

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 660 778**

51 Int. Cl.:

G10L 19/008 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **21.10.2014 PCT/EP2014/072570**

87 Fecha y número de publicación internacional: **30.04.2015 WO15059153**

96 Fecha de presentación y número de la solicitud europea: **21.10.2014 E 14792778 (4)**

97 Fecha y número de publicación de la concesión europea: **17.01.2018 EP 3061089**

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

30 Prioridad:

21.10.2013 US 201361893770 P
03.04.2014 US 201461974544 P
15.08.2014 US 201462037693 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
26.03.2018

73 Titular/es:

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

72 Inventor/es:

VILLEMoes, LARS;
LEHTONEN, HEIDI-MARIA;
PURNHAGEN, HEIKO y
HIRVONEN, TONI

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 660 778 T3

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

DESCRIPCIÓN

Reconstrucción paramétrica de señales de audio

Campo técnico de la invención

5 La invención que se divulga en la presente solicitud está relacionada en general con la codificación y decodificación de señales de audio, y en particular con la reconstrucción paramétrica de una señal de audio multicanal a partir de una señal downmix (mezcla con reducción de canales) y los metadatos asociados.

Antecedentes de la invención

10 Los sistemas de reproducción de audio que comprenden múltiples altavoces se utilizan frecuentemente para reproducir un escenario de audio representado por una señal de audio multicanal, en donde los respectivos canales de la señal de audio multicanal se reproducen en los altavoces respectivos. La señal de audio multicanal puede haber sido grabada, por ejemplo, mediante una pluralidad de transductores acústicos o puede haber sido generada por un equipo de creación de audio. En muchas situaciones existen limitaciones de ancho de banda para transmitir la señal de audio al equipo de reproducción y/o un espacio limitado para almacenar la señal de audio en la memoria de un ordenador o en un dispositivo de almacenamiento portátil. Existen sistemas de codificación de audio para
15 codificación paramétrica de señales de audio, cuya función es reducir el ancho de banda o la cantidad de almacenamiento necesarios. En el codificador, estos sistemas suelen mezclar la señal de audio multicanal en una señal downmix, que normalmente es una señal downmix mono (un canal) o estéreo (dos canales) y extraen la información lateral que describe las propiedades de los canales por medio de parámetros tales como diferencias de nivel y correlación cruzada. A continuación se codifican la señal downmix y la información lateral y se envían a un decodificador. En el decodificador se reconstruye, esto es, se aproxima, la señal de audio multicanal a partir de la señal downmix bajo el control de los parámetros de la información lateral.

20 La solicitud de patente internacional WO 2008/131903 A1 divulga un equipo y un método para sintetizar una señal generada de salida que representa múltiples objetos de audio. En el equipo se genera una señal de descorrelación a partir de una señal downmix multicanal. A continuación, se forma una combinación ponderada de la señal downmix y la señal descorrelacionada a partir de la información paramétrica de los objetos de audio, la información de la señal downmix y la información de generación objetivo.

25 Sin embargo, a la vista de la amplia gama de tipos diferentes de dispositivos y sistemas disponibles para la reproducción de contenido de audio multicanal, incluyendo un segmento emergente dirigido a los usuarios finales en sus hogares, se necesitan formas nuevas y alternativas de codificación eficiente de contenidos de audio multicanal con el fin de reducir los requisitos de ancho de banda y/o la cantidad de memoria requeridos para almacenar y/o facilitar la reconstrucción de la señal de audio multicanal en el decodificador.

Breve descripción de los dibujos

A continuación, se describirán con mayor detalle algunos modos de realización de ejemplo, haciendo referencia a los dibujos adjuntos, en los cuales:

35 la Fig. 1 es un diagrama de bloques generalizado de una sección de reconstrucción paramétrica para reconstruir una señal de audio multicanal a partir de una señal downmix de un solo canal y los parámetros asociados de upmix (mezcla con aumento de canales) seca (dry, original) y húmeda (wet, procesada), de acuerdo con un modo de realización de ejemplo;

40 la Fig. 2 es un diagrama de bloques generalizado de un sistema de decodificación de audio que comprende la sección de reconstrucción paramétrica representada en la Fig. 1, de acuerdo con un modo de realización de ejemplo;

la Fig. 3 es un diagrama de bloques generalizado de una sección de codificación paramétrica para codificar una señal de audio multicanal como una señal downmix de un solo canal junto con los metadatos asociados, de acuerdo con un modo de realización de ejemplo;

45 la Fig. 4 es un diagrama de bloques generalizado de un sistema de codificación de audio que comprende la sección de codificación paramétrica representada en la Fig. 3, de acuerdo con un modo de realización de ejemplo;

las figuras 5-11 ilustran formas alternativas de representación de una señal de audio de 11.1 canales por medio de señales downmix, de acuerdo con los modos de realización de ejemplo;

50 las figuras 12-13 ilustran formas alternativas de representación de una señal de audio de 13.1 canales por medio de señales downmix, de acuerdo con los modos de realización de ejemplo; y

las figuras 14-16 ilustran formas alternativas de representación de una señal de audio de 22.2 canales por medio de señales downmix, de acuerdo con los modos de realización de ejemplo.

Todas las figuras son esquemáticas y en general solo muestran los componentes necesarios para aclarar la invención, en tanto que otros componentes se pueden omitir o simplemente sugerir.

Descripción de los modos de realización de ejemplo

5 Tal como se utiliza en la presente solicitud, una *señal de audio* puede ser una señal de audio pura, la parte de audio de una señal audiovisual o señal multimedia o cualquiera de estas en combinación con unos metadatos.

Tal como se utiliza en la presente solicitud, un *canal* es una señal de audio asociada con una posición/orientación espacial predefinida/fija o una posición espacial indefinida como, por ejemplo, "izquierda" o "derecha".

I. Visión general

10 De acuerdo con un primer aspecto, los modos de realización de ejemplo proponen sistemas de decodificación de audio, así como métodos y productos de programas de ordenador para reconstruir una señal de audio. Los sistemas de decodificación, métodos y productos de programas de ordenador propuestos, de acuerdo con el primer aspecto, pueden compartir generalmente las mismas características y ventajas.

15 De acuerdo con los modos de realización de ejemplo, se proporciona un método para reconstruir una señal de audio de N canales, en donde $N \geq 3$. El método comprende recibir una señal downmix de un solo canal, o un canal de una señal downmix multicanal que transporta datos para la reconstrucción de más señales de audio, junto con los parámetros de upmix seca y húmeda asociados; calcular una primera señal con una pluralidad de (N) canales, denominada señal upmix seca, mediante un mapeo lineal de la señal downmix, en donde, como parte del cálculo de la señal upmix seca, a la señal downmix se le aplica un conjunto de coeficientes de upmix seca; generar una señal descorrelacionada de ($N - 1$) canales a partir de la señal downmix; calcular una señal adicional con una pluralidad de (N) canales, denominada señal upmix húmeda, mediante un mapeo lineal de la señal descorrelacionada, en donde, como parte del cálculo de la señal upmix húmeda, a los canales de la señal descorrelacionada se les aplica un conjunto de coeficientes de upmix húmeda; y combinar las señales de upmix seca y húmeda para obtener una señal reconstruida multidimensional correspondiente a la señal de audio de N canales que se desea reconstruir. El método comprende, además, determinar el conjunto de coeficientes de upmix seca a partir de los parámetros de upmix seca recibidos; rellenar una matriz intermedia que tiene un número de elementos mayor que el número de parámetros de upmix húmeda recibidos, a partir de los parámetros de upmix húmeda recibidos y sabiendo que la matriz intermedia pertenece a una clase de matrices predefinidas; y obtener el conjunto de coeficientes de upmix húmeda multiplicando la matriz intermedia por una matriz predefinida, en donde el conjunto de coeficientes de upmix húmeda corresponde a la matriz resultante de la multiplicación y contiene más coeficientes que el número de elementos de la matriz intermedia.

