al hecho de que el canal C se

transmite al lado de decodificador y se puede reconstruir sin ninguna descorrelación , p.ej., empleando solo un coeficiente de mezcla ascendente seco con el valor 1.

En algunas realizaciones ejemplares, la sección 304 de control puede seleccionar formatos de codificación para las dos señales de audio de cinco canales L, LS, LB, TFL, TBL y R, RS, RB, TFR, TBR independientemente una de otra, en base a los coeficientes γ_L , β_L de mezcla ascendente húmedos y secos y los coeficientes γ_R , β_R de mezcla ascendente húmedos y secos adicionales, respectivamente.

El sistema 300 de codificación de audio puede entonces emitir la señal de mezcla descendente L_1, L_2, y la señal de señal mezcla descendente R_1, R_2 adicional, del formato de codificación seleccionado, los parámetros α de mezcla ascendente a partir de los cuales los coeficientes β_L, γ_L de mezcla ascendente secos y húmedos y los coeficientes β_R, γ_R de mezcla ascendente secos y húmedos adicionales asociados con el formato de codificación seleccionado, pueden ser derivados, y la señalización S que indica el formato de codificación seleccionado.

En la presente realización ejemplar, la sección 304 de control emite la señal de mezcla descendente L_1, L_2, y la señal de mezcla descendente R_1, R_2 adicional del formato de codificación seleccionado, los parámetros α de mezcla ascendente a partir de los cuales los coeficientes β_L, γ_L de mezcla ascendente secos y húmedos y los coeficientes β_R, γ_R de mezcla ascendente secos y húmedos adicionales, asociados con el formato de codificación seleccionado, pueden ser derivados, y la señalización S que indica el formato de codificación seleccionado. La señal de mezcla descendente L_1, L_2 y la señal de mezcla descendente R_1, R_2 adicional son transformadas de vuelta desde el dominio de QMF por una sección 305 (o banco de filtros) de síntesis de QMF y son transformadas a un dominio de la transformada discreta del coseno modificada (MDCT) por una sección 306 de transformación. Una sección 307 de cuantización cuantifica los parámetros α de mezcla ascendente. Por ejemplo, se puede emplear cuantificación uniforme con un tamaño de paso de 0,1 o 0,2 (adimensional), seguida de codificación entrópica en forma de codificación de Huffman. Se puede por ejemplo emplear una cuantificación más gruesa con tamaño de paso de 0,2 para ahorrar ancho de banda de transmisión, y se puede por ejemplo emplear una cuantificación más fina con tamaño de paso de 0,1 para mejorar la fidelidad de la reconstrucción en un lado de decodificador. Los canales C y LFE son transformados también a un dominio de MDCT por una sección 308 de transformación. Las señales y canales de mezcla descendente transformados mediante MDCT, los parámetros de mezcla ascendente cuantificados, y la señalización, son combinados entonces en un flujo de bits B por un multiplexor 309, para su transmisión a un lado de decodificador. El sistema 300 de codificación de audio también puede comprender un codificador central (no mostrado en la Figura 3) 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 utilizando un códec de audio perceptual, tal como Dolby Digital, MPEG AAC o un desarrollo de los mismos, antes de que las señales de mezcla descendente y los canales C y LFE sean proporcionados al demultiplexor 309. Se puede por ejemplo aplicar una ganancia de clip, p.ej., correspondiente a -8,7 dB, 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. Alternativamente, dado que los parámetros son independientes del nivel absoluto, las ganancias de clip también se pueden aplicar a los canales de entrada antes de formar las combinaciones lineales correspondientes a L_1, L_2.

También se pueden contemplar realizaciones en las que la sección 304 de control solo recibe los coeficientes γ_L, γ_R , β_L, β_R de mezcla ascendente húmedos y secos para los diferentes formatos de codificación F_1, F_2, F_3 (o sumas de cuadrados de los coeficientes de mezcla ascendente húmedos y secos para los diferentes formatos de codificación) para seleccionar un formato de codificación, i.e., la sección 304 de control no necesita necesariamente recibir las señales de mezcla descendente L_1, L_2, R_1, R_2 para los diferentes formatos de codificación. En tales realizaciones, la sección 304 de control puede por ejemplo controlar las secciones 100, 303 de codificación para entregar las señales de mezcla descendente L_1, L_2, R_1, R_2, los coeficientes β_L, β_R de mezcla ascendente secos y los coeficientes γ_L, γ_R de mezcla ascendente húmedos para el formato de codificación seleccionado como salida del sistema 300 de codificación de audio, o como entrada al multiplexor 309.

Si el formato de codificación seleccionado conmuta entre los formatos de codificación, entonces se puede por ejemplo realizar interpolación entre los valores de los coeficientes de mezcla descendente empleados antes y después de la conmutación de formato de codificación para formar la señal de mezcla descendente de acuerdo con la ecuación (1). Esto es en general equivalente a una interpolación de las señales de mezcla descendente producidas de acuerdo con los respectivos conjuntos de valores de los coeficientes de mezcla descendente.

Mientras que la Figura 3 ilustra cómo la señal de mezcla descendente puede ser generada en el dominio de QMF y luego posteriormente transformada de vuelta al dominio del tiempo, se puede implementar un codificador alternativo que cumple las mismas funciones sin las secciones 302, 305 de QMF, por lo que calcula la señal de mezcla descendente directamente en el dominio del tiempo. Esto es posible en situaciones donde los coeficientes de mezcla descendente no son dependientes de la frecuencia, lo que en general es cierto. Con el codificador alternativo, se pueden manejar las transiciones de formato de codificación mediante desvanecimiento cruzado entre las dos señales de mezcla descendente para los respectivos formatos de codificación o mediante interpolación entre los coeficientes de mezcla descendente (incluyendo los coeficientes que son de valor cero en uno de los formatos) que producen las señales de mezcla descendente. Tal codificador alternativo puede tener menor retardo/latencia y/o menor complejidad de cálculo.

La Figura 2 es un diagrama de bloques generalizado de una sección 200 de codificación similar a la sección 100 de codificación, descrita con referencia a la Figura 1, según una realización ejemplar. La sección 200 de codificación comprende una sección 210 de mezcla descendente y una sección 220 de análisis. Como en la sección 100 de decodificación, descrita con referencia a la Figura 1, la sección 210 de mezcla descendente calcula una señal de mezcla descendente de dos canales L_1, L_2 en base a la señal de audio de cinco canales L, LS, LB, TFL, TBL para cada uno de los formatos de codificación F_1, F_2, F_3, y la sección 220 de análisis determina respectivos conjuntos de coeficientes β_L de mezcla ascendente secos, y calcula las diferencias Δ_L entre una matriz de covarianza de la señal de audio de cinco canales L, LS, LB, TFL, TBL recibida y las matrices de covarianza de la señal de audio de cinco canales aproximadas por los respectivos mapeos lineales de las respectivas señales de mezcla descendente.

En contraste con la sección 120 de análisis en la sección 100 de codificación, descrita con referencia a la Figura 1, la sección 220 de análisis no calcula los parámetros de mezcla ascendente húmedos para todos los formatos de codificación. En su lugar, las diferencias Δ_L calculadas se proporcionan a la sección 304 de control (véase la Figura 3) para la selección de un formato de codificación. Una vez que se ha seleccionado un formato de codificación en base a las diferencias Δ_L calculadas, los coeficientes de mezcla ascendente húmedos (a ser incluidos en un conjunto de parámetros de mezcla ascendente) para el formato de codificación seleccionado pueden entonces ser determinados por la sección 304 de control. Alternativamente, la sección 304 de control es responsable de seleccionar el formato de codificación en base a las diferencias Δ_L calculadas entre las matrices de covarianza discutidas anteriormente, pero ordena a la sección 220 de análisis, a través de señalización en la dirección aguas arriba, que calcule los coeficientes γ_L de mezcla ascendente húmedos; según esta alternativa (no mostrada), la sección 220 de análisis tiene la capacidad de emitir tanto las diferencias como los coeficientes de mezcla ascendente húmedos.

En la presente realización ejemplar, el conjunto de coeficientes de mezcla ascendente húmedos se determina de tal manera que una matriz de covarianza de una señal obtenida por un mapeo lineal de la señal descorrelacionada, definida por los coeficientes de mezcla ascendente húmedos, suplementa una matriz de covarianza de la señal de audio de cinco canales aproximada por el mapeo lineal de la señal de mezcla descendente del formato de codificación seleccionado. En otras palabras, los parámetros de mezcla ascendente húmedos no necesitan necesariamente ser determinados para lograr la reconstrucción completa de la covarianza cuando al reconstruir la señal de audio de cinco canales L, LS, LB, TFL, TBL en un lado de decodificador. Se pueden determinar los parámetros de mezcla ascendente húmedos para mejorar la fidelidad de la señal de audio de cinco canales reconstruida, pero, si por ejemplo el número de descorreladores en el lado de decodificador es limitado, se pueden determinar los parámetros de mezcla ascendente húmedos para permitir la reconstrucción tanto como sea posible de la matriz de covarianza de la señal de audio de cinco canales L, LS, LB, TFL, TBL.

Se pueden contemplar realizaciones, en las que los sistemas de codificación de audio similares al sistema 300 de codificación de audio, descrito con referencia a la Figura 3, comprenden una o más secciones 200 de codificación del tipo descrito con referencia a la Figura 2.

La Figura 4 es un diagrama de flujo de un método 400 de codificación de audio para codificar una señal de audio de M canales como una señal de mezcla descendente de dos canales y los parámetros de mezcla ascendente asociados, según una realización ejemplar. El método 400 de codificación de audio se ejemplifica en la presente memoria mediante un método realizado por un sistema de codificación de audio que comprende la sección 200 de codificación, descrita con referencia a la Figura 2.

El método 400 de codificación de audio comprende: recibir 410 la señal de audio de cinco canales L, LS, LB, TFL, TBL; calcular 420, de acuerdo con un primero de los formatos de codificación F_1, F_2, F_3 descritos con referencia a las Figuras 6 – 8, la señal de mezcla descendente de dos canales L_1, L_2 en base a la señal de audio de cinco canales L, LS, LB, TFL, TBL; determinar 430 el conjunto de coeficientes β_L de mezcla ascendente secos de acuerdo con el formato de codificación; y calcular 440 la diferencia Δ_L de acuerdo con el formato de codificación. El método 400 de codificación de audio comprende: determinar 450 si se han calculado las diferencias Δ_L para cada uno de los formatos de codificación F_1, F_2, F_3. Siempre que quede por calcular una diferencia Δ_L para al menos un formato de codificación, el método 400 de codificación de audio vuelve a calcular 420 la señal de mezcla descendente L_1, L_2 de acuerdo con el siguiente formato de codificación en la línea, lo que está indicado por N en el diagrama de flujo.

Si se han calculado las diferencias Δ_L para cada uno de los formatos de codificación F_1, F_2, F_3 indicados por Y en el diagrama de flujo, el método 400 procede seleccionando 460 uno de los formatos de codificación F_1, F_2, F_3, en base a las respectivas diferencias Δ_L calculadas; y determinando 470 el conjunto de coeficientes de mezcla ascendente húmedos, que junto con los coeficientes β_L de mezcla ascendente secos del formato de codificación seleccionado permite la reconstrucción paramétrica de la señal de audio de cinco canales L, LS, LB, TFL, TBLM según la ecuación (2). El método 400 de codificación de audio comprende además: emitir 480 la señal de mezcla descendente L_1, L_2 del formato de codificación seleccionado, y los parámetros de mezcla ascendente a partir de los cuales se pueden derivar los coeficientes de mezcla ascendente secos y húmedos asociados con el formato de codificación seleccionado; y emitir 490 la señalización S que indica el formato de codificación seleccionado.

La Figura 5 es un diagrama de flujo de un método 500 de codificación de audio para codificar una señal de audio de M canales como una señal de mezcla descendente de dos canales y los parámetros de mezcla ascendente asociados,

según una realización ejemplar. El método 500 de codificación de audio se ejemplifica en la presente memoria mediante un método realizado por el sistema 300 de codificación de audio, descrito con referencia a la Figura 3.

Similarmente al método 400 de codificación de audio descrito con referencia a la Figura 4, el método 500 de codificación de audio comprende: recibir 410 la señal de audio de cinco canales L, LS, LB, TFL, TBL; calcular 420, de acuerdo con un primero de los formatos de codificación F_1, F_2, F_3, la señal de mezcla descendente de dos canales L_1, L_2 en base a la señal de audio de cinco canales L, LS, LB, TFL, TBL; determinar 430 el conjunto de coeficientes β_L de mezcla ascendente secos de acuerdo con el formato de codificación; y calcular 440 la diferencia Δ_L de acuerdo con el formato de codificación. El método 500 de codificación de audio comprende además determinar 560 el conjunto de coeficientes γ_L de mezcla ascendente húmedos que junto con los coeficientes β_L de mezcla ascendente secos del formato de codificación permite la reconstrucción paramétrica de la señal de audio de M canales de acuerdo con la ecuación (2). El método 500 de codificación de audio comprende: determinar 550 si se han calculado los coeficientes γ_L, β_L de mezcla ascendente húmedos y secos para cada uno de los formatos de codificación F_1, F_2, F_3. Siempre que queden por calcular coeficientes γ_L, β_L de mezcla ascendente húmedos y secos para al menos un formado de codificación, el método de método 500 de codificación de audio método vuelve a calcular 420 la señal de mezcla descendente L_1, L_2 de acuerdo con el siguiente formato de codificación en la línea, lo que está indicado por N en el diagrama de flujo.

Si se han calculado los coeficientes γ_L, β_L de mezcla ascendente húmedos y secos para cada uno de los formatos de codificación F_1, F_2, F_3 indicados por Y en el diagrama de flujo, el método 500 de codificación de audio procede seleccionando 570 uno de los formatos de codificación F_1, F_2, F_3, en base a los respectivos coeficientes γ_L, β_L de mezcla ascendente húmedos y secos calculados; emitiendo 480 la señal de mezcla descendente L_1, L_2 del formato de codificación seleccionado, y los parámetros de mezcla ascendente a partir de los cuales se pueden derivar los coeficientes β_L, γ_L de mezcla ascendente secos y húmedos asociados con el formato de codificación seleccionado; y emitir 490 señalización que indica el formato de codificación seleccionado.