25 En este modo de realización de ejemplo, el número de coeficientes de upmix húmeda utilizados para reconstruir la señal de audio de N canales es mayor que el número de parámetros de upmix húmeda recibidos. Utilizando el conocimiento de la matriz predefinida y la clase de matrices predefinidas para obtener los coeficientes de upmix húmeda a partir de los parámetros de upmix húmeda recibidos, se puede reducir la cantidad de información necesaria para permitir la reconstrucción de la señal de audio de N canales, lo que permite una reducción de la cantidad de metadatos transmitidos junto con la señal downmix por parte del codificador. Mediante la reducción de 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 N canales, y/o la cantidad de memoria requerida para almacenar dicha representación.

35 La señal descorrelacionada de ($N - 1$) canales sirve para aumentar la dimensionalidad del contenido de la señal de audio de N canales reconstruida, tal como la percibe quien la escucha. Los canales de la señal descorrelacionada de ($N - 1$) canales pueden tener al menos aproximadamente el mismo espectro que la señal downmix de un solo canal, o pueden tener unos espectros correspondientes a versiones reajustadas/normalizadas del espectro de la señal downmix de un solo canal, y pueden formar, junto con la señal downmix de un solo canal, N al menos aproximadamente canales mutuamente incorrelacionados. Con el fin de proporcionar una reconstrucción fiel de los canales de la señal de audio de N canales, cada uno de los canales de la señal descorrelacionada tiene preferiblemente tales propiedades que un oyente lo percibe similar a la señal downmix. Como resultado, aunque es posible sintetizar señales mutuamente incorrelacionadas con un espectro dado a partir de, por ejemplo, ruido blanco, los canales de la señal descorrelacionada se obtienen preferiblemente procesando la señal downmix, por ejemplo, incluyendo la aplicación de filtros paso todo respectivos a la señal downmix, o recombinando porciones de la señal downmix, con el fin de preservar tantas propiedades como sea posible, especialmente las propiedades localmente estacionarias, de la señal downmix, incluyendo propiedades relativamente más sutiles, condicionadas psicoacústicamente, de la señal downmix, tales como el timbre.

40 La combinación de las señales de upmix húmeda y seca puede incluir la incorporación de contenido de audio de los canales respectivos de la señal upmix húmeda al contenido de audio de los respectivos canales correspondientes de la señal upmix seca como, por ejemplo, una mezcla aditiva por muestra o por coeficiente de transformación.

55 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 como, por ejemplo, ciertas relaciones entre algunos de

- los elementos de la matriz, o algunos elementos de la matriz que son cero. El conocimiento de estas propiedades permite rellenar la matriz intermedia sobre la base de un número de parámetros de upmix húmeda menor que la cantidad total de elementos de la matriz intermedia. El 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 a partir de los pocos parámetros de upmix húmeda.
- El que la señal upmix seca sea un mapeo lineal de la señal downmix, significa que la señal upmix seca se obtiene aplicando una primera transformación lineal a la señal downmix. Esta una primera transformación utiliza un canal como entrada y produce N canales como salida, y los coeficientes de upmix seca son coeficientes que definen las propiedades cuantitativas de esta una primera transformación lineal.
- El que la señal upmix húmeda sea un mapeo lineal de la señal descorrelacionada, significa que la señal upmix húmeda se obtiene aplicando una segunda transformación lineal a la señal descorrelacionada. Esta segunda transformación utiliza $N - 1$ canales como entrada y produce N canales como salida, y los coeficientes de upmix húmeda son coeficientes que definen las propiedades cuantitativas de esta segunda transformación lineal.
- En un modo de realización de ejemplo, la recepción de los parámetros de upmix húmeda puede incluir recibir $N(N - 1)/2$ parámetros de upmix húmeda. En el presente modo de realización de ejemplo, rellenar la matriz intermedia puede incluir la obtención de valores para $(N - 1)^2$ elementos de la matriz a partir de los $N(N - 1)/2$ parámetros de upmix húmeda recibidos y sabiendo que la matriz intermedia pertenece a la clase de matrices predefinidas. Esto puede incluir insertar los valores de los parámetros de upmix húmeda directamente como elementos de la matriz, o procesar los parámetros de upmix húmeda de una manera apropiada para obtener valores para los elementos de la matriz. En el presente modo de realización de ejemplo, la matriz predefinida puede incluir $N(N - 1)$ elementos, y el conjunto de coeficientes de upmix húmeda puede incluir $N(N - 1)$ coeficientes. Por ejemplo, la recepción de los parámetros de upmix húmeda puede incluir recibir no más de $N(N - 1)/2$ parámetros de upmix húmeda asignables de forma independiente y/o el número de parámetros de upmix húmeda recibidos puede no ser más de la mitad del número de coeficientes de upmix húmeda utilizados para reconstruir el señal de audio de N canales.
- Se debe entender que omitir la contribución de un canal de la señal descorrelacionada al formar un canal de la señal upmix húmeda mediante un mapeo lineal de los canales de la señal descorrelacionada corresponde a aplicar un coeficiente con el valor cero a ese canal, esto es, omitir la contribución de un canal no afecta al número de coeficientes aplicados como parte del mapeo lineal.
- En un modo de realización de ejemplo, rellenar la matriz intermedia puede incluir la utilización de los parámetros de upmix húmeda recibidos como elementos en la matriz intermedia. Como los parámetros de upmix húmeda recibidos se utilizan como elementos en la matriz intermedia sin ningún procesamiento adicional, se puede reducir la complejidad de los cálculos requeridos para rellenar la matriz intermedia, y obtener los coeficientes de upmix, lo que permite una reconstrucción computacionalmente más eficiente de la señal de audio de N canales.
- En un modo de realización de ejemplo, la recepción de los parámetros de upmix seca puede incluir recibir $(N - 1)$ parámetros de upmix seca. En el presente modo de realización de ejemplo, el conjunto de coeficientes de upmix seca puede incluir N coeficientes, y el conjunto de coeficientes de upmix seca se determina a partir de los $(N - 1)$ parámetros de upmix seca recibidos y de acuerdo con una relación predefinida entre los coeficientes del conjunto de coeficientes de upmix seca. Por ejemplo, la recepción de los parámetros de upmix seca puede incluir recibir no más de $(N - 1)$ parámetros de upmix seca asignables de forma independiente. Por ejemplo, la señal downmix se puede obtener, de acuerdo con una regla predefinida, mediante un mapeo lineal de la señal de audio de N canales que se desea reconstruir, y la relación predefinida entre los coeficientes de upmix seca se puede basar en la regla predefinida.
- En un modo de realización de ejemplo, la clase de matrices predefinidas puede ser una de las siguientes: 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 sean cero; matrices simétricas, en donde las propiedades conocidas de todas las matrices de la clase incluyen que algunos elementos de la matriz predefinida (en cualquiera de los lados de la diagonal principal) sean 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 enumeradas más arriba es que su dimensionalidad es menor que el número total de elementos de la matriz.
- En un modo de realización de ejemplo, la señal downmix se puede obtener, de acuerdo con una regla predefinida, mediante un mapeo lineal de la señal de audio de N canales que se desea reconstruir. En el presente modo de realización de ejemplo, la regla predefinida puede definir una operación de downmix predefinida, y la matriz predefinida se puede basar en vectores que abarcan el espacio del núcleo de la operación de downmix predefinida. Por ejemplo, las filas o columnas de la matriz predefinida pueden ser vectores que forman una base, por ejemplo una base ortonormal, para el espacio del núcleo de la operación de downmix predefinida.

En un modo de realización de ejemplo, la recepción de la señal downmix de un solo canal junto con los parámetros de upmix seca y húmeda asociados puede incluir recibir un segmento de tiempo o un átomo de tiempo/frecuencia de la señal downmix junto con los parámetros de upmix seca y húmeda asociados a ese segmento de tiempo o a ese átomo de tiempo/frecuencia. En el presente modo de realización de ejemplo, la señal multidimensional reconstruida puede corresponder a un segmento de tiempo o un átomo de tiempo/frecuencia de la señal de audio de N canales que se desea reconstruir. En otras palabras, la reconstrucción de la señal de audio de N canales se puede realizar, al menos en algunos modos de realización de ejemplo, en un segmento de tiempo o un átomo de tiempo/frecuencia a la vez. Los sistemas de codificación/decodificación de audio típicamente dividen el espacio de tiempo-frecuencia en átomos de tiempo/frecuencia, por ejemplo aplicando bancos de filtros apropiados a las señales de audio de entrada. Por átomo de tiempo/frecuencia se pretende indicar generalmente una porción del espacio tiempo-frecuencia que corresponde a un intervalo/segmento de tiempo y a una sub-banda de frecuencia.

De acuerdo con los modos de realización de ejemplo, se proporciona un sistema de decodificación de audio que comprende una primera sección de reconstrucción paramétrica configurada para reconstruir una señal de audio de N canales a partir de una primera señal downmix de un solo canal y los parámetros de upmix seca y húmeda asociados, en donde $N \geq 3$. La primera sección de reconstrucción paramétrica comprende una primera sección de descorrelación configurada para recibir la primera señal downmix y generar, en función de la misma, una primera señal descorrelacionada de $N - 1$ canales. La primera sección de reconstrucción paramétrica también comprende una primera sección de upmix seca configurada para: recibir los parámetros de upmix seca y la señal downmix; determinar un primer conjunto de coeficientes de upmix seca a partir de los parámetros de upmix seca; y generar una primera señal upmix seca calculada mediante un mapeo lineal de la primera señal downmix de acuerdo con el primer conjunto de coeficientes de upmix seca. En otras palabras, los canales de la primera señal upmix seca se obtienen multiplicando la señal downmix de un solo canal por los coeficientes respectivos, que pueden ser los propios coeficientes de upmix seca, o que pueden ser coeficientes controlables a través de los coeficientes de upmix seca. La primera sección de reconstrucción paramétrica comprende, además, una primera sección de upmix húmeda configurada para: recibir los parámetros de upmix húmeda y la primera señal descorrelacionada; rellenar una primera matriz intermedia que tiene más elementos que el número de parámetros de upmix húmeda recibidos, a partir de los parámetros de upmix húmeda recibidos y sabiendo que la primera matriz intermedia pertenece a una primera clase de matrices predefinidas, esto es, utilizando las propiedades de ciertos elementos de la matriz que se sabe que son comunes a todas las matrices de la clase de matrices predefinidas; obtener un primer conjunto de coeficientes de upmix húmeda multiplicando la primera matriz intermedia por una primera matriz predefinida, en donde el primer conjunto de coeficientes de upmix húmeda corresponde a la matriz resultante de la multiplicación e incluye más coeficientes que el número de elementos de la primera matriz intermedia; y generar una primera señal upmix húmeda calculada mediante un mapeo lineal de la primera señal descorrelacionada de acuerdo con el primer conjunto de coeficientes de upmix húmeda, esto es, formando combinaciones lineales de los canales de la señal descorrelacionada utilizando los coeficientes de upmix húmeda. La primera sección de reconstrucción paramétrica también comprende una primera sección de combinación configurada para recibir la primera señal upmix seca y la primera señal upmix húmeda y combinar dichas señales para obtener una primera señal multidimensional reconstruida correspondiente a la señal de audio N dimensional que se desea reconstruir.

En un modo de realización de ejemplo, el sistema de decodificación de audio puede comprender, además, una segunda sección de reconstrucción paramétrica que funciona de forma independiente de la primera sección de reconstrucción paramétrica y está configurada para reconstruir una señal de audio de N_2 canales a partir de una segunda señal downmix de un solo canal y los parámetros de upmix seca y húmeda asociados, en donde $N_2 \geq 2$. Por ejemplo, puede suceder que $N_2 = 2$ o que $N_2 \geq 3$. En el presente modo de realización de ejemplo, la segunda sección de reconstrucción paramétrica puede comprender una segunda sección de descorrelación, una segunda sección de upmix seca, una segunda sección de upmix húmeda y una segunda sección de combinación, y las secciones de la segunda sección de reconstrucción paramétrica se pueden configurar de forma análoga a las secciones correspondientes de la primera sección de reconstrucción paramétrica. En el presente modo de realización de ejemplo, la segunda sección de upmix húmeda se puede configurar para utilizar una segunda matriz intermedia que pertenece a una segunda clase de matrices predefinidas y una segunda matriz predefinida. La segunda clase de matrices predefinidas y la segunda matriz predefinida pueden ser diferentes de, o iguales a, la primera clase de matrices predefinidas y la primera matriz predefinida, respectivamente.

En un modo de realización de ejemplo, el sistema de decodificación de audio se puede adaptar para reconstruir una señal de audio multicanal a partir de una pluralidad de canales downmix y los parámetros de upmix seca y húmeda asociados. En el presente modo de realización de ejemplo, el sistema de decodificación de audio puede comprender: una pluralidad de secciones de reconstrucción, que incluyen secciones de reconstrucción paramétrica que se utilizan para reconstruir de forma independiente conjuntos respectivos de canales de señal de audio a partir de los respectivos canales downmix y los respectivos parámetros de upmix seca y húmeda asociados; y una sección de control configurada para recibir una señalización que indica un formato de codificación de la señal de audio multicanal correspondiente a una partición de los canales de la señal de audio multicanal en conjuntos de canales representados por los respectivos canales downmix y, para al menos algunos de los canales downmix, por los respectivos parámetros de upmix seca y húmeda asociados. En el presente modo de realización de ejemplo, el formato de codificación puede corresponder además a un conjunto de matrices predefinidas para la obtención de los coeficientes de upmix húmeda asociados con al menos algunos de los conjuntos de canales respectivos en función

de los parámetros de upmix húmeda respectivos. Opcionalmente, el formato de codificación también puede corresponder a un conjunto de clases de matrices predefinidas indicando cómo se deben rellenar las matrices intermedias respectivas de acuerdo con los conjuntos respectivos de parámetros de upmix húmeda.

5 En el presente modo de realización de ejemplo, el sistema de decodificación se puede configurar para reconstruir la señal de audio multicanal utilizando un primer subconjunto de la pluralidad de secciones de reconstrucción, en respuesta a una señalización recibida que indica un primer formato de codificación. En el presente modo de realización de ejemplo, el sistema de decodificación se puede configurar para reconstruir la señal de audio multicanal utilizando un segundo subconjunto de la pluralidad de secciones de reconstrucción, en respuesta a una señalización recibida que indica un segundo formato de codificación, y al menos uno entre el primer y el segundo subconjuntos de secciones de reconstrucción puede comprender la primera sección de reconstrucción paramétrica.