La Figura 9 es un diagrama de bloques generalizado de una sección 900 de decodificación para reconstruir una señal de audio de M canales en base a una señal de mezcla descendente de dos canales y los parámetros α_L de mezcla ascendente asociados, según una realización ejemplar.

En la presente realización ejemplar, la señal de mezcla descendente se ejemplifica mediante la señal de mezcla descendente L_1, L_2 emitida por la sección 100 de codificación, descrita con referencia a la Figura 1. En la presente realización ejemplar, los parámetros β_L, γ_L, de mezcla ascendente secos y húmedos emitidos por la sección 100 de codificación, y que se adaptan para la reconstrucción paramétrica de la señal de audio de cinco canales L, LS, LB, TFL, TBL se pueden derivar a partir de los parámetros α_L de mezcla ascendente. Sin embargo, también se pueden contemplar realizaciones en las que los parámetros α_L de mezcla ascendente se adaptan para la reconstrucción paramétrica de una señal de audio de M canales, donde M = 4, o M ≥ 6.

La sección 900 de decodificación comprende una sección 910 de descorrelación previa, una sección 920 de descorrelación y una sección 930 de mezcla. La sección 910 de descorrelación previa determina un conjunto de coeficientes de descorrelación previa en base a un formato de codificación seleccionado empleado en un lado de codificador para codificar la señal de audio de cinco canales L, LS, LB, TFL, TBL. Como se describe a continuación con referencia a la Figura 10, el formato de codificación seleccionado se puede indicar a través de señalización desde el lado de codificador. La sección 910 de descorrelación previa calcula una señal de entrada de descorrelación D_1, D_2, D_3 como un mapeo lineal de la señal de mezcla descendente L_1, L_2, donde se aplica el conjunto de coeficientes de descorrelación previa a la señal de mezcla descendente L_1, L_2.

La sección 920 de descorrelación genera una señal descorrelacionada en base a la señal de entrada de descorrelación D_1, D_2, D_3. La señal descorrelacionada se ejemplifica en la presente memoria mediante tres canales, cada uno generado procesando uno de los canales de la señal de entrada de descorrelación en un descorrelador 921 – 923 de la sección 920 de descorrelación, p.ej., incluyendo aplicar filtros lineales a los respectivos canales de la señal de entrada de descorrelación D_1, D_2, D_3.

La sección 930 de mezcla determina los conjuntos de coeficientes β_L, γ_L de mezcla ascendente húmedos y secos en base a los parámetros α_L de mezcla ascendente recibidos y el formato de codificación seleccionado empleado en un lado de codificador para codificar la señal de audio de cinco canales L, LS, LB, TFL, TBL. La sección 930 de mezcla realiza reconstrucción paramétrica de la señal de audio de cinco canales L, LS, LB, TFL, TBL de acuerdo con la ecuación (2), i.e., calcula una señal de mezcla ascendente seca como un mapeo lineal de la señal de mezcla descendente L_1, L_2, en donde el conjunto de coeficientes β_L de mezcla ascendente secos se aplica a la señal de mezcla descendente L_1, L_2; calcula una señal de mezcla ascendente húmeda como un mapeo lineal de la señal descorrelacionada, donde el conjunto de coeficientes γ_L de mezcla ascendente húmedos se aplica a la señal descorrelacionada; y combina las señales de mezcla ascendente secas y húmedas para obtener una señal reconstruida multidimensional $\widetilde{L}$, $\widetilde{LS}$, $\widetilde{LB}$, $\widetilde{TFL}$, $\widetilde{TBL}$ correspondiente a la señal de audio de cinco canales L, LS, LB, TFL, TBL a reconstruir.

En algunas realizaciones ejemplares, los parámetros α_L de mezcla ascendente recibidos pueden incluir los propios coeficientes β_L, γ_L de mezcla ascendente húmedos y secos, o pueden corresponder a una forma más compacta,

incluyendo menos parámetros que el número de coeficientes β_L, γ_L de mezcla ascendente húmedos y secos, a partir de los cuales se pueden derivar los coeficientes β_L, γ_L de mezcla ascendente húmedos y secos en el lado de decodificador en base al conocimiento de la forma compacta particular empleada.

La Figura 11 ilustra la operación de la sección 930 de mezcla, descrita con referencia a la Figura 9, en un escenario ejemplar donde la señal de mezcla descendente L_1, L_2 representa la señal de audio de cinco canales L, LS, LB, TFL, TBL de acuerdo con el primer formato de codificación F_1, descrito con referencia a la Figura 6. Se apreciará que la operación de la sección 930 de mezcla puede ser similar en escenarios ejemplares donde la señal de mezcla descendente L_1, L_2 representa la señal de audio de cinco canales L, LS, LB, TFL, TBL de acuerdo con cualquiera del segundo y tercer formatos de codificación F_2, F_3. En particular, la sección 930 de mezcla puede activar temporalmente otras instancias de las secciones de mezcla ascendente y combinar secciones que se describirán de forma inminente, para permitir un desvanecimiento cruzado entre dos formatos de codificación, lo que puede requerir disponibilidad simultánea de las señales de mezcla descendente calculadas.

En el presente escenario ejemplar, el primer canal L_1 de la señal de mezcla descendente representa los tres canales L, LS, LB, y el segundo canal L_2 de la señal de mezcla descendente representa los dos canales TFL, TBL. La sección 910 de descorrelación previa determina los coeficientes de descorrelación previos de tal manera que se generan dos canales de la señal descorrelacionada en base al primer canal L_1 de la señal de mezcla descendente y de tal manera que se genera un canal de la señal descorrelacionada en base al segundo canal L_2 de la señal de mezcla descendente.

Una primera sección 931 de mezcla ascendente seca proporciona una señal X_1 de mezcla ascendente seca de tres canales como un mapeo lineal del primer canal L_1 de la señal de mezcla descendente, donde un subconjunto de los coeficientes de mezcla ascendente secos, que se pueden derivar a partir de los parámetros α_L de mezcla ascendente recibidos, se aplica al primer canal L_1 de la señal de mezcla descendente. Una primera sección 932 de mezcla ascendente húmeda proporciona una señal Y_1 de mezcla ascendente húmeda de tres canales como un mapeo lineal de los dos canales de la señal descorrelacionada, donde un subconjunto de los coeficientes de mezcla ascendente húmedos, que se pueden derivar de los parámetros α_L de mezcla ascendente recibidos, se aplica a los dos canales de la señal descorrelacionada. Una primera sección 933 de combinación combina la primera señal X_1 de mezcla ascendente seca y la primera señal Y_1 de mezcla ascendente húmeda en versiones reconstruidas $\widetilde{L}$, $\widetilde{LS}$, $\widetilde{LB}$ de los canales L, LS, LB.

Similarmente, una segunda sección 934 de mezcla ascendente seca proporciona una señal X_2 de mezcla ascendente seca de dos canales como un mapeo lineal del segundo canal L_2 de la señal de mezcla descendente, y una segunda sección 935 de mezcla ascendente húmeda proporciona una señal Y_2 de mezcla ascendente húmeda de dos canales como una combinación lineal de un canal de la señal descorrelacionada. Una segunda sección 936 de combinación combina la segunda señal X_2 de mezcla ascendente seca y la segunda señal Y_2 de mezcla ascendente húmeda en versiones reconstruidas $\widetilde{TFL}$, $\widetilde{TBL}$ de los canales TFL, TBL.

La Figura 10 es un diagrama de bloques generalizado de un sistema 1000 de decodificación de audio que comprende la sección 900 de decodificación, descrita con referencia a la Figura 9, según una realización ejemplar. Una sección 1001 de recepción, p.ej., que incluye un multiplexor, recibe el flujo de bits B transmitido desde el sistema 300 de codificación de audio, descrito con referencia a la Figura 3, y extrae la señal de mezcla descendente L_1, L_2, la señal de mezcla descendente R_1, R_2 adicional, y los parámetros α de mezcla ascendente, así como los canales C y LFE, del flujo de bits B. Los parámetros α de mezcla ascendente pueden por ejemplo comprender primer y segundo subconjuntos α_L y α_R, asociados con el lado izquierdo y el lado derecho, respectivamente, de la señal de audio de 11.1 canales L, LS, LB, TFL, TBL, R, RS, RB, TFR, TBR, C, LFE a reconstruir.

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 utilizando un códec de audio perceptual tal como Dolby Digital, MPEG AAC, o desarrollos de los mismos, el sistema 1000 de decodificación de audio puede comprender un decodificador central (no mostrado en la Figura 10) configurado para decodificar las respectivas señales y canales cuando se extraen del flujo de bits B.

Una sección 1002 de transformación transforma la señal de mezcla descendente L_1, L_2 realizando MDCT inversa y una sección 1003 de análisis de QMF transforma la señal de mezcla descendente L_1, L_2 a un dominio de QMF para su procesamiento por la sección 900 de decodificación de la señal de mezcla descendente L_1, L_2 en forma de baldosas de tiempo/frecuencia. Una sección 1004 de descuantificación descuantifica el primer subconjunto de parámetros α_L de mezcla ascendente, p.ej., a partir de un formato codificado de entropía, antes de suministrarlo a la sección 900 de decodificación. Como se describió con referencia a la Figura 3, la cuantificación se puede haber realizado con uno de dos tamaños de paso diferentes, p.ej., 0,1 o 0,2. El tamaño de paso real empleado se puede predefinir, o se puede señalizar al sistema 1000 de decodificación de audio desde el lado de codificador, p.ej., a través del flujo de bits B.

En la presente realización ejemplar, el sistema 1000 de decodificación de audio comprende una sección 1005 de decodificación adicional análoga a la sección 900 de decodificación. La sección 1005 de decodificación adicional se configura para recibir la señal de mezcla descendente de dos canales R_1, R_2 adicional descrita con referencia a la Figura 3, y el segundo subconjunto α_R de parámetros de mezcla ascendente, y para proporcionar una versión

reconstruida $\widetilde{R}$, $\widetilde{RS}$, $\widetilde{RB}$, $\widetilde{TFR}$, $\widetilde{TBR}$ de la señal de audio de cinco canales R, RS, RB, TFR, TBR adicional en base a la señal de mezcla descendente R_1, R_2 adicional y el segundo subconjunto α_R de parámetros de mezcla ascendente.

Una sección 1006 de transformación transforma la señal de mezcla descendente R_1, R_2 adicional realizando MDCT inversa y una sección 1007 de análisis de QMF transforma la señal de mezcla descendente R_1, R_2 adicional a un dominio de QMF para su procesamiento por la sección 1005 de decodificación adicional de la señal de mezcla descendente R_1, R_2 adicional en forma de baldosas de tiempo/frecuencia. Una sección 1008 de descuantificación descuantifica el segundo subconjunto de parámetros α_R de mezcla ascendente, p.ej., a partir de un formato codificado de entropía, antes de suministrarlos a la sección 1005 de decodificación adicional.

En realizaciones ejemplares donde se ha aplicado una ganancia de clip 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 un lado de codificador, se puede aplicar una ganancia correspondiente, p.ej., correspondiente a 8,7 dB, a estas señales en el sistema 1000 de decodificación de audio para compensar la ganancia de clip.

Una sección 1009 de control recibe la señalización S que indica uno seleccionado de los formatos de codificación F_1, F_2, F_3 empleados en el lado de codificador para codificar la señal de audio de 11.1 canales en la señal de mezcla descendente L_1, L_2 y la señal de mezcla descendente R_1, R_2 adicional y los parámetros α de mezcla ascendente asociados. La sección 1009 de control controla la sección 900 de decodificación (p.ej., la sección 910 de descorrelación previa y la sección 920 de mezcla en la misma) y la sección (1005) de decodificación adicional para realizar la reconstrucción paramétrica de acuerdo con el formato de codificación indicado.

En la presente realización ejemplar, las versiones reconstruidas de la señal de audio de cinco canales L, LS, LB, TFL, TBL y la señal de audio de cinco canales R, RS, RB, TFL, TBL adicional emitidas por la sección 900 de decodificación y la sección 1005 de decodificación adicional, respectivamente, son transformadas de vuelta desde el dominio de QMF por una sección 1011 de síntesis de QMF antes de ser proporcionadas junto con los canales C y LFE como salida del sistema 1000 de decodificación de audio para la reproducción en el sistema 1012 de altavoces múltiples. Una sección 1010 de transformación transforma los canales C y LFE al dominio del tiempo realizando MDCT inversa antes de que estos canales sean incluidos en la salida del sistema 1000 de decodificación de audio.

Los canales C y LFE se pueden por ejemplo extraer del flujo de bits B en una forma codificada de manera discreta y el sistema 1000 de decodificación de audio puede por ejemplo comprender secciones de decodificación de un solo canal (no mostradas en la Figura 10) configuradas para decodificar los respectivos canales codificados de manera discreta. Las secciones de decodificación de un solo canal pueden por ejemplo incluir decodificadores centrales para decodificar el contenido de audio codificado utilizando un códec de audio perceptual tal como Dolby Digital, MPEG AAC, o desarrollos de los mismos.

En la presente realización ejemplar, los coeficientes de descorrelación previa son determinados por la sección 910 de descorrelación previa de tal manera que, en cada uno de los formatos de codificación F_1, F_2, F_3, cada uno de los canales de la señal de entrada de descorrelación D_1, D_2, D_3 coincide con un canal de la señal de mezcla descendente L_1, L_2, de acuerdo con la Tabla 1.

Tabla 1

Canal de señal de entrada de descorrelación	Formato de codificación F_1	Formato de codificación F_2	Formato de codificación F_3
D1	L_1 = L + LS + LB	L_1 = L + TFL	L_2 = LS + LB + TFL + TBL
D2	L_1 = L + LS + LB	L_2 = LS + LB + TBL	L_2 = LS + LB + TFL + TBL
D3	L_2 = TFL + TBL	L_2 = LS + LB + TBL	L_2 = LS + LB + TFL + TBL

Como se puede ver en la Tabla 1, el canal TBL contribuye, a través de la señal de mezcla descendente L_1, L_2, a un tercer canal D_3 de la señal de entrada de descorrelación en los tres formatos de codificación F_1, F_2, F_3, mientras que cada uno de los pares de canales LS, LB y TFL, TBL contribuye, a través de la señal de mezcla descendente L_1, L_2, al tercer canal D3 de la señal de entrada de descorrelación en al menos dos de los formatos de codificación, respectivamente.

La Tabla 1 muestra que cada uno de los canales L y TFL contribuye, a través de la señal de mezcla descendente L_1, L_2, a un primer canal D1 de la señal de entrada de descorrelación en dos de los formatos de codificación,

respectivamente, y el par de canales LS, LB contribuye, a través de la señal de mezcla descendente L_1, L_2, al primer canal D1 de la señal de entrada de descorrelación en al menos dos de los formatos de codificación.

La tabla 1 también muestra que los tres canales LS, LB, TBL contribuyen, a través de la señal de mezcla descendente L_1, L_2, a un segundo canal D2 de la señal de entrada de descorrelación tanto en el segundo como en el tercer formatos de codificación F_3, F_3, mientras que el par de canales LS, LB contribuye, a través de la señal de mezcla descendente L_1, L_2, al segundo canal D2 de la señal de entrada de descorrelación en los tres formatos de codificación F_1, F_2, F_3.

Cuando el formato de codificación indicado conmuta entre diferentes formatos de codificación, la entrada a los descorreladores 921 – 923 cambia. En la presente realización ejemplar, al menos algunas porciones de las señales de entrada de descorrelación D1, D2, D3 permanecerán durante la conmutación, i.e., al menos un canal de la señal de audio de cinco canales L, LS, LB, TFL, TBL permanecerá en cada canal de la señal de entrada de descorrelación D1, D2, D3 en cualquier conmutación entre dos de los formatos de codificación F_1, F_2, F_3, lo que permite una transición más suave entre los formatos de codificación, tal como es percibido por un oyente durante la reproducción de la señal de audio de M canales reconstruida.

Los inventores se han dado cuenta de que dado que la señal descorrelacionada se puede generar en base a una sección de la señal de mezcla descendente L_1, L_2 correspondiente a varias tramas de tiempo, durante las que se puede producir una conmutación de formato de codificación, se pueden potencialmente generar artefactos audibles en la señal descorrelacionada como resultado de la conmutación de formatos de codificación. Incluso si se interpolan los coeficientes β_L, γ_L de mezcla ascendente húmedos y secos en respuesta a una transición entre formatos de codificación, los artefactos provocados en la señal descorrelacionada pueden todavía persistir en la señal de audio de cinco canales L, LS, LB, TFL, TBL reconstruida. Proporcionar la señal de entrada de descorrelación D1, D2, D3 de acuerdo con la Tabla 1 puede suprimir los artefactos audibles en la señal descorrelacionada provocados por la conmutación de formato de codificación, y puede mejorar la calidad de reproducción de la señal de audio de cinco canales L, LS, LB, TFL, TBL reconstruida.

Aunque la Tabla 1 se expresa en términos de formatos de codificación F_1, F_2, F_3 para los que los canales de la señal de mezcla descendente L_1, L_2 son generados como sumas del primer y segundo grupos de canales, respectivamente, se pueden por ejemplo emplear los mismos valores para los coeficientes de descorrelación previa cuando los canales de la señal de mezcla descendente han sido formados como combinaciones lineales del primer y segundo grupos de canales, respectivamente, de tal manera que los canales de la señal de entrada de descorrelación D1, D2, D3 coinciden con los canales de la señal de mezcla descendente L_1, L_2 de acuerdo con la Tabla 1. Se apreciará que la calidad de reproducción de la señal de audio de cinco canales reconstruida se puede mejorar de esta manera también en cuando los canales de la señal de mezcla descendente se forman como combinaciones lineales del primer y segundo grupos de canales, respectivamente.

Para mejorar adicionalmente la calidad de reproducción de la señal de audio de cinco canales reconstruida, se puede por ejemplo realizar interpolación de valores de los coeficientes de descorrelación previa en respuesta a la conmutación del formato de codificación. En el primer formato de codificación F_1, la señal de entrada de descorrelación D1, D2, D3 se puede determinar como

$$\begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \end{bmatrix},$$

mientras que en el segundo formato de codificación F_2, la señal de entrada de descorrelación D1, D2, D3 se puede determinar como

$$\begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \end{bmatrix}.$$

En respuesta a una conmutación del primer formato de codificación F_1 al segundo formato de codificación F_2, se puede por ejemplo realizar interpolación continua o lineal entre la matriz de descorrelación previa en la ecuación (3) y la matriz de descorrelación previa en la ecuación (4).

La señal de mezcla descendente L_1, L_2, en las ecuaciones (3) y (4) puede por ejemplo estar en el dominio de QMF, y cuando se conmuta entre formatos de codificación, los coeficientes de mezcla descendente empleados en un lado de codificador para calcular la señal de mezcla descendente L_1, L_2 según la ecuación (1) se pueden haber interpolado durante p.ej., intervalos de 32 QMF. La interpolación de los coeficientes (o matrices) de descorrelación previa se puede por ejemplo sincronizar con la interpolación de los coeficientes de mezcla descendente, p.ej., se puede realizar durante los mismos intervalos de 32 QMF. La interpolación de los coeficientes de descorrelación previa puede ser por ejemplo

una interpolación de banda ancha, p.ej., empleada para todas las bandas de frecuencia decodificadas por el sistema 1000 de decodificación de audio.

Los coeficientes β_L, γ_L de mezcla ascendente secos y húmedos también se pueden interpolar. Las interpolaciones de los coeficientes β_L, γ_L de mezcla ascendente secos y húmedos se pueden por ejemplo controlar a través de la señalización S desde el lado de codificador para mejorar el manejo de transitorios. En caso de una conmutación de formato de codificación, el esquema de interpolación seleccionado en el lado de codificador, para interpolar los coeficientes β_L, γ_L de mezcla ascendente secos y húmedos en el lado de decodificador, puede por ejemplo ser un esquema de interpolación apropiado para una conmutación de formato de codificación, que puede ser diferente de los esquemas de interpolación empleados para los coeficientes β_L, γ_L de mezcla ascendente secos y húmedos cuando no se produce conmutación de formato de codificación.

En algunas realizaciones ejemplares, se puede emplear al menos un esquema de interpolación diferente en la sección 900 de decodificación que en la sección 1005 de decodificación adicional.

La Figura 12 es un diagrama de flujo de un método 1200 de decodificación de audio para reconstruir una señal de audio de M canales en base a una señal de mezcla descendente de dos canales y los parámetros de mezcla ascendente asociados, según una realización ejemplar. El método 1200 de decodificación se ejemplifica en la presente memoria mediante un método de decodificación que puede ser realizado por el sistema 1000 de decodificación de audio descrito con referencia a la Figura 10.

El método 1200 de decodificación de audio comprende: recibir 1201 la señal de mezcla descendente de dos canales L_1, L_2 y los parámetros α_L 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 Figuras 6 – 8, en base a la señal de mezcla descendente L_1, L_2; recibir 1202 la señalización S que indica uno seleccionado de los formatos de codificación F_1, F_2, F_3, descritos con referencia a las Figuras 6 – 8; y determinar 1203 el conjunto de coeficientes de descorrelación previa en base al formato de codificación indicado.

El método 1200 de decodificación de audio comprende detectar 1204 si el formato indicado conmuta de un formato de codificación a otro. Si no se detecta una conmutación, indicado por N en el diagrama de flujo, el siguiente paso es calcular 1205 la señal de entrada de descorrelación D_1, D_2, D_3 como un mapeo lineal de la señal de mezcla descendente L_1, L_2, en donde el conjunto de coeficientes de descorrelación previa se aplica a la señal de mezcla descendente. Si, por otro lado, se detecta una conmutación de formato de codificación, indicado por Y en el diagrama de flujo, el siguiente paso es en su lugar realizar 1206 interpolación en forma de una transición gradual de valores de los coeficientes de descorrelación previa de un formato de codificación a valores de los coeficientes de descorrelación previa de otro formato de codificación, y luego calcular 1205 la señal de entrada de descorrelación D_1, D_2, D_3 empleando los valores de los coeficientes de descorrelación previa interpolados.

El método 1200 de decodificación de audio comprende generar 1207 una señal descorrelacionada en base a la señal de entrada de descorrelación D_1, D_2, D_3; y determinar 1208 los conjuntos de coeficientes β_L, γ_L de mezcla ascendente húmedos y secos en base a los parámetros de mezcla ascendente recibidos y el formato de codificación indicado.

Si no se detecta conmutación de formato de codificación, indicado por una rama N de una caja 1209 de decisión, el método 1200 continúa calculando 1210 una señal de mezcla ascendente seca como un mapeo lineal de la señal de mezcla descendente, donde se aplica el conjunto de coeficientes β_L de mezcla ascendente secos a la señal de mezcla descendente L_1, L_2; y calculando 1211 una señal de mezcla ascendente húmeda como un mapeo lineal de la señal descorrelacionada, donde se aplica el conjunto de coeficientes γ_L de mezcla ascendente húmedos a la señal descorrelacionada. Si, por otro lado, el formato de codificación indicado conmuta de un formato de codificación a otro indicado por la rama Y de la caja 1209 de decisión, el método en su lugar continúa: realizando 1212 interpolación de valores de los coeficientes de mezcla ascendente secos y húmedos (incluyendo los coeficientes de valor cero) aplicables para un formato de codificación, a valores de los coeficientes de mezcla ascendente secos y húmedos (incluyendo los coeficientes de valor cero) aplicables para otro formato de codificación; calculando 1210 una señal de mezcla ascendente seca como un mapeo lineal de la señal de mezcla descendente L_1, L_2, donde el conjunto de coeficientes de mezcla ascendente secos interpolados se aplica a la señal de mezcla descendente L_1, L_2; y calculando 1211 una señal de mezcla ascendente húmeda como un mapeo lineal de la señal descorrelacionada, donde el conjunto de coeficientes de mezcla ascendente húmedos interpolados se aplica a la señal descorrelacionada. El método también comprende: combinar 1213 las señales de mezcla ascendente secas y húmedas para obtener la señal reconstruida multidimensional $\widetilde{L}$, $\widetilde{LS}$, $\widetilde{LB}$, $\widetilde{TFL}$, $\widetilde{TBL}$ correspondiente a la señal de audio de cinco canales a reconstruir.

La Figura 13 es un diagrama de bloques generalizado de una sección 1300 de decodificación para reconstruir una señal de audio de 13.1 canales en base a una señal de audio de 5.1 canales y los parámetros α de mezcla ascendente asociados, según una realización ejemplar.

En la presente realización ejemplar, la señal de audio de 13.1 canales se ejemplifica mediante los canales LW (ancho izquierdo), LSCRN (pantalla izquierdo), TFL (izquierdo frontal superior), LS (lateral izquierdo), LB (posterior izquierdo), TBL (izquierdo posterior superior), RW (ancho derecho), RSCRN (pantalla derecho), TFR (derecho frontal superior),

RS (lateral derecho), RB (posterior derecho), TBR (derecho posterior superior), C (central), y LFE (efectos de baja frecuencia). La señal de 5.1 canales comprende: una señal de mezcla descendente L_1, L_2, para la que un primer canal L_1 corresponde a una combinación lineal de los canales LW, LSCRN, TFL, y para la que un segundo canal L_2 corresponde a una combinación lineal de los canales LS, LB, TBL; una señal de mezcla descendente R_1, R_2 adicional para la que un primer canal R_1 corresponde a una combinación lineal de los canales RW, RSCRN, TFR, y para la que un segundo canal R_2 corresponde a una combinación lineal de los canales RS, RB, TBR; y los canales C y LFE.

Una primera sección 1310 de mezcla ascendente reconstruye los canales LW, LSCRN y TFL en base al primer canal L_1 de la señal de mezcla descendente bajo el control de al menos algunos de los parámetros α de mezcla ascendente; una segunda sección 1320 de mezcla ascendente reconstruye los canales LS, LB, TBL en base al segundo canal L_2 de la señal de mezcla descendente bajo el control de al menos de algunos de los parámetros α de mezcla ascendente; una tercera sección 1330 de mezcla ascendente reconstruye los canales RW, RSCRN, TFR en base al primer canal R_1 de la señal de mezcla descendente adicional bajo el control de al menos de algunos de los parámetros α de mezcla ascendente, y una cuarta sección 1340 de mezcla ascendente reconstruye los canales RS, RB, TBR en base al segundo canal R_2 de la señal de mezcla descendente bajo el control de al menos algunos de los parámetros α de mezcla ascendente. Se puede proporcionar una versión reconstruida $\widetilde{LW}$, $\widetilde{LSCRN}$, $\widetilde{TFL}$, $\widetilde{LS}$, $\widetilde{LB}$, $\widetilde{TBL}$, $\widetilde{RW}$, $\widetilde{RSCRN}$, $\widetilde{TFR}$, $\widetilde{RS}$, $\widetilde{RB}$, $\widetilde{TBR}$ de la señal de audio de 13.1 canales como salida de la sección 1310 de decodificación.

En una realización ejemplar, el sistema 1000 de decodificación de audio, descrito con referencia a la Figura 10 puede comprender la sección 1300 de decodificación además de las secciones 900 y 1005 de decodificación, o puede al menos ser operable reconstruir la señal de 13.1 canales por un método similar al realizado por la sección 1300 de decodificación. La señalización S extraída del flujo de bits B puede por ejemplo indicar si la señal de audio de 5.1 canales recibida L_1, L_2, R_1, R_2, C, LFE y los parámetros de mezcla ascendente asociados representan una señal de 11.1 canales, como se describe con referencia a la Figura 10, o si representa una señal de audio de 13.1 canales, como se describe con referencia a la Figura 13.