10 En función de la composición del contenido de audio de la señal de audio multicanal, el ancho de banda disponible para la transmisión desde un codificador a un decodificador, la calidad de reproducción requerida tal como la percibe un oyente y/o la fidelidad requerida de la señal de audio reconstruida por parte del decodificador, el formato de codificación más apropiado puede diferir entre aplicaciones y/o periodos de tiempo diferentes. Mediante el soporte de múltiples formatos de codificación para la señal de audio multicanal, el sistema de decodificación de audio del presente modo de realización de ejemplo permite que un codificador utilice un formato de codificación más apropiado específicamente para las circunstancias actuales.

15 En un modo de realización de ejemplo, la pluralidad de secciones de reconstrucción puede incluir una sección de reconstrucción de un solo canal que se utiliza para reconstruir de forma independiente un canal de audio individual en un canal downmix en el que no se ha codificado más de un canal de audio individual. En el presente modo de realización de ejemplo, al menos uno entre el primer y el segundo subconjuntos de secciones de reconstrucción puede comprender la sección de reconstrucción de un solo canal. Algunos canales de la señal de audio multicanal pueden ser particularmente importantes para la impresión global de la señal de audio multicanal, tal como la percibe un oyente. Mediante la utilización de la sección de reconstrucción de un solo canal para codificar por ejemplo dicho canal por separado en su propio canal downmix, en tanto que otros canales se codifican paraméricamente en conjunto en otros canales downmix, se puede aumentar la fidelidad de la señal de audio multicanal reconstruida. En algunos modos de realización de ejemplo, el contenido de audio de un canal de la señal de audio multicanal puede ser de un tipo diferente que el contenido de audio de los otros canales de la señal de audio multicanal, y se puede aumentar la fidelidad de la señal de audio multicanal reconstruida utilizando un formato de codificación en el que ese canal se codifique por separado en un canal downmix propio.

20 En un modo de realización de ejemplo, el primer formato de codificación puede corresponder a la reconstrucción de la señal de audio multicanal a partir de un número de canales downmix menor que en el segundo formato de codificación. Utilizando un menor número de canales downmix, se puede reducir el ancho de banda requerido para la transmisión desde un codificador a un decodificador. Utilizando un mayor número de canales downmix, se puede aumentar la fidelidad y/o la calidad de audio percibida de la señal de audio multicanal reconstruida.

25 De acuerdo con un segundo aspecto, los modos de realización de ejemplo proponen sistemas de codificación de audio, así como métodos y productos de programas de ordenador para codificar una señal de audio multicanal. Los sistemas de codificación, métodos y productos de programas de ordenador propuestos de acuerdo con el segundo aspecto pueden compartir, en general, las mismas características y ventajas. Además, las ventajas que se han presentado más arriba para las características de los sistemas, métodos y productos de programas de ordenador de decodificación, de acuerdo con el primer aspecto, pueden ser válidas, en general, para las características correspondientes de los sistemas, métodos y productos de programas de ordenador de codificación de acuerdo con el segundo aspecto.

30 De acuerdo con los modos de realización de ejemplo, se proporciona un método para codificar una señal de audio de N canales como una señal downmix de un solo canal junto con los metadatos apropiados para la reconstrucción paramétrica de la señal de audio a partir de la señal downmix y una señal descorrelacionada de $(N - 1)$ canales determinada a partir de la señal downmix, en donde $N \geq 3$. El método comprende: recibir la señal de audio; calcular, de acuerdo con una regla predefinida, la señal downmix de un solo canal mediante un mapeo lineal de la señal de audio; y determinar un conjunto de coeficientes de upmix seca para definir un mapeo lineal de la señal downmix que aproxima la señal de audio, por ejemplo a través de una aproximación del error cuadrático medio asumiendo que para la reconstrucción solo se dispone de la señal downmix. El método comprende, además, determinar una matriz intermedia basada en la diferencia entre la covarianza de la señal de audio recibida y la covarianza de la señal de audio aproximada mediante el mapeo lineal de la señal downmix, en donde la matriz intermedia, cuando se multiplica por una matriz predefinida, corresponde a un conjunto de coeficientes de upmix húmeda que definen un mapeo lineal de la señal descorrelacionada como parte de la reconstrucción paramétrica de la señal de audio, y en donde el conjunto de coeficientes de upmix húmeda contiene más coeficientes que el número de elementos de la matriz intermedia. El método comprende, además, generar la señal downmix junto con los parámetros de upmix seca, a partir de los cuales se puede derivar el conjunto de coeficientes de upmix seca, y los parámetros de upmix húmeda, en donde la matriz intermedia tiene más elementos que el número de parámetros de upmix húmeda generados, y en donde la matriz intermedia está definida unívocamente por los parámetros de upmix húmeda generados siempre que la matriz intermedia pertenezca a una clase de matrices predefinidas.

Una copia de la señal de audio reconstruida de forma paramétrica en un decodificador incluye, a modo de contribución, una señal upmix seca formada mediante el mapeo lineal de la señal downmix y, a modo de contribución adicional, una señal upmix húmeda formada mediante el mapeo lineal de la señal descorrelacionada. El conjunto de coeficientes de upmix seca define el mapeo lineal de la señal downmix, y el conjunto de coeficientes de upmix húmeda define el mapeo lineal de las señales descorrelacionadas. Mediante la generación de los parámetros de upmix húmeda en un número inferior al número de coeficientes de upmix húmeda, y de los que se pueden derivar los coeficientes de upmix húmeda a partir de la matriz predefinida y la clase de matrices predefinidas, se puede reducir la cantidad de información que se le envía al decodificador para permitir la reconstrucción de la señal de audio de N canales. Mediante la reducción de 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 N canales y/o la cantidad de memoria requerida para almacenar dicha representación.

La matriz intermedia se puede determinar basándose en la diferencia entre la covarianza de la señal de audio recibida y la covarianza de la señal de audio aproximada mediante el mapeo lineal de la señal downmix, por ejemplo una covarianza de la señal obtenida mediante el mapeo lineal de la señal descorrelacionada que suplementa la covarianza de la señal de audio aproximada por el mapeo lineal de la señal downmix.

En un modo de realización de ejemplo, la determinación de la matriz intermedia puede incluir determinar la matriz intermedia de tal modo que la covarianza de la señal obtenida mediante el mapeo lineal de la señal descorrelacionada, definida por el conjunto de coeficientes de upmix húmeda aproxime, o coincida sustancialmente con, la diferencia entre la covarianza de la señal de audio recibida y la covarianza de la señal de audio aproximada mediante el mapeo lineal de la señal downmix. En otras palabras, la matriz intermedia se puede determinar de forma que la copia reconstruida de la señal de audio, obtenida como suma de una señal upmix seca resultante del mapeo lineal de la señal downmix y una señal upmix húmeda resultante del mapeo lineal de la señal descorrelacionada restablezca completa, o al menos aproximadamente, la covarianza de la señal de audio recibida.

En un modo de realización de ejemplo, la generación de los parámetros de upmix húmeda puede incluir generar no más de $N(N - 1)/2$ parámetros de upmix húmeda asignables de forma independiente. En el presente modo de realización de ejemplo, la matriz intermedia puede tener $(N - 1)^2$ elementos, y estos pueden quedar definidos de forma unívoca por los parámetros de upmix húmeda generados, siempre que la matriz intermedia pertenezca a la clase de matrices predefinidas. En el presente modo de realización de ejemplo, el conjunto de coeficientes de upmix húmeda puede incluir $N(N - 1)$ coeficientes.