La sección 1009 de control puede detectar si la señalización S recibida indica una configuración de 11.1 canales o una configuración de 13.1 canales y puede controlar otras secciones del sistema 1000 de decodificación de audio para realizar la reconstrucción paramétrica de la señal de audio de 11.1 canales, como se describe con referencia a la Figura 10, o la señal de audio de 13.1 canales, como se describe con referencia a la Figura 13. Se puede por ejemplo emplear un único formato de codificación para la configuración de los 13.1 canales, en lugar de dos o tres formatos de codificación, como para la configuración de 11.1 canales. En caso de que la señalización S indique una configuración de 13.1 canales, el formato de codificación puede por lo tanto estar implícitamente indicado, y puede que no haya necesidad de que la señalización S indique explícitamente un formato de codificación seleccionado.

Se apreciará que aunque las realizaciones ejemplares descritas con referencia a las Figuras 1 – 5 se han formulado en términos de la señal de audio de 11.1 canales descrita con referencia a las Figuras 6 - 8, se pueden contemplar sistemas de codificación que pueden incluir cualquier número de secciones de codificación, y que se pueden configurar para codificar cualquier número de señales de audio de M canales, donde M ≥ 4. Similarmente, se apreciará que aunque las realizaciones ejemplares descritas con referencia a las Figuras 9 – 12 se han formulado en términos de la señal de audio de 11.1 canales descrita con referencia a las Figuras 6 – 8, se pueden contemplar sistemas de decodificación que pueden incluir cualquier número de secciones de decodificación, y que se pueden configurar para reconstruir cualquier número de señales de audio de M canales, donde M ≥ 4.

En algunas realizaciones, el lado de codificador puede seleccionar entre los tres formatos de codificación F_1, F_2, F_3. En otras realizaciones ejemplares, el lado de codificador puede seleccionar entre solo dos formatos de codificación, p.ej., el primer y segundo formatos de codificación F_1, F_2.

La Figura 14 es un diagrama de bloques generalizado de una sección 1400 de codificación para codificar una señal de audio de M canales como una señal de mezcla descendente de dos canales y los coeficientes de mezcla ascendente secos y húmedos asociados, según una realización ejemplar. La sección 1400 de codificación se puede disponer en un sistema de codificación de audio del tipo mostrado en la Figura 3. Más precisamente, se puede disponer en la ubicación ocupada por la sección 100 de codificación. Como quedará claro cuando se describan los funcionamientos internos de los componentes mostrados, la sección 1400 de codificación es operable en dos formatos de codificación distintos; se pueden implementar sin embargo secciones de codificación similares, sin desviarse del alcance de la invención, que sean operables en tres o más formatos de codificación.

La sección 1400 de codificación comprende una sección 1410 de mezcla descendente y una sección 1420 de análisis. Para al menos uno seleccionado (véase la descripción a continuación de una sección 1430 de control de la sección 1400 de codificación) de los formatos de codificación F_1, F_2, que pueden ser uno de los descritos con referencia a las Figuras 6 – 7 o pueden ser formatos diferentes, la sección 1410 de mezcla descendente calcula, de acuerdo con el formato de codificación, una señal de mezcla descendente de dos canales L_1, L_2 en base a la señal de audio de cinco canales L, LS, LB, TFL, TBL. En por ejemplo el primer formato de codificación F_1, el primer canal L_1 de la señal de mezcla descendente se forma como una combinación lineal (p.ej., una suma) de un primer grupo de canales de la señal de audio de cinco canales L, LS, LB, TFL, TBL, y el segundo canal L_2 de la señal de mezcla descendente se forma como una combinación lineal (p.ej., una suma) de un segundo grupo de canales de la señal de audio de cinco

canales L, LS, LB, TFL, TBL. La operación realizada por la sección 1410 de mezcla descendente se puede por ejemplo expresar como la ecuación (1).

Para al menos dicho uno seleccionado de los formatos de codificación F_1, F_2, la sección 1420 de análisis determina un conjunto de coeficientes β_L de mezcla ascendente secos que definen un mapeo lineal de la respectiva señal de mezcla descendente L_1, L_2 que se aproxima a la señal de audio de cinco canales L, LS, LB, TFL, TBL. Para cada uno de los formatos de codificación F_1, F_2, la sección 1420 de análisis determina además un conjunto de coeficientes γ_L de mezcla ascendente húmedos, en base a la respectiva diferencia calculada, que junto con los coeficientes β_L de mezcla ascendente secos permite la reconstrucción paramétrica según la ecuación (2) de la señal de audio de cinco canales L, LS, LB, TFL, TBL a partir de la señal de mezcla descendente L_1, L_2 y a partir de una señal descorrelacionada de tres canales determinada en un lado de decodificador en base a la señal de mezcla descendente L_1, L_2. El conjunto de coeficientes γ_L de mezcla ascendente húmedos define un mapeo lineal de la señal descorrelacionada de tal manera que la matriz de covarianza de la señal obtenida por el mapeo lineal de la señal descorrelacionada se aproxima a la diferencia entre la matriz de covarianza de la señal de audio de cinco canales L, LS, LB, TFL, TBL recibida y la matriz de covarianza de la señal de audio de cinco canales aproximada por el mapeo lineal de la señal de mezcla descendente L_1, L_2.

La sección 1410 de mezcla descendente puede por ejemplo calcular la señal de mezcla descendente L_1, L_2 en el dominio del tiempo, i.e., en base a una representación de dominio del tiempo de la señal de audio de cinco canales L, LS, LB, TBL, TBL, o en un dominio de la frecuencia, i.e., en base a una representación de dominio de la frecuencia de la señal de audio de cinco canales L, LS, LB, TFL, TBL. Es posible calcular L_1, L_2 en el dominio del tiempo al menos si la decisión sobre un formato de codificación no es selectiva en frecuencia, y por tanto aplica para todos los componentes de frecuencia de la señal de audio de M canales; este es el caso preferido actualmente.

La sección 1420 de análisis puede por ejemplo determinar los coeficientes β_L de mezcla ascendente secos y los coeficientes γ_L de mezcla ascendente húmedos en base a un análisis de dominio de la frecuencia de la señal de audio de cinco canales L, LS, LB, TFL, TBL. El análisis de dominio de la frecuencia se puede realizar sobre una sección de ventana de la señal de audio de M canales. Para las ventanas, se pueden por ejemplo utilizar ventanas rectangulares disjuntas o triangulares solapadas. La sección 1420 de análisis puede por ejemplo recibir la señal de mezcla descendente L_1, L_2 calculada por la sección 1410 de mezcla descendente (no mostrada en la Figura 14), o puede calcular su propia versión de la señal de mezcla descendente L_1, L_2, para el propósito específico de determinar los coeficientes β_L de mezcla ascendente secos y los coeficientes γ_L Y de mezcla ascendente húmedos.

La sección 1400 de codificación comprende además una sección 1430 de control, que es responsable de seleccionar un formato de codificación a ser utilizado actualmente. No es esencial que la sección 1430 de control utilice un criterio particular o razonamiento particular para decidir el formato de codificación a seleccionar. El valor de la señalización S generada por la sección 1430 de control indica el resultado de la toma de decisión de la sección 1430 de control para una sección actualmente considerada (p.ej., una trama de tiempo) de la señal de audio de M canales. La señalización S se puede incluir en un flujo de bits B producido por el sistema 300 de codificación en el que está incluida la sección 1400 de codificación, para facilitar la reconstrucción de la señal de audio codificada. Adicionalmente, la señalización S se alimenta a cada una de la sección 1410 de mezcla descendente y la sección 1420 de análisis, para informar a estas secciones del formato de codificación a utilizar. Como la sección 1420 de análisis, la sección 1430 de control puede considerar secciones de ventana de la señal de M canales. Se observa para exhaustividad que la sección 1410 de mezcla descendente puede operar con retardos de 1 o 2 tramas y posiblemente con adelanto adicional, con respecto a la sección 1430 de control. Opcionalmente, la señalización S puede también contener información relativa a un desvanecimiento cruzado de la señal de mezcla descendente que produce la sección 1410 de mezcla descendente y/o información relativa a una interpolación del lado de decodificador de valores discretos de los coeficientes de mezcla ascendente secos y húmedos que proporciona la sección 1420 de análisis, para asegurar la sincronicidad en una escala de tiempo de subtrama.

Como un componente adicional, la sección 1400 de codificación puede incluir un estabilizador 1440 dispuesto inmediatamente aguas abajo de la sección 1430 de control y que actúa sobre su señal de salida inmediatamente antes de que sea procesada por otros componentes. En base a esta señal de salida, el estabilizador 1440 suministra la información lateral S a los componentes aguas abajo. El estabilizador 1440 puede implementar el objetivo deseable de no cambiar el formato de codificación seleccionado demasiado frecuentemente. Para este propósito, el estabilizador 1440 puede considerar una serie de selecciones de formato de codificación para tramas de tiempo anteriores de la señal de audio de M canales y asegurar que se mantiene un formato de codificación elegido durante al menos un número predefinido de tramas de tiempo. Alternativamente, el estabilizador puede aplicar un filtro de promediado a una serie de selecciones de formatos de codificación anteriores (p.ej., representadas como una variable discreta), lo que puede provocar un efecto de suavizado. Como otra alternativa más, el estabilizador 1440 puede comprender una máquina de estado configurada para suministrar información lateral S para todas las tramas de tiempo en una ventana de tiempo móvil si la máquina de estado determina que la selección de formato de codificación proporcionada por la sección 1430 de control ha permanecido estable a lo largo de la ventana de tiempo móvil. La ventana de tiempo móvil puede corresponder a una memoria intermedia que almacena las selecciones formato de codificación para una serie de tramas de tiempo anteriores. Como se da cuenta fácilmente el experto en la técnica que estudie esta descripción, tales funcionalidades de estabilización pueden necesitar ir acompañadas por un aumento en el retardo operacional

entre el estabilizador 1440 y al menos la sección 1410 de mezcla descendente y la sección 1420 de análisis. El retardo se puede implementar por medio de secciones de almacenamiento en memoria intermedia de la señal de audio de M canales.

Se recuerda que la Figura 14 es una vista parcial del sistema de codificación en la Figura 3. Mientras que los componentes mostrados en la Figura 14 solo se refieren al procesamiento de los canales laterales izquierdos L, LS, LB, TFL, TBL, el sistema de codificación procesa al menos los canales laterales derechos R, RS, RB, TFR, TBR también. Por ejemplo, una instancia adicional (p.ej., una réplica funcionalmente equivalente) de la sección 1400 de codificación puede operar en paralelo para codificar una señal lateral derecha que incluye dichos canales R, RS, RB, TFR, TBR. Aunque los canales laterales izquierdos y laterales derechos contribuyen a dos señales de mezcla descendente separadas (o al menos a grupos separados de canales de una señal de mezcla descendente común), se prefiere utilizar un formato de codificación común para todos los canales. Es decir, la sección 1430 de control dentro de la sección 1400 de codificación lateral izquierda puede ser responsable de decidir sobre un formato de codificación común a utilizar tanto para los canales laterales izquierdos como laterales derechos; es entonces preferible que la sección 1430 de control tenga acceso a los canales laterales derechos R, RS, RB, TFR, TBR también o a cantidades derivadas a partir de estas señales, tal como una covarianza, una señal de mezcla descendente, etc., y puede tenerlas en cuenta al decidir sobre un formato de codificación a utilizar. La señalización S se proporciona luego no solo a la sección 1410 de mezcla descendente y la sección 1420 de análisis de la sección 1430 de control (lateral izquierda), sino también a las secciones equivalentes de una sección de codificación lateral derecha (no mostrada). Alternativamente, el propósito de utilizar un formato de codificación común para todos los canales se puede lograr permitiendo que la propia sección 1430 de control sea común tanto a una instancia lateral izquierda de la sección 1400 de codificación como a una instancia lateral derecha de la misma. En un diseño del tipo representado en la Figura 3, la sección 1430 de codificación se puede proporcionar fuera tanto de la sección 100 de codificación como de la sección 303 de codificación adicional, que son responsables de los canales laterales izquierdos y laterales derechos, respectivamente, que reciben todos los canales laterales derechos y laterales izquierdos L, LS, LB, TFL, TBL, R, RS, RB, TFR, TBR y emiten la señalización S, que indica una selección de un formato de codificación y se suministra al menos a la sección 100 de codificación y la sección 303 de codificación adicional.

La Figura 15 representa esquemáticamente una posible implementación de una sección 1410 de mezcla descendente configurada para alternar, de acuerdo con la señalización S, entre dos formatos de codificación F_1, F_2 predefinidos y proporcionar un desvanecimiento cruzado de estos. La sección 1410 de mezcla descendente comprende dos subsecciones 1411, 1412 de mezcla descendente configuradas para recibir la señal de audio de M canales y emitir una señal de mezcla descendente de dos canales. Las dos subsecciones 1411, 1412 de mezcla descendente pueden ser copias funcionalmente equivalentes de un diseño, aunque configuradas con diferentes configuraciones de mezcla descendente (p.ej., valores de los coeficientes para producir la señal de mezcla descendente L_1, L_2 en base a la señal de audio de M canales). En operación normal, las dos subsecciones 1411, 1412 de mezcla descendente juntas proporcionan una señal de mezcla descendente $L_1(F_1)$, $L_2(F_1)$ de acuerdo con el primer formato de codificación F_1 y/o una señal de mezcla descendente $L_1(F_2)$, $L_2(F_2)$ de acuerdo con el segundo formato de codificación F_2. Aguas abajo de las subsecciones 1411, 1412 de mezcla descendente, hay dispuestas una primera sección 1413 de interpolación de mezcla descendente y una segunda sección 1414 de interpolación de mezcla descendente. La primera sección 1413 de interpolación de mezcla descendente se configura para interpolar, incluyendo desvanecimiento cruzado, un primer canal L_1 de la señal de mezcla descendente, y la segunda sección 1414 de interpolación de mezcla descendente se configura para interpolar, incluyendo desvanecimiento cruzado, un segundo canal L_2 de la señal de mezcla descendente. La primera sección 1413 de interpolación de mezcla descendente es operable en al menos los siguientes estados:

a) primer formato de codificación solamente ($L_1 = L_1(F_1)$), como se puede utilizar en la operación de estado estable en el primer formato de codificación;

b) segundo formato de codificación solamente ($L_1 = L_1(F_2)$), como se puede utilizar en la operación de estado estable en el segundo formato de codificación; y

c) mezcla de canales de mezcla descendente según ambos formatos de codificación ($L_1 = \alpha_1 L_1(F_1) + \alpha_2 L_1(F_2)$, en donde $0 < \alpha_1 < 1$ y $0 < \alpha_2 < 1$), como se puede utilizar en una transición del primer al segundo formato de codificación o viceversa.

Mezclar el estado (c) puede requerir que las señales de mezcla descendente estén disponibles desde tanto la primera como la segunda subsecciones 1411, 1412 de mezcla descendente. Preferiblemente, la primera sección 1413 de interpolación de mezcla descendente es operable en una pluralidad de estados de mezcla (c), de modo que es posible una transición en subpasos finos, o incluso un desvanecimiento cruzado casi continuo. Esto tiene la ventaja de hacer un desvanecimiento cruzado menos perceptible. Por ejemplo, en un diseño interpolador donde $\alpha_1 + \alpha_2 = 1$, es posible un desvanecimiento cruzado de cinco pasos si se definen los siguientes valores de (α_1, α_2): (0,2; 0,8) (0,4; 0,6), (0,6; 0,4), (0,8; 0,2). La segunda sección 1414 de interpolación de mezcla descendente puede tener capacidades idénticas o similares.

En una variación a la realización anterior de la sección 1410 de mezcla descendente, como se sugiere por la línea discontinua en la Figura 15, la señalización S se puede alimentar a la primera y segunda subsecciones 1411, 1412 de

mezcla descendente también. Como se explicó anteriormente, se puede entonces suprimir la generación de la señal de mezcla descendente asociada con el formato de codificación no seleccionado. Esto puede reducir la carga computacional promedio.

Adicionalmente o alternativamente a esta variación, el desvanecimiento cruzado entre señales de mezcla descendente de dos formatos de codificación diferentes se puede lograr mediante el desvanecimiento cruzado de los coeficientes de mezcla descendente. La primera subsección 1411 de mezcla descendente puede entonces ser alimentada por coeficientes de mezcla descendente interpolados, que son producidos por un interpolador de coeficientes (no mostrado) que almacena valores predefinidos de coeficientes de mezcla descendente a ser utilizados en los formatos de codificación F_1, F_2 disponibles, y recibe como entrada la señalización S. En esta configuración, la segunda subsección 1412 de mezcla descendente y la primera y segunda subsecciones 1413, 1414 de interpolación se pueden eliminar o desactivar permanentemente.

La señalización S que recibe la sección 1410 de mezcla descendente se suministra al menos a las secciones 1413, 1414 de interpolación de mezcla descendente, pero no necesariamente a las subsecciones 1411, 1412 de mezcla descendente. Es necesario suministrar la señalización S a las subsecciones 1411, 1412 de mezcla descendente si se desea una operación alterna, es decir, si se debe disminuir la cantidad de mezcla descendente redundante fuera de las transiciones entre formatos de codificación. La señalización puede ser comandos de bajo nivel, p.ej., que se refieren a diferentes modos operacionales de las secciones 1413, 1414 de interpolación de mezcla descendente, o se puede referir a instrucciones de alto nivel, tal como una orden para ejecutar un programa de desvanecimiento cruzado predefinido (p.ej., una sucesión de los modos operacionales en donde cada uno tiene una duración predefinida) en un punto de inicio indicado.

Volviendo a la Figura 16, se representa una posible implementación de una sección 1420 de análisis configurada para alternar, de acuerdo con la señalización S, entre dos formatos de codificación F_1, F_2 predefinidos. La sección 1420 de análisis comprende dos subsecciones 1421, 1422 de análisis configuradas para recibir la señal de audio de M canales y emitir los coeficientes de mezcla ascendente secos y húmedos. Las dos subsecciones 1421, 1422 de análisis pueden ser copias funcionalmente equivalentes de un diseño. En operación normal, las dos subsecciones 1421, 1422 de análisis juntas proporcionan un conjunto de coeficientes $\beta_L(F_1)$, $\gamma_L(F_1)$ de mezcla ascendente secos y húmedos de acuerdo con el primer formato de codificación F_1 y/o un conjunto de coeficientes $\beta_L(F_2)$, $\gamma_L(F_2)$ de mezcla ascendente secos y húmedos de acuerdo con el segundo formato de codificación F_2.

Como se explicó anteriormente para la sección 1420 análisis en su conjunto, se puede recibir la señal de mezcla descendente actual de la sección 1410 de mezcla descendente, o se puede producir un duplicado de esta señal en la sección 1420 de análisis. Más precisamente, la primera subsección 1421 de análisis puede recibir la señal de mezcla descendente $L_1(F_1)$, $L_2(F_1)$ según el primer formato de codificación F_1 de la primera subsección 1411 de mezcla descendente en la sección 1410 de mezcla descendente, o puede producir un duplicado por sí misma. Similarmente, la segunda subsección 1422 de análisis puede recibir la señal de mezcla descendente $L_1(F_2)$, $L_2(F_2)$ según el segundo formato de codificación F_2 de la segunda subsección 1412 de mezcla descendente, o puede producir un duplicado de esta señal por sí misma.

Aguas abajo de las secciones 1421, 1422 de análisis, hay dispuesto un selector 1423 de coeficientes de mezcla ascendente secos y un selector 1424 de coeficientes de mezcla ascendente húmedos. El selector 1423 de coeficientes de mezcla ascendente secos se configura para reenviar un conjunto de coeficientes β_L de mezcla ascendente secos desde la primera o segunda subsección 1421, 1422 de análisis, y el selector 1424 de coeficientes de mezcla ascendente húmedos se configura para reenviar un conjunto de coeficientes γ_L de mezcla ascendente húmedos desde la primera o segunda subsección 1421, 1422 de análisis. El selector 1423 de coeficientes de mezcla ascendente secos es operable en al menos los estados (a) y (b) discutidos anteriormente para la primera sección 1413 de interpolación de mezcla descendente. Sin embargo, si el sistema de codificación de la Figura 3, del cual se está describiendo aquí una porción, se configura para cooperar con un sistema de decodificación que, como el mostrado en la Figura 9, realiza la reconstrucción paramétrica en base a valores discretos interpolados de los coeficientes de mezcla ascendente que recibe, entonces no hay necesidad de configurar un estado de mezcla como (c) definido por las secciones 1413, 1414 de interpolación de mezcla descendente. El selector 1424 de coeficientes de mezcla ascendente húmedos puede tener capacidades similares.

La señalización S que recibe la sección 1420 de análisis se suministra al menos a los selectores 1423, 1424 de coeficientes de mezcla ascendente húmedos y secos. No es necesario que las subsecciones 1421, 1422 de análisis reciban la señalización, aunque esto es ventajoso para evitar el cálculo redundante de los coeficientes de mezcla ascendente fuera de las transiciones. La señalización puede ser comandos de bajo nivel, p.ej., que se refieren a diferentes modos operaciones de los selectores 1423, 1424 de coeficientes de mezcla ascendente secos y húmedos, o se puede referir a instrucciones de alto nivel, tal como una orden para pasar de un formato de codificación a otro en una trama de tiempo dada. Como se explicó anteriormente, esto preferiblemente no implica una operación de desvanecimiento cruzado pero puede equivaler a definir los valores de los coeficientes de mezcla ascendente para un punto adecuado en el tiempo, o definir estos valores para aplicarlos en un punto adecuado en el tiempo.

Se describirá ahora un método 1700 que es una variación del método para codificar una señal de audio de M canales como una señal de mezcla descendente de dos canales, según una realización ejemplar, que se representó

esquemáticamente como un diagrama de flujo en la Figura 17. El método ejemplificado aquí puede ser realizado por un sistema de codificación de audio que comprende la sección 1400 de codificación que se ha descrito anteriormente con referencia a las Figuras 14 – 16.

El método 1700 de codificación de audio comprende: recibir 1710 la señal de audio de M canales L, LS, LB, TFL, TBL; seleccionar 1720 uno de al menos dos de los formatos de codificación F_1, F_2, F_3 descritos con referencia a las Figuras 6 – 8; calcular 1730, para el formato de codificación seleccionado, una señal de mezcla descendente de dos canales L_1, L_2 en base a la señal de audio de M canales L, LS, LB, TFL, TBL; emitir 1740 la señal de mezcla descendente L_1, L_2 del formato de codificación seleccionado e información lateral α que permite la reconstrucción paramétrica de la señal de audio de M canales en base a la señal de mezcla descendente; y emitir 1750 la señalización S que indica el formato de codificación seleccionado. El método se repite, p.ej., para cada trama de tiempo de la señal de audio de M canales. Si el resultado de la selección 1720 es un formato de codificación diferente al seleccionado inmediatamente antes, entonces la señal de mezcla descendente se reemplaza, durante una duración adecuada, por un desvanecimiento cruzado entre señales de mezcla descendente de acuerdo con los formatos de codificación previo y actual. Como ya se ha discutido, no es necesario o no es posible realizar un desvanecimiento cruzado de la información lateral, que puede estar sujeta a interpolación inherente del lado de decodificador.

Se observa que el método descrito aquí se puede implementar sin uno o más de los cuatro pasos 430, 440, 450 y 470 representados en la Figura 4.

IV. Equivalentes, extensiones, alternativas y misceláneos

Incluso aunque la presente descripción describe y representa realizaciones ejemplares específicas, la invención no se limita a estos ejemplos específicos. Se pueden hacer modificaciones y variaciones a las realizaciones ejemplares anteriores sin desviarse del alcance de la invención, que está definida por las reivindicaciones adjuntas solamente.

En las reivindicaciones, la palabra “comprender” no excluye otros elementos o pasos, y el artículo indefinido “un” o “una” no excluye una pluralidad. El mero hecho de que ciertas medidas sean enumeradas en reivindicaciones dependientes mutuamente diferentes no indica que no se pueda utilizar una combinación de estas medidas ventajosamente. Cualquier signo de referencia que aparezca en las reivindicaciones no debe entenderse como limitativo de su alcance.

Los dispositivos y métodos descritos anteriormente se pueden implementar como software, firmware, 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 corresponde necesariamente a la división en unidades físicas; al contrario, un componente físico puede tener múltiples funcionalidades, y una tarea puede ser llevada a cabo de manera 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 señal o microprocesador, o se pueden implementar como hardware o como un circuito integrado de aplicación específica. Tal software puede ser distribuido en medios legibles por ordenador, que pueden comprender medios (o medios no transitorios) de almacenamiento informático y medios (o medios transitorios) de comunicación. Como es bien conocido para un experto en la técnica, el término medios de almacenamiento informático incluye medios tanto volátiles como no volátiles, extraíbles y no extraíbles implementados en cualquier método o tecnología para el almacenamiento de información tal como instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos. Los medios de almacenamiento informático incluyen, pero no se limitan a, RAM, ROM, EEPROM, memoria flash 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 magnéticos, o cualquier otro medio que se pueda utilizar para almacenar la información deseada y que pueda ser accedido por un ordenador. Además, es bien conocido para el experto en la técnica que los medios de comunicación típicamente representan 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 incluyen cualquier medio de entrega de información.

REIVINDICACIONES

1. Un método (1200) de decodificación de audio que comprende:

recibir (1201) una señal de mezcla descendente de dos canales (L_1, L_2) y parámetros (α_L) de mezcla ascendente 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 ≥ 4;

recibir (1202) señalización (S) que indica uno seleccionado de al menos dos formatos de codificación (F_1, F_2, F_3) de la señal de audio de M canales, en donde los formatos de codificación corresponden a respectivas particiones diferentes de los canales de la señal de audio de M canales en respectivos primer y segundo grupos (601, 602) de uno o más canales, en donde, en el formato de codificación indicado, un primer canal de la señal de mezcla descendente corresponde a una combinación lineal del 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 del segundo grupo de uno o más canales de la señal de audio de M canales;

determinar (1203) un conjunto de coeficientes de descorrelación previa en base al formato de codificación indicado;

calcular (1205) una señal de entrada de descorrelación (D_1, D_2, D_3) como un mapeo lineal de la señal de mezcla descendente, en donde el conjunto de coeficientes de descorrelación previa se aplica a la señal de mezcla descendente, en donde los coeficientes de descorrelación previa se determinan de tal manera que un primer canal (TBL) de la señal de audio de M canales contribuye, a través de la señal de mezcla descendente, a un primer canal fijo (D3) de la señal de entrada de descorrelación en al menos dos de los formatos de codificación;

generar (1207) una señal descorrelacionada en base a la señal de entrada de descorrelación;

determinar (1208) conjuntos de coeficientes (γ_L, β_L) de mezcla ascendente húmedos y secos en base a los parámetros de mezcla ascendente recibidos y el formato de codificación indicado;

calcular (1210) una señal (X_1, X_2) de mezcla ascendente seca como un mapeo lineal de la señal de mezcla descendente, en donde el conjunto de coeficientes de mezcla ascendente secos se aplica a la señal de mezcla descendente;

calcular (1211) una señal (Y_1, Y_2) de mezcla ascendente húmeda como un mapeo lineal de la señal descorrelacionada, en donde el conjunto de coeficientes de mezcla ascendente húmedos se aplica a la señal descorrelacionada; y

combinar (1213) las señales de mezcla ascendente secas y húmedas para obtener una señal reconstruida multidimensional ($\widetilde{L}$, $\widetilde{LS}$, $\widetilde{LB}$, $\widetilde{TFL}$, $\widetilde{TBL}$) correspondiente a la señal de audio de M canales a reconstruir;

caracterizado por que

la señal de audio de M canales tiene una configuración de canales predefinida, y el formato de codificación seleccionado indicado conmuta entre los al menos dos formatos de codificación.