En un modo de realización de ejemplo, el conjunto de coeficientes de upmix seca puede incluir N coeficientes. En los presentes modos de realización de ejemplo, la generación de los parámetros de upmix seca puede incluir la generación de no más de $N - 1$ parámetros de upmix seca, y el conjunto de coeficientes de upmix seca se puede derivar de los $N - 1$ parámetros de upmix seca mediante la regla predefinida.

En un modo de realización de ejemplo, el conjunto determinado de coeficientes de upmix seca puede definir un mapeo lineal de la señal downmix correspondiente a una aproximación del mínimo error cuadrático medio de la señal de audio, esto es, el conjunto determinado de coeficientes de upmix seca puede definir, entre el conjunto de mapeos lineales de la señal downmix, el mapeo lineal que mejor aproxime la señal de audio desde el punto de vista del mínimo cuadrático medio.

De acuerdo con los modos de realización de ejemplo, se proporciona un sistema de codificación de audio que comprende una sección de codificación paramétrica configurada para codificar una señal de audio de N canales como una señal downmix de un solo canal junto con los metadatos apropiados para la reconstrucción paramétrica de la señal de audio a partir de la señal downmix y una señal descorrelacionada de $(N - 1)$ canales determinada a partir de la señal downmix, en donde $N \geq 3$. La sección de codificación paramétrica comprende: una sección de downmix configurada para recibir la señal de audio y calcular, de acuerdo con una regla predefinida, la señal downmix de un solo canal mediante un mapeo lineal de la señal de audio; y una primera sección de análisis configurada para determinar un conjunto de coeficientes de upmix seca para definir un mapeo lineal de la señal downmix que aproxime la señal de audio. La sección de codificación paramétrica comprende, además, una segunda sección de análisis configurada para determinar una matriz intermedia a partir de la diferencia entre la covarianza de la señal de audio recibida y la covarianza de la señal de audio aproximada por el mapeo lineal de la señal downmix, en donde, el producto de la matriz intermedia por una matriz predefinida corresponde a un conjunto de coeficientes de upmix húmeda que definen un mapeo lineal de la señal descorrelacionada como parte de la reconstrucción paramétrica de la señal de audio, en donde el conjunto de coeficientes de upmix húmeda incluye un número de coeficientes mayor que el número de elementos de la matriz intermedia. La sección de codificación paramétrica está configurada, además, para generar la señal downmix junto con los parámetros de upmix seca, a partir de los cuales se puede obtener el conjunto de coeficientes de upmix seca, y los parámetros de upmix húmeda, en donde la matriz intermedia tiene un número de elementos mayor que el número de parámetros de upmix húmeda generados, y en donde la matriz intermedia está definida de forma unívoca por los parámetros de upmix húmeda generados siempre que la matriz intermedia pertenezca a una clase de matrices predefinidas.

En un modo de realización de ejemplo, el sistema de codificación de audio se puede configurar para proporcionar una representación de una señal de audio multicanal en forma de una pluralidad de canales downmix junto con los

parámetros de upmix seca y húmeda asociados. En el presente modo de realización de ejemplo, el sistema de codificación de audio puede comprender: una pluralidad de secciones de codificación, incluyendo secciones de codificación paramétrica que se utilizan para calcular de forma independiente los respectivos canales downmix y los parámetros respectivos de upmix asociados a partir de los respectivos conjuntos de canales de señal de audio. En el presente modo de realización de ejemplo, el sistema de codificación de audio puede comprender, además, una sección de control configurada para determinar un formato de codificación para la señal de audio multicanal correspondiente a una partición de los canales de la señal de audio multicanal en conjuntos de canales para ser representados por los respectivos canales downmix y, para al menos algunos de los canales downmix, por los respectivos parámetros de upmix seca y húmeda asociados. En el presente modo de realización de ejemplo, el formato de codificación puede corresponder además a un conjunto de reglas predefinidas para calcular al menos algunos de los respectivos canales downmix. En el presente modo de realización de ejemplo, el sistema de codificación de audio se puede configurar para codificar la señal de audio multicanal utilizando un primer subconjunto de la pluralidad de secciones de codificación si el formato de codificación determinado es un primer formato de codificación. En el presente modo de realización de ejemplo, el sistema de codificación de audio se puede configurar para codificar la señal de audio multicanal utilizando un segundo subconjunto de la pluralidad de secciones de codificación si el formato de codificación determinado es un segundo formato de codificación, y al menos uno entre el primer y segundo subconjuntos de secciones de codificación puede comprender la primera sección de codificación paramétrica. En el presente modo de realización de ejemplo, la sección de control puede determinar, por ejemplo, el formato de codificación en función del ancho de banda disponible para transmitir una versión codificada de la señal de audio multicanal a un decodificador, de acuerdo con el contenido de audio de los canales de la señal de audio multicanal y/o una señal de entrada que indica el formato de codificación deseado.

En un modo de realización de ejemplo, la pluralidad de secciones de codificación puede incluir una sección de codificación de un solo canal que se utiliza para codificar de forma independiente no más de un canal de audio individual en un canal downmix, y al menos uno entre el primer y segundo subconjuntos de las secciones de codificación puede comprender la sección de codificación de un solo canal.

De acuerdo con los modos de realización de ejemplo, se proporciona un producto de programa para ordenador que comprende un medio legible por ordenador con instrucciones para ejecutar cualquiera de los métodos del primer y segundo aspectos.

De acuerdo con los modos de realización de ejemplo, se puede asumir que $N = 3$ o $N = 4$ en cualquiera de los métodos, sistemas de codificación, sistemas de decodificación y productos de programa para ordenador del primer y segundo aspectos.

En las reivindicaciones dependientes se definen algunos modos de realización de ejemplo adicionales. Se debe observar que los modos de realización de ejemplo incluyen todas las combinaciones de características, incluso si se detallan en reivindicaciones mutuamente diferentes.

II. Modos de realización de ejemplo

En un codificador, que se describirá haciendo referencia a las Fig. 3 y 4, se calcula una señal Y downmix de un solo canal como un mapeo lineal de una señal de audio $X = [x_1 \dots x_N]^T$ de N canales de acuerdo con

$$Y = [d_1 \dots d_N] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \sum_{n=1}^N d_n x_n = DX, \quad (1)$$

en donde d_n , $n = 1, \dots, N$, son coeficientes downmix representados por una matriz D de downmix. En un decodificador, que se describirá haciendo referencia a las Fig. 1 y 2, se realiza la reconstrucción paramétrica de la señal X de audio de N canales de acuerdo con

$$\hat{X} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} Y + \begin{bmatrix} p_{11} & \dots & p_{1,N-1} \\ p_{21} & \dots & p_{2,N-1} \\ \vdots & \ddots & \vdots \\ p_{N,1} & \dots & p_{N,N-1} \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_{N-1} \end{bmatrix} = CY + PZ, \quad (2)$$

en donde c_n , $n = 1, \dots, N$, son los coeficientes de upmix seca representados por una matriz C de upmix seca, $p_{n,k}$, $n = 1, \dots, N$, $k = 1, \dots, N - 1$, son los coeficientes de upmix húmeda representados por una matriz P de upmix húmeda, y z_k , $k = 1, \dots, N - 1$ son los canales de una señal descorrelacionada Z de $(N - 1)$ canales generada a partir de la señal Y downmix. Si los canales de cada una de las señales de audio se representan como filas, la matriz de covarianza de la señal X de audio original se puede expresar como $R = XX^T$, y la matriz de covarianza de la señal \hat{X} de audio reconstruida se puede expresar como $\hat{R} = \hat{X}\hat{X}^T$. Se debe observar que si, por ejemplo, las señales de audio se representan como filas que comprenden coeficientes de transformación con valores complejos, en lugar de XX^T se

puede considerar, por ejemplo, la parte real de XX^* , en donde X^* es la transpuesta conjugada compleja de la matriz X .