2. El método de decodificación de audio de la reivindicación 1, en donde la señal de entrada de descorrelación y la señal descorrelacionada comprende cada una M – 2 canales, en donde un canal de la señal descorrelacionada se genera en base a no más de un canal de la señal de entrada de descorrelación, y en donde los coeficientes de descorrelación previa se determinan de tal manera que, en cada uno de los formatos de codificación, un canal de la señal de entrada de descorrelación recibe una contribución de no más de un canal de la señal de mezcla descendente.

3. El método de decodificación de audio de cualquiera de las reivindicaciones precedentes, en donde los coeficientes de descorrelación previa se determinan de tal manera que, adicionalmente, un segundo canal (L) de la señal de audio de M canales contribuye, a través de la señal de mezcla descendente, a un segundo canal fijo (D1) de la señal de entrada de descorrelación en al menos dos de los formatos de codificación; y/o en donde los coeficientes de descorrelación previa se determinan de tal manera que un par de canales (LS, LB) de la señal de audio de M canales contribuye, a través de la señal de mezcla descendente, a un tercer canal fijo (D2) de la señal de entrada de descorrelación en al menos dos de los formatos de codificación.

4. El método de decodificación de audio de cualquiera de las reivindicaciones precedentes, que comprende además:

en respuesta a la detección de una conmutación del formato de codificación indicado de un primer formato de codificación a un segundo formato de codificación, realizar (1206) una transición gradual de valores de los coeficientes de descorrelación previa asociados con el primer formato de codificación a valores de los coeficientes de descorrelación previa asociados con el segundo formato de codificación.

5. El método de decodificación de audio de cualquiera de las reivindicaciones precedentes, en donde la señal de audio de M canales comprende tres canales (L, 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 verticalmente separadas de las de dichos tres canales en dicho entorno de reproducción.

6. El método de decodificación de audio de la reivindicación 5, en donde, en un primer formato de codificación (F_1), dicho segundo grupo comprende dichos dos canales y/o, en donde, en un primer formato de codificación (F_1), dicho primer grupo comprende dichos tres canales y dicho segundo grupo comprende dichos dos canales y/o, en donde, en un segundo formato de codificación (F_2), cada uno del primer y segundo grupos comprende uno de dichos dos canales.

7. El método de decodificación de audio de cualquiera de las reivindicaciones precedentes, en donde, en un formato de codificación (F_1, F_2) particular, dicho primer grupo consta de N canales, en donde $N \geq 3$, y en donde, en respuesta a que el formato de codificación indicado sea el formato de codificación particular:

los coeficientes de descorrelación previa se determinan de tal manera que se generan N – 1 canales de la señal descorrelacionada en base al primer canal de la señal de mezcla descendente; y

los coeficientes de mezcla ascendente secos y húmedos se determinan de tal manera que dicho primer grupo se reconstruye como un mapeo lineal del primer canal de la señal de mezcla descendente y dichos N – 1 canales de la señal descorrelacionada, en donde un subconjunto de los coeficientes de mezcla ascendente secos se aplica al primer canal de la señal de mezcla descendente y un subconjunto de los coeficientes de mezcla ascendente húmedos se aplica a dichos N – 1 canales de la señal descorrelacionada.

8. Un método de decodificación de audio que comprende:

recibir señalización (S) que indica una de al menos dos configuraciones de canales predefinidas;

en respuesta a la detección de la señalización recibida que indica una primera configuración de canales (L, LS, LB, TFL, TBL) predefinida, realizar el método de decodificación de audio de cualquiera de las reivindicaciones precedentes; y

en respuesta a la detección de la señalización recibida que indica una segunda configuración de canales (LW, LSCRN, TFL, LS, LB, TBL) predefinida

recibir una señal de mezcla descendente de dos canales (L_1, L_2) y los parámetros (α) de mezcla ascendente asociados,

realizar reconstrucción paramétrica de una primera señal de audio de tres canales (LW, LSCRN, TFL) en base a un primer canal (L_1), de la señal de mezcla descendente y al menos algunos de los parámetros de mezcla ascendente, y

realizar reconstrucción paramétrica de una segunda señal de audio de tres canales (LS, LB, TBL) en base a un segundo canal (L_2), de la señal de mezcla descendente y al menos algunos de los parámetros de mezcla ascendente.

9. Un sistema (1000) de decodificación de audio que comprende uno o más componentes configurados para realizar el método de cualquiera de las reivindicaciones 1 – 8.

10. El sistema de decodificación de audio de la reivindicación 9, en donde el uno o más componentes se configuran además para:

reconstruir una señal de audio de M canales (R, RS, RB, TFR, TBR) adicional en base a una señal de mezcla descendente de dos canales (R_1, R_2) adicional y los parámetros (α_R) de mezcla ascendente adicionales asociados;

recibir señalización (S) que indica uno seleccionado de al menos dos formatos de codificación de la señal de audio de M canales adicional, correspondiendo los formatos de codificación de la señal de audio de M canales adicional a respectivas particiones diferentes de los canales de la señal de audio de M canales adicional en respectivos primer y segundo grupos (603, 604) de uno o más canales, en donde, en el formato de codificación indicado de la señal de audio de M canales adicional, un primer canal (R_1) de la señal de mezcla descendente adicional corresponde a una combinación lineal del primer grupo de uno o más canales de la señal de audio de M canales adicional y un segundo canal (R_2) de la señal de mezcla descendente adicional corresponde a una combinación lineal del segundo grupo de uno o más canales de la señal de audio de M canales adicional;

determinar un conjunto de coeficientes de descorrelación previa adicional en base al formato de codificación indicado de la señal de audio de M canales adicional;

calcular una señal de entrada de descorrelación adicional como un mapeo lineal de la señal de mezcla descendente adicional, en donde el conjunto de coeficientes de descorrelación previa adicional se aplica a la señal de mezcla descendente adicional;

generar una señal descorrelacionada adicional en base a la señal de entrada de descorrelación adicional;

determinar conjuntos de coeficientes de mezcla ascendente húmedos y secos adicionales en base a los parámetros de mezcla ascendente adicionales recibidos y el formato de codificación indicado de la señal de audio de M canales adicional;

calcular una señal de mezcla ascendente seca adicional como un mapeo lineal de la señal de mezcla descendente adicional, en donde el conjunto de coeficientes de mezcla ascendente secos adicional se aplica a la señal de mezcla descendente adicional;

calcular una señal de mezcla ascendente húmeda adicional como un mapeo lineal de la señal descorrelacionada adicional, en donde el conjunto de coeficientes de mezcla ascendente húmedos adicional se aplica a la señal descorrelacionada adicional; y

combinar las señales de mezcla ascendente secas y húmedas adicionales para obtener una señal reconstruida multidimensional ($\widetilde{R}$, $\widetilde{RS}$, $\widetilde{RB}$, $\widetilde{TFR}$, $\widetilde{TBR}$) adicional correspondiente a la señal de audio de M canales adicional a reconstruir.

11. El sistema de decodificación de audio de cualquiera de las reivindicaciones 9 – 10, en donde el uno o más componentes se configura además para:

extraer, a partir de un flujo de bits (B), la señal de mezcla descendente, los parámetros de mezcla ascendente asociados con la señal de mezcla descendente, y un canal (C) de audio codificado de manera discreta; y

decodificar dicho canal de audio codificado de manera discreta.

12. Un método (1700) de codificación de audio, que comprende:

recibir (1710) una señal de audio de M canales (L, LS, LB, TFL, TBL), donde M ≥ 4;

seleccionar (1720) repetidamente uno de al menos dos formatos de codificación (F_1, F_2, F_3) correspondientes a respectivas particiones diferentes de los canales de la señal de audio de M canales en respectivos primer y segundo grupos (601, 602) de uno o más canales cada uno, en donde cada uno de los formatos de codificación define una señal de mezcla descendente de dos canales (L_1, L_2), en la que un primer canal (L_1) de la señal de mezcla descendente se forma como una combinación lineal del primer grupo de uno o más canales de la señal de audio de M canales, y en donde un segundo canal (L_2) de la señal de mezcla descendente se forma como una combinación lineal del segundo grupo de uno o más canales de la señal de audio de M canales;

para el formato de codificación actualmente seleccionado, determinar un conjunto de coeficientes (β_L) de mezcla ascendente secos y un conjunto de coeficientes (γ_L) de mezcla ascendente húmedos;

calcular (1730), de acuerdo con el formato de codificación actualmente seleccionado, una señal de mezcla descendente de dos canales (L_1, L_2) en base a la señal de audio de M canales;

emitir (1740) la señal de mezcla descendente del formato de codificación actualmente seleccionado, estando la señal de mezcla descendente segmentada en tramas de tiempo, y permitiendo la información lateral la reconstrucción paramétrica de la señal de audio de M canales en base a la señal de mezcla descendente y una señal descorrelacionada determinada en base a al menos un canal de la señal de mezcla descendente del formato de codificación seleccionado, comprendiendo la información lateral valores discretos de los conjuntos de coeficientes (β_L, γ_L) de mezcla ascendente secos y húmedos, en donde se emite al menos un valor discreto por trama de tiempo; y

emitir (1750) señalización (S) que indica el formato de codificación actualmente seleccionado, caracterizado por que

la señal de audio de M canales tiene una configuración de canales predefinida,

en respuesta a un cambio de un primer formato de codificación seleccionado a un segundo formato de codificación seleccionado distinto, se calcula una señal de mezcla descendente según el segundo formato de codificación seleccionado, y se emite un desvanecimiento cruzado de la señal de mezcla descendente según el primer formato de codificación seleccionado y la señal de mezcla descendente según el segundo formato de codificación seleccionado en lugar de la señal de mezcla descendente, y

la reconstrucción paramétrica de la señal de audio de M canales entre los valores discretos debe basarse en valores interpolados de los conjuntos de coeficientes (β_L, γ_L) de mezcla ascendente secos y húmedos según una regla de interpolación predefinida, en donde el desvanecimiento cruzado de la señal de mezcla descendente y los valores discretos de los conjuntos de coeficientes de mezcla ascendente secos y húmedos se emiten de tal manera que dicho desvanecimiento cruzado e interpolación serán síncronos.

13. Un sistema (300) de codificación de audio que comprende una sección (1400) de codificación configurada para codificar una señal de audio de M canales (L, LS, LB, TFL, TBL) como una señal de mezcla descendente de dos canales y los parámetros de mezcla ascendente asociados, donde M ≥ 4, comprendiendo la sección de codificación:

una sección (1411, 1412) de mezcla descendente configurada para, para al menos uno de al menos dos formatos de codificación (F_1, F_2, F_3) correspondientes a respectivas particiones diferentes de los canales de la señal de audio de M canales en respectivos primer y segundo grupos (601, 602) de uno o más canales cada uno, calcular, de acuerdo con el formato de codificación, una señal de mezcla descendente de dos canales (L_1, L_2) en base a la señal de audio de M canales, estando la señal de mezcla descendente segmentada en tramas de tiempo, en donde un primer canal (L_1) de la señal de mezcla descendente se forma como una combinación lineal del primer grupo de uno o más canales de la señal de audio de M canales y un segundo canal (L_2) de la señal de mezcla descendente se forma como una combinación lineal del segundo grupo de uno o más canales de la señal de audio de M canales;

una sección (1430) de control configurada para seleccionar repetidamente uno de los formatos de codificación,

en donde el sistema de codificación de audio se configura para, para el formato de codificación actualmente seleccionado, determinar un conjunto de coeficientes (β_L) de mezcla ascendente secos y un conjunto de coeficientes (γ_L) de mezcla ascendente húmedos, y emitir señalización (S) que indica el formato de codificación actualmente seleccionado e información lateral (α) que permite la reconstrucción paramétrica de la señal de audio de M canales en base a la señal de mezcla descendente y una señal descorrelacionada determinada en base a al menos un canal de la señal de mezcla descendente del formato de codificación seleccionado, comprendiendo la información lateral valores discretos de los conjuntos de coeficientes (β_L, γ_L) de mezcla ascendente secos y húmedos, en donde se emite al menos un valor discreto por trama de tiempo,

caracterizado por que

la señal de audio de M canales tiene una configuración de canales predefinida,

el sistema (300) de codificación de audio comprende un interpolador (1413, 1414) de mezcla descendente configurado para producir un desvanecimiento cruzado de la señal de mezcla descendente según un primer formato de codificación, que ha sido seleccionado por la sección de control, y la señal de mezcla descendente según un segundo formato de codificación, que ha sido seleccionado por la sección de control inmediatamente después del primer formato de codificación, y

la reconstrucción paramétrica de la señal de audio de M canales entre los valores discretos debe basarse en valores interpolados de los conjuntos de coeficientes (β_L, γ_L) de mezcla ascendente secos y húmedos según una regla de interpolación predefinida, en donde el sistema de codificación de audio se configura para emitir el desvanecimiento cruzado de la señal de mezcla descendente y los valores discretos de los conjuntos de coeficientes de mezcla ascendente secos y húmedos de tal manera que dicho desvanecimiento cruzado e interpolación serán síncronos.

14. El sistema de codificación de audio de la reivindicación 13, configurado para codificar además una señal de audio de M_2 canales (R, RS, RB, TFR, TBR),

en donde la sección de control se configura para seleccionar repetidamente uno de los formatos de codificación con efecto para la señal de audio de M canales y la señal de audio de M_2 canales,

comprendiendo además el sistema una sección de codificación adicional, que se acopla comunicativamente a la sección de control y se configura para codificar la señal de audio de M_2 canales de acuerdo con el formato de codificación seleccionado por la sección de control.

15. Un producto de programa de ordenador que comprende un medio legible por ordenador con instrucciones que, cuando son ejecutadas por un ordenador, provocan que el ordenador realice el método de cualquiera de las reivindicaciones 1 – 8 y 12.

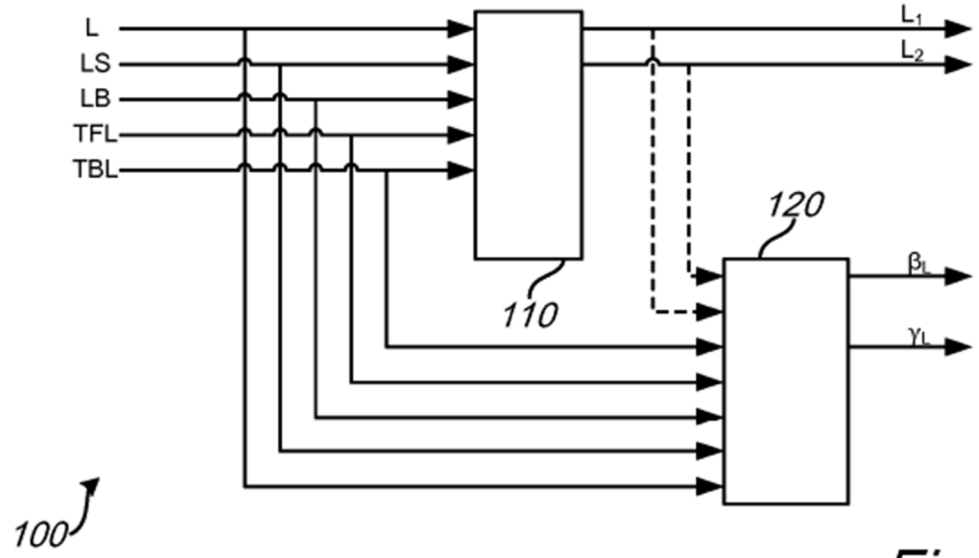
L
LS
LB
TFL
TBL
L1
L2
120
110
βL
γL
100

Fig. 1

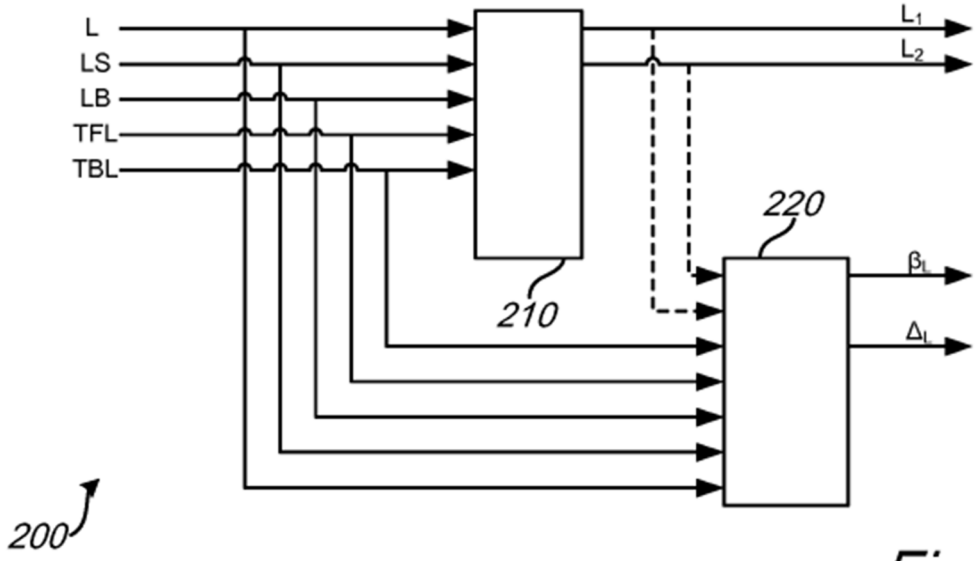
L
LS
LB
TFL
TBL
L1
L2
220
210
βL
ΔL
200

Fig. 2

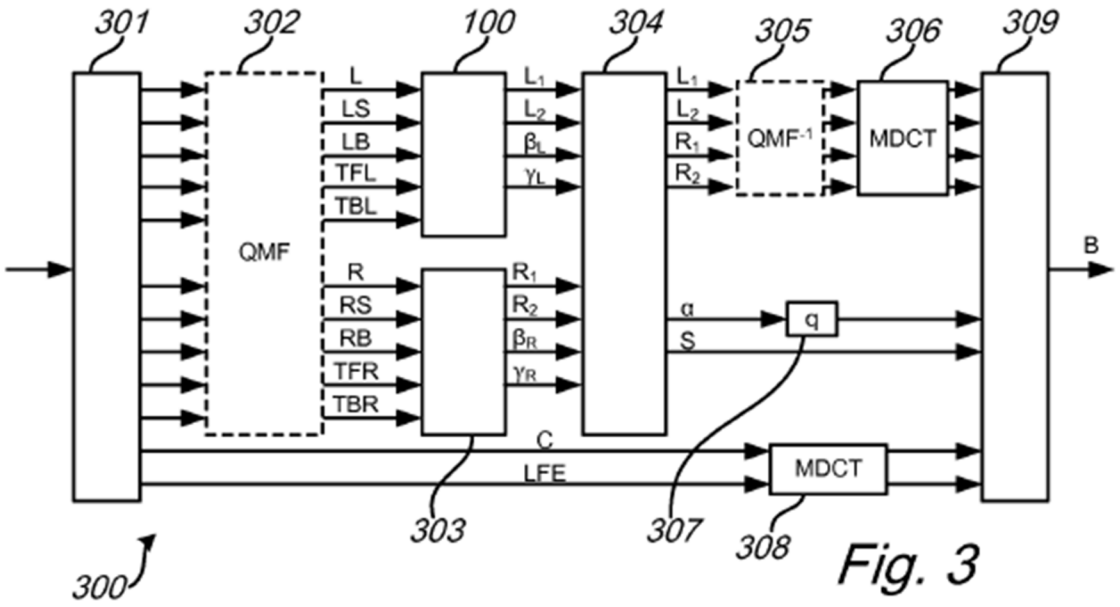

Fig. 3

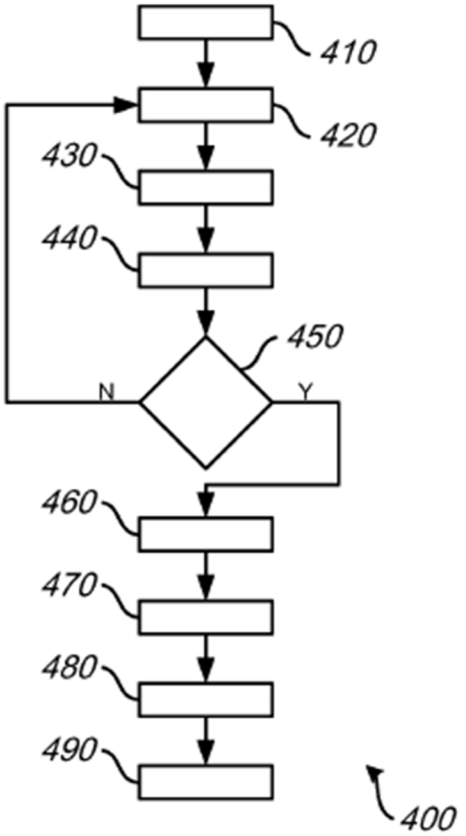

Fig. 4

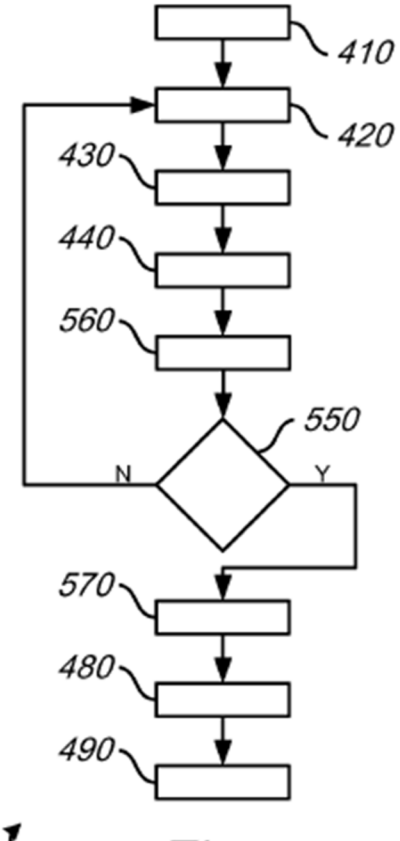

Fig. 5

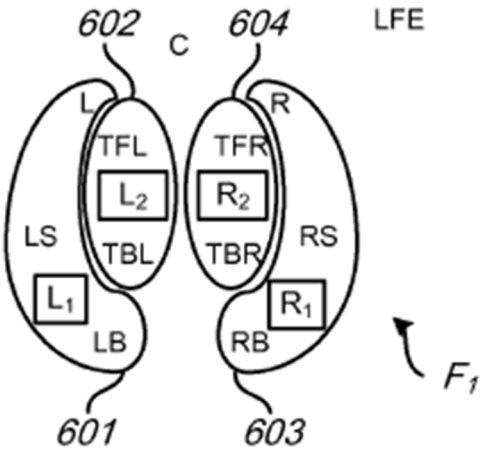

Fig. 6

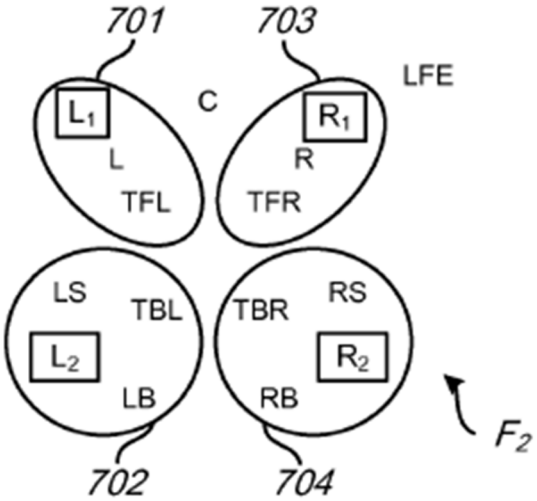

Fig. 7

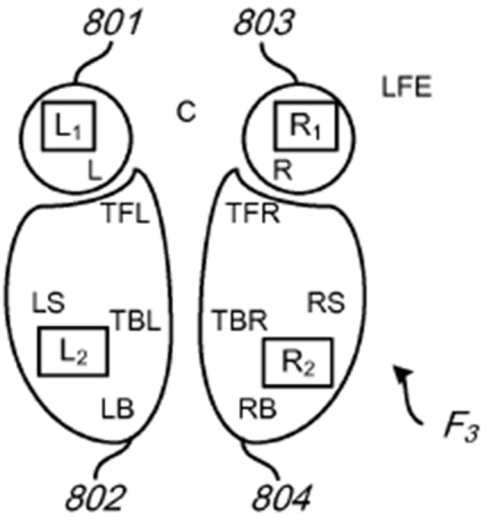

Fig. 8

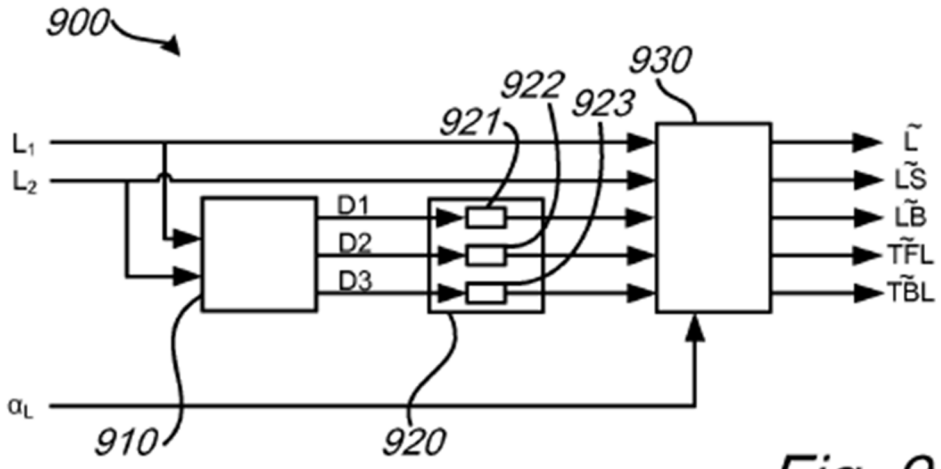
900
930
922
923
921
L1
L2
D1
D2
D3
L̃
L̃S
L̃B
T̃FL
T̃BL
αL
910
920

Fig. 9

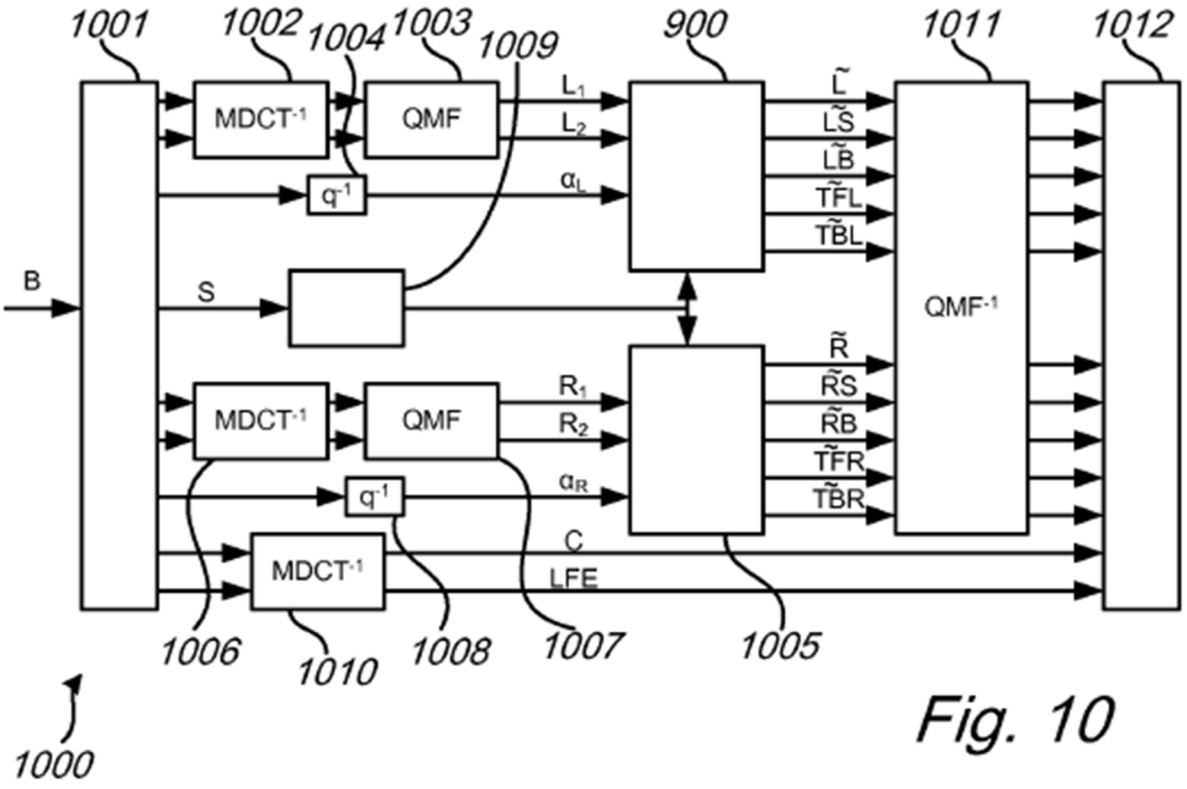
1001
1002
1004
1003
1009
900
1011
1012
MDCT-1
QMF
q-1
L1
L2
αL
L̃
L̃S
L̃B
T̃FL
T̃BL
B
S
QMF-1
R̃
R̃S
R̃B
T̃FR
T̃BR
R1
R2
αR
C
LFE
1006
1010
1008
1007
1005
1000