Para proporcionar una reconstrucción fiel de la señal X de audio original, para la reconstrucción expresada por la ecuación (2) puede ser ventajoso restablecer la covarianza total, esto es, puede ser ventajoso utilizar las matrices C y P de upmix seca y húmeda, de modo que

$$R = \widehat{R}. \quad (3)$$

Un enfoque consiste en encontrar en primer lugar una matriz C de upmix seca que produzca la mejor upmix posible "seca" $\widehat{X}_0 = CY$ desde el punto de vista de mínimos cuadrados, resolviendo las ecuaciones normales

$$CYY^T = XY^T. \quad (4)$$

10 Para $\widehat{X}_0 = CY$, con una matriz C resolviendo la ecuación (4), se cumple que

$$R = \widehat{X}_0 \widehat{X}_0^T + (\widehat{X}_0 - X)(\widehat{X}_0 - X)^T = R_0 + \Delta R. \quad (5)$$

Suponiendo que los canales de la señal Z descorrelacionada están mutuamente incorrelacionados y todos tienen la misma energía $\|Y\|^2$ igual a la de la señal Y downmix de un solo canal, la falta de covarianza definida positiva ΔR se puede descomponer en factores de acuerdo con

$$15 \quad \Delta R = PP^T \|Y\|^2. \quad (6)$$

La covarianza completa se puede restablecer de acuerdo con la ecuación (3) utilizando una matriz C de upmix seca que resuelve la ecuación (4) y una matriz P de upmix húmeda que resuelve la ecuación (6). Las ecuaciones (1) y (4) implican que $DCYY^T = YY^T$, y por tanto que

$$\sum_{n=1}^N d_n c_n = DC = 1, \quad (7)$$

20 para las matrices D de downmix no degeneradas. Las ecuaciones (5) y (7) implican que $D(X_0 - X) = DCY - Y = 0$ y

$$D\Delta R = 0. \quad (8)$$

En consecuencia, la falta de covarianza ΔR tiene un rango $N - 1$, y ciertamente se puede proporcionar utilizando una señal Z descorrelacionada con $N - 1$ canales mutuamente incorrelacionados.

25 Las ecuaciones (6) y (8) implican que $DP = 0$, de modo que las columnas de la matriz P de upmix húmeda que resuelve la ecuación (6) se pueden construir a partir de vectores que abarcan el espacio del núcleo de la matriz D de downmix. Los cálculos para encontrar una matriz P de upmix húmeda apropiada se pueden trasladar en consecuencia a ese espacio de menor dimensión.

30 Sea V una matriz de tamaño $N(N - 1)$ que contiene una base ortonormal para el espacio del núcleo de la matriz D de downmix, esto es, un espacio lineal de vectores v con $Dv = 0$. Ejemplos de tales matrices V predefinidas para $N = 2$, $N = 3$, y $N = 4$, respectivamente, son

$$\frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} \\ 0 & -2/\sqrt{6} \\ -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix} \quad \text{y} \quad \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix}. \quad (9)$$

35 En la base proporcionada por V , la falta de covarianza se puede expresar como $R_v = V^T(\Delta R)V$. Para encontrar una matriz P de upmix húmeda que resuelva la ecuación (6), se puede encontrar por lo tanto en primer lugar una matriz H resolviendo $R_v = HH^T$ y, a continuación, obtener P como $P = VH/\|Y\|$, en donde $\|Y\|$ es la raíz cuadrada de la energía de la señal Y downmix de un solo canal. Se pueden obtener otras matrices P de upmix apropiadas como $P = VH0/\|Y\|$, en donde 0 es una matriz ortogonal. Alternativamente, se puede reajustar la falta de covarianza R_v mediante la energía $\|Y\|^2$ de la señal Y downmix de un solo canal y en su lugar resolver la ecuación

$$\frac{R_v}{\|Y\|^2} = H_R H_R^T, \quad (10)$$

en donde $H = H_R/\|Y\|$, y obtener P como

$$P = VH_R. \quad (11)$$

Cuando se cuantifican las entradas de H_R y el resultado deseado tiene un canal en silencio, las propiedades de la matriz V predefinida como se han indicado más arriba pueden no ser convenientes. A modo de ejemplo, para $N = 3$, una opción mejor para la segunda matriz de (9) sería

$$\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & -1/\sqrt{2} \\ -1/\sqrt{2} & 0 \end{bmatrix}. \quad (12)$$

5 Afortunadamente, se puede descartar el requisito de que las columnas de la matriz V sean ortogonales dos a dos siempre que dichas columnas sean linealmente independientes. A continuación, se obtiene la solución deseada R_v para $\Delta R = VR_vV^T$ mediante $R_v = W^T(\Delta R)W$ con $W = V(V^T V)^{-1}$, la pseudoinversa de V .

10 La matriz R_v es una matriz semidefinida positiva de tamaño $(N - 1)^2$ y hay varios enfoques para encontrar soluciones a la ecuación (10), que dan lugar a soluciones dentro de las respectivas clases de matrices de dimensión $N(N - 1)/2$, esto es, en las que las matrices están definidas de forma unívoca por $N(N - 1)/2$ elementos. Las soluciones se pueden obtener, por ejemplo, mediante la utilización de:

- a. Factorización de Cholesky, lo que da lugar a una triangular inferior H_R ;
- b. Raíz cuadrada positiva lo que da lugar a una semidefinida positiva simétrica H_R ; o
- c. Polar, lo que da lugar a una H_N de la forma $H_R = O\Lambda$, donde O es ortogonal y Λ es diagonal.

15 Además, existe una versión normalizada de las opciones a) y b) en las que H_R se puede expresar como $H_R = \Lambda H_0$, en donde Λ es diagonal y H_0 tiene todos los elementos de la diagonal iguales a uno. Las alternativas a, b y c indicadas más arriba proporcionan soluciones H_R en diferentes clases de matrices, esto es, matrices triangulares inferiores, matrices simétricas y productos de matrices diagonales y ortogonales. Si la clase de matrices a la cual pertenece H_R es conocida por parte del decodificador, esto es, si se sabe que H_R pertenece a una clase de matrices predefinidas, por ejemplo de acuerdo con una cualquiera de las alternativas a, b y c indicadas más arriba, H_R se puede rellenar en solo $N(N - 1)/2$ de sus elementos. Si la matriz V también es conocida por parte del decodificador, por ejemplo si se sabe que V es una de las matrices que aparecen en (9), la matriz P de upmix húmeda, necesaria para la reconstrucción de acuerdo con la ecuación (2), se puede obtener a continuación mediante la ecuación (11).

25 La Fig. 3 es un diagrama de bloques generalizado de una sección 300 de codificación paramétrica de acuerdo con un modo de realización de ejemplo. La sección 300 de codificación paramétrica está configurada para codificar una señal X de audio de N canales como una señal Y downmix de un solo canal junto con los metadatos apropiados para la reconstrucción paramétrica de la señal X de audio de acuerdo con la ecuación (2). La sección 300 de codificación paramétrica comprende una sección 301 de downmix, que recibe la señal X de audio y calcula, de acuerdo con una regla predefinida, la señal Y downmix de un solo canal mediante un mapeo lineal de la señal X de audio. En el presente modo de realización de ejemplo, la sección 301 de downmix calcula la señal Y downmix de acuerdo con la ecuación (1), en donde la matriz D de downmix está predefinida y corresponde a la regla predefinida. Una primera sección 302 de análisis determina un conjunto de coeficientes de upmix seca, representados por la matriz C de upmix seca con el fin de definir un mapeo lineal de la señal Y downmix que aproxime la señal X de audio. Este mapeo lineal de la señal Y downmix se denota por CY en la ecuación (2). En el presente modo de realización de ejemplo, los N coeficientes C de upmix seca se determinan de acuerdo con la ecuación (4), de forma que el mapeo lineal CY de la señal Y downmix corresponde a una aproximación del mínimo cuadrático medio de la señal X de audio. Una segunda sección 303 de análisis determina una matriz intermedia H_R basada en la diferencia entre la matriz de covarianza de la señal X de audio recibida y la matriz de covarianza de la señal de audio aproximada por el mapeo lineal CY de la señal Y downmix. En el presente modo de realización de ejemplo, las matrices de covarianza se calculan mediante una primera y segunda secciones 304, 305 de procesamiento, respectivamente y, a continuación, se entregan a la segunda sección 303 de análisis. En el presente modo de realización de ejemplo, la matriz intermedia H_R se determina de acuerdo con el enfoque b descrito más arriba para resolver la ecuación (10), lo que da lugar a una matriz intermedia H_R que es simétrica. Tal como se ha indicado en las ecuaciones (1) y (11), la matriz intermedia H_R , al multiplicarse por una matriz V predefinida, define, mediante un conjunto de parámetros P de upmix húmeda, un mapeo lineal PZ de una señal Z descorrelacionada como parte de la reconstrucción paramétrica de la señal X de audio por parte del decodificador. En el presente modo de realización de ejemplo, la matriz intermedia V es la segunda matriz en (9) para el caso $N = 3$, y la tercera matriz en (9) para el caso $N = 4$. La sección 300 de codificación paramétrica genera la señal Y downmix junto con los parámetros \tilde{C} de upmix seca y los parámetros \tilde{P} de upmix húmeda. En el presente modo de realización de ejemplo, $N - 1$ de los N coeficientes C de upmix seca son los parámetros \tilde{C} de upmix seca, y el coeficiente de upmix seca restante se puede derivar de los parámetros \tilde{C} de upmix seca mediante la ecuación (7) si la matriz D de downmix predefinida es conocida. Como la matriz intermedia H_R pertenece a la clase de matrices simétricas, está definida de forma unívoca por $N(N - 1)/2$ de sus $(N - 1)^2$ elementos. En el presente modo de realización de ejemplo, $N(N - 1)/2$ de los elementos

de la matriz intermedia H_R son, por lo tanto, parámetros \tilde{P} de upmix húmeda a partir de los cuales se puede derivar el resto de la matriz intermedia H_R sabiendo que es simétrica.

La Fig. 4 es un diagrama de bloques generalizado de un sistema 400 de codificación de audio de acuerdo con un modo de realización de ejemplo, que comprende la sección 300 de codificación paramétrica descrita en relación con la Fig. 3. En el presente modo de realización de ejemplo, se proporciona un contenido de audio, por ejemplo grabado por uno o más transductores acústicos 401, o generado por un equipo 401 de producción de audio, en forma de una señal X de audio de N canales. Una sección 402 de análisis de filtro espejo en cuadratura (QMF) transforma la señal X de audio, segmento a segmento de tiempo, en un dominio QMF para su procesamiento por parte de la sección 300 de codificación paramétrica de la señal X de audio en forma de átomos de tiempo/frecuencia. La señal Y downmix generada por la sección 300 de codificación paramétrica se vuelve a transformar desde el dominio de QMF mediante una sección 403 de síntesis QMF y se transforma en un dominio de la transformada discreta del coseno modificada (MDCT) mediante una sección 404 de transformación. Las secciones 405 y 406 de cuantificación cuantifican los parámetros \tilde{C} de upmix seca y los parámetros \tilde{P} de upmix húmeda, respectivamente. Por ejemplo, se puede utilizar una cuantificación uniforme con un tamaño de paso de 0,1 ó 0,2 (adimensional), seguida por una codificación entrópica en forma de codificación de Huffman. Se puede utilizar una cuantificación más gruesa con un tamaño de paso de 0,2, por ejemplo, para economizar ancho de banda de transmisión, y se puede utilizar una cuantificación más fina con un tamaño de paso de 0,1, por ejemplo, para mejorar la fidelidad de la reconstrucción por parte del decodificador. A continuación, se combinan la señal Y downmix transformada mediante MDCT y los parámetros \tilde{C} de upmix seca y los parámetros \tilde{P} de upmix húmeda cuantificados en un flujo B de bits mediante un multiplexor 407, para su transmisión a un decodificador. El sistema 400 de codificación de audio también puede comprender un codificador central (en la Fig. 4 no se muestra) configurado para codificar la señal Y downmix utilizando un codec de audio perceptual como, por ejemplo, Dolby Digital o MPEG AAC, antes de entregarle la señal Y downmix al multiplexor 407.

La Fig. 1 es un diagrama de bloques generalizado de una sección 100 de reconstrucción paramétrica de acuerdo con un modo de realización de ejemplo, configurada para reconstruir la señal X de audio de N canales a partir de una señal Y downmix de un solo canal y los parámetros \tilde{C} de upmix seca y los parámetros \tilde{P} de upmix húmeda asociados. La sección 100 de reconstrucción paramétrica está adaptada para realizar la reconstrucción de acuerdo con la ecuación (2), esto es, utilizando los parámetros C de upmix seca y los parámetros P de upmix húmeda. Sin embargo, en lugar de recibir los propios parámetros C de upmix seca y los parámetros P de upmix húmeda se reciben los parámetros \tilde{C} de upmix seca y los parámetros \tilde{P} de upmix húmeda de los cuales se pueden derivar los parámetros C de upmix seca y los parámetros P de upmix húmeda. Una sección 101 de descorrelación recibe la señal Y downmix y genera, a partir de la misma, una señal $Z = [z_1 \dots z_{N-1}]^T$ de $(N - 1)$ canales. En el presente modo de realización de ejemplo, los canales de la señal Z descorrelacionada se derivan procesando la señal Y downmix, incluyendo la aplicación de filtros paso todo a la señal Y downmix, con el fin de proporcionar canales que están incorrelacionados con la señal Y downmix, y con un contenido de audio que es espectralmente similar, y también percibido por un oyente como similar, al de la señal Y downmix. La señal Z descorrelacionada de $(N - 1)$ canales sirve para aumentar la dimensionalidad de la versión reconstruida \hat{X} de la señal X de audio de N canales, tal como la percibe un oyente. En el presente modo de realización de ejemplo, los canales de la señal Z descorrelacionada tienen al menos aproximadamente el mismo espectro que el de la señal Y downmix de un solo canal y forman, junto con la señal Y downmix de un solo canal, N canales al menos aproximadamente mutuamente incorrelacionados. Una sección 102 de upmix seca recibe los parámetros \tilde{C} de upmix seca y la señal Y downmix. En el presente modo de realización de ejemplo, los parámetros \tilde{C} de upmix seca coinciden con los primeros $N - 1$ de los N coeficientes C de upmix seca, y el coeficiente de upmix seca restante se determina a partir de una relación predefinida entre los coeficientes C de upmix seca dada por la ecuación (7). La sección 102 de upmix seca genera una señal upmix seca calculada mediante un mapeo lineal de la señal Y downmix de acuerdo con el conjunto de coeficientes C de upmix seca, y denotado por CY en la ecuación (2). Una sección 103 de upmix húmeda recibe los parámetros \tilde{P} de upmix húmeda y la señal Z descorrelacionada. En el presente modo de realización de ejemplo, los parámetros \tilde{P} de upmix húmeda son $N(N - 1)/2$ elementos de la matriz intermedia H_R determinada por parte del codificador de acuerdo con la ecuación (10). En el presente modo de realización de ejemplo, la sección 103 de upmix húmeda rellena los elementos restantes de la matriz intermedia H_R sabiendo que la matriz intermedia H_R pertenece a una clase de matrices predefinidas, esto es, que es simétrica, y explotando las relaciones correspondientes entre los elementos de la matriz. La sección 103 de upmix húmeda obtiene a continuación un conjunto de coeficientes P de upmix húmeda utilizando la ecuación (11), esto es, multiplicando la matriz intermedia H_R por la matriz V predefinida, esto es, la segunda matriz en (9) para el caso $N = 3$, y la tercera matriz en (9) para el caso $N = 4$. En consecuencia, los $N(N - 1)$ coeficientes P de upmix húmeda se derivan de los $N(N - 1)/2$ parámetros \tilde{P} de upmix húmeda asignables de forma independiente recibidos. La sección 103 de upmix húmeda genera una señal upmix húmeda calculada mediante un mapeo lineal de la señal Z descorrelacionada de acuerdo con el conjunto de coeficientes P de upmix húmeda, y denotado por PZ en la ecuación (2). Una sección 104 de combinación recibe la señal CY de upmix seca y la señal PZ de upmix húmeda y combina estas señales para obtener una primera señal \hat{X} reconstruida multidimensional correspondiente a la señal X de audio de N canales que se desea reconstruir. En el presente modo de realización de ejemplo, la sección 104 de combinación obtiene los canales respectivos de la señal \hat{X} reconstruida combinando el contenido de audio de los respectivos canales de la

señal CY de upmix seca con los canales respectivos de la señal PZ de upmix húmeda, de acuerdo con la ecuación (2).

La Fig. 2 es un diagrama de bloques generalizado de un sistema 200 de decodificación de audio de acuerdo con un modo de realización de ejemplo. El sistema 200 de decodificación de audio comprende la sección 100 de reconstrucción paramétrica descrita en relación con la Fig. 1. Una sección 201 de recepción que incluye, por ejemplo, un demultiplexor, recibe el flujo B de bits transmitido desde el sistema 400 de codificación de audio descrito en relación con la Fig. 4, y extrae la señal Y downmix y los parámetros \tilde{C} de upmix seca y los parámetros \tilde{P} de upmix húmeda asociados a partir del flujo B de bits. En caso de que la señal Y downmix esté codificada en el flujo B de bits utilizando un codec de audio perceptual como Dolby Digital o MPEG AAC, el sistema 200 de decodificación de audio puede comprender un decodificador central (en la Fig. 2 no se muestra) configurado para decodificar la señal Y downmix al extraerla del flujo B de bits. Una sección 202 de transformación transforma la señal Y downmix realizando una MDCT inversa, y una sección 203 de análisis QMF transforma la señal Y downmix en un dominio QMF para su procesamiento por la sección 100 de reconstrucción paramétrica de la señal Y downmix en forma de átomos de tiempo/frecuencia. Unas secciones 204 y 205 de descuantificación descuantifican los parámetros \tilde{C} de upmix seca y los parámetros \tilde{P} de upmix húmeda, por ejemplo, a partir de un formato de codificación entrópica, antes de entregárselos a la sección 100 de reconstrucción paramétrica. Tal como se ha descrito en relación con la Fig. 4, la cuantificación se puede haber realizado con uno de dos tamaños de paso diferentes, por ejemplo 0,1 ó 0,2. El tamaño real de paso utilizado puede estar predefinido, o se le puede señalar al sistema 200 de decodificación de audio desde el codificador, por ejemplo, a través del flujo B de bits. En algunos modos de realización de ejemplo, los coeficientes C de upmix seca y los coeficientes P de upmix húmeda se pueden derivar a partir de los parámetros \tilde{C} de upmix seca y los parámetros \tilde{P} de upmix húmeda, respectivamente, ya en las respectivas secciones 204 y 205 de descuantificación, las cuales pueden considerarse opcionalmente como parte de la sección 102 de upmix seca y la sección 103 de upmix húmeda, respectivamente. En el presente modo de realización de ejemplo, la señal \hat{X} de audio reconstruida generada por la sección 100 de reconstrucción paramétrica se vuelve a transformar desde el dominio de QMF mediante una sección 206 de síntesis QMF antes de ser enviada como salida del sistema 200 de decodificación de audio para su reproducción en un sistema 207 con múltiples altavoces.

Las Fig. 5-11 ilustran formas alternativas de representación de una señal de audio de 11.1 canales por medio de canales downmix, de acuerdo con los modos de realización de ejemplo. En los presentes modos de realización de ejemplo, la señal de audio de 11.1 canales comprende los canales: izquierdo (L), derecho (R), central (C), efectos de baja frecuencia (LFE), lateral izquierdo (LS), lateral derecho (RS), izquierdo posterior (LB), derecho posterior (RB), izquierdo frontal superior (TFL), derecho frontal superior (TFR), izquierdo posterior superior (TBL) y derecho posterior superior (TBR), que se indican en las Fig. 5-11 mediante letras mayúscula. Las formas alternativas de representación de la señal de audio de 11.1 canales corresponden a particiones alternativas de los canales en conjuntos de canales, estando representado cada conjunto por una sola señal downmix, y opcionalmente por los parámetros de upmix húmeda y seca asociados. La codificación de cada uno de los conjuntos de canales en su respectiva señal downmix de un solo canal (y metadatos) se puede realizar de forma independiente y en paralelo. De modo análogo, la reconstrucción de los respectivos conjuntos de canales a partir de sus respectivas señales downmix de un solo canal se puede realizar de forma independiente y en paralelo.

Se debe entender que, en los modos de realización de ejemplo descritos en relación con las Fig. 5-11 (y también más abajo en relación con las Fig. 13-16), ninguno de los canales reconstruidos puede comprender contribuciones de más de un canal downmix y cualesquiera señales descorrelacionadas derivadas a partir de esa única señal downmix, esto es, contribuciones de múltiples canales downmix, no se combinan/mezclan durante la reconstrucción paramétrica.

En la Fig. 5, los canales LS, TBL y LB forman un grupo 501 de canales representados por el canal downmix individual ls (y sus metadatos asociados). La sección 300 de codificación paramétrica descrita en relación con la Fig. 3 se puede utilizar con $N = 3$ para representar los tres canales de audio LS, TBL y LB por el canal downmix individual ls y los parámetros de upmix seca y húmeda asociados. Como por parte del decodificador son conocidas una matriz V predefinida y una clase de matrices predefinidas de una matriz intermedia H_R , ambas asociadas con la codificación realizada en la sección 300 de codificación paramétrica, se puede utilizar la sección 100 de reconstrucción paramétrica, descrita en relación con la Fig. 1, para reconstruir los tres canales LS, TBL y LB a partir de la señal ls downmix y los parámetros de upmix seca y húmeda asociados. De modo análogo, los canales RS, TBR y RB forman un grupo 502 de canales