Fig. 10

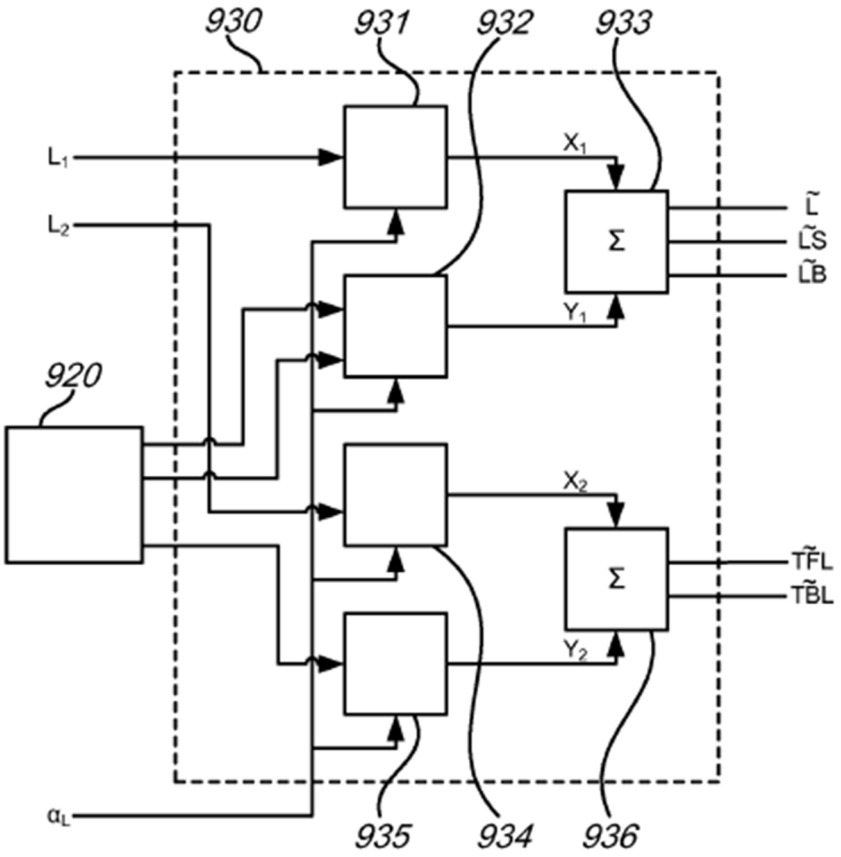
930
931
932
933
L1
L2
X1
Σ
L̃
L̃S
L̃B
Y1
920
X2
Σ
T̃FL
T̃BL
Y2
αL
935
934
936

Fig. 11

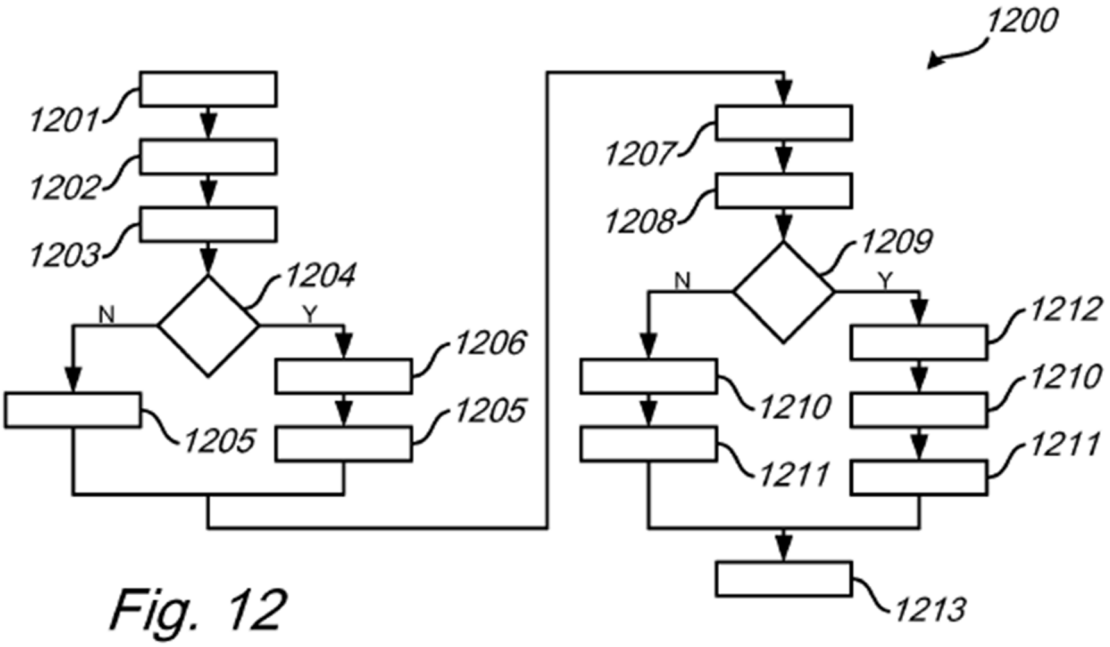
1200
1201
1202
1203
1204
N
Y
1206
1205
1207
1208
1209
N
Y
1212
1210
1211
1210
1211
1213

Fig. 12

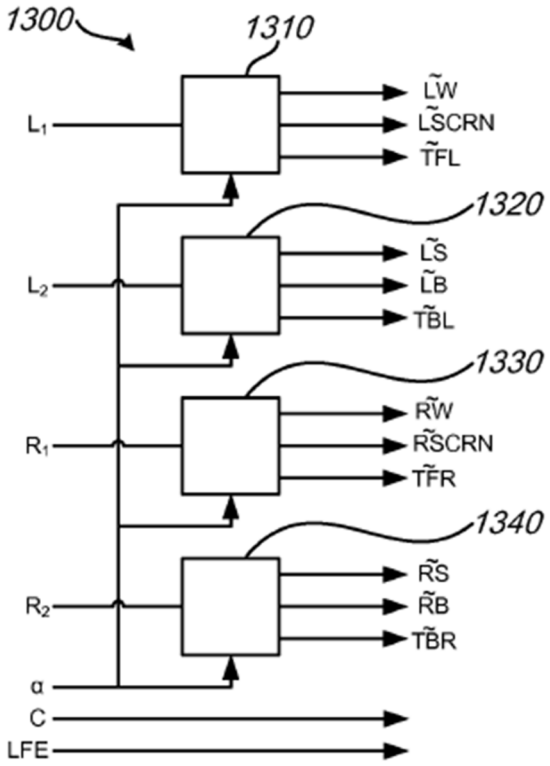

Fig. 13

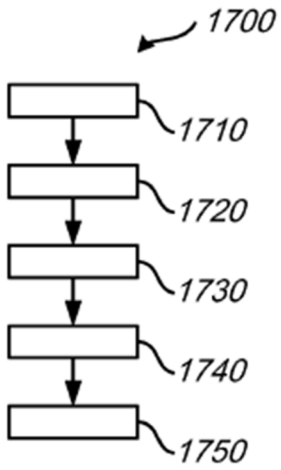

Fig. 17

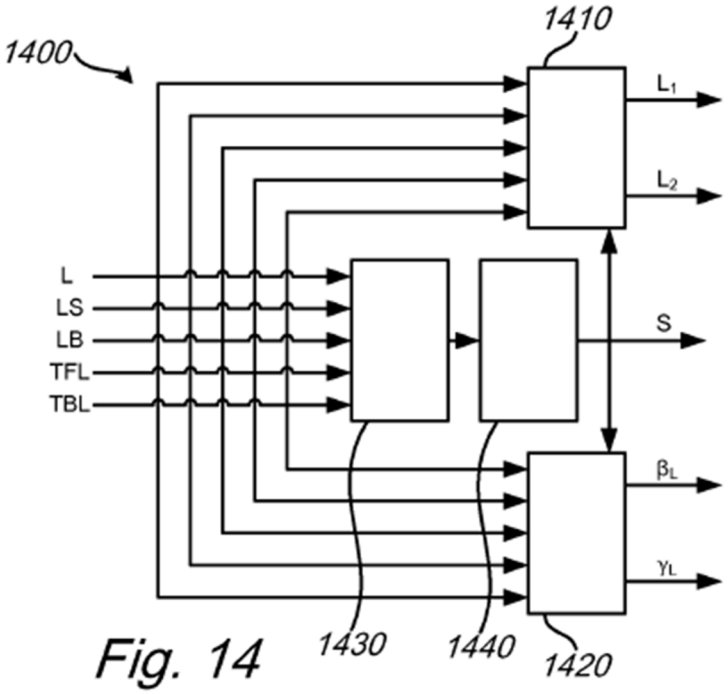

Fig. 14

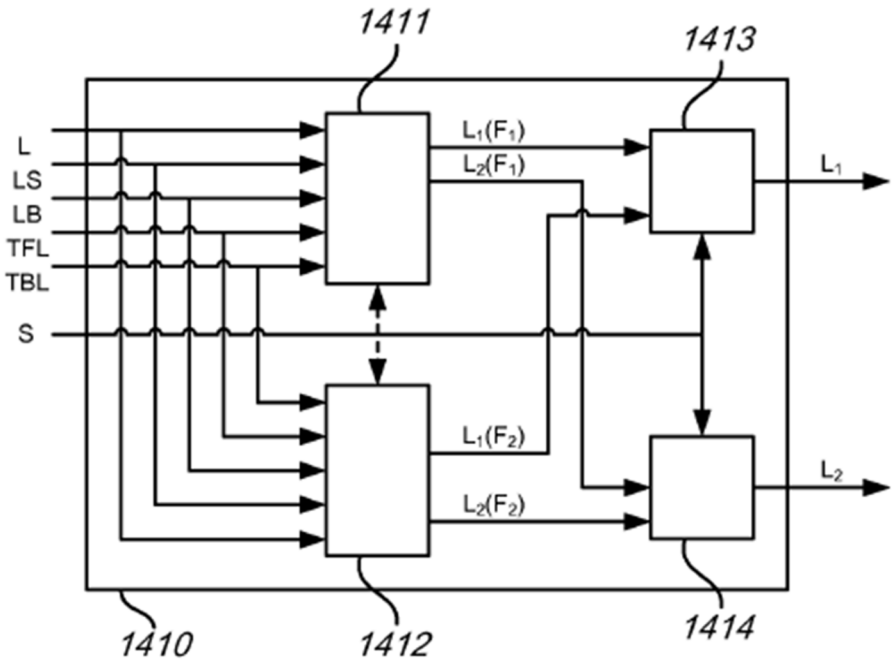
1411
1413
L
LS
LB
TFL
TBL
S
L1(F1)
L2(F1)
L1
L1(F2)
L2(F2)
L2
1410
1412
1414

Fig. 15

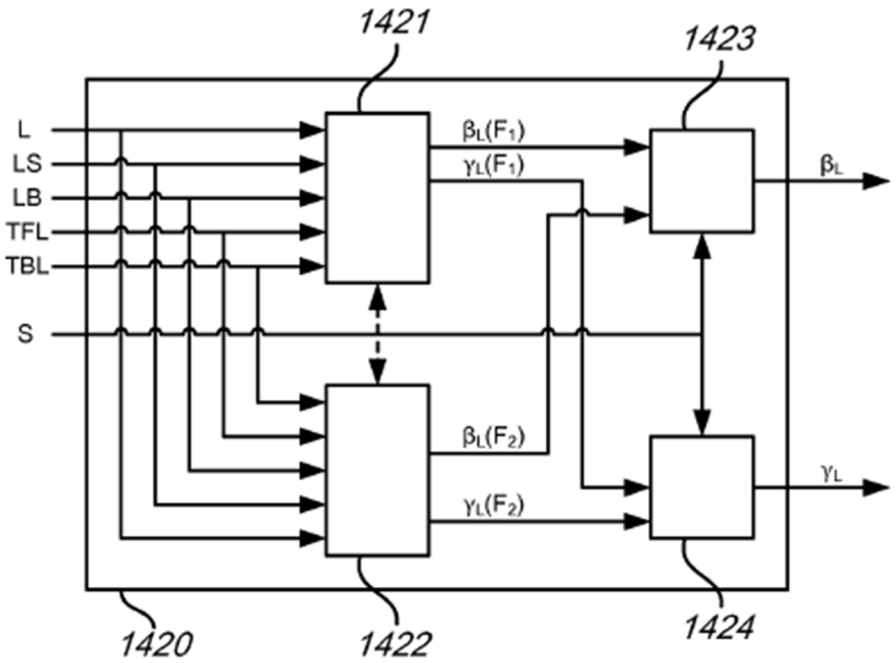
1421
1423
L
LS
LB
TFL
TBL
S
βL(F1)
γL(F1)
βL
βL(F2)
γL(F2)
γL
1420
1422
1424

Fig. 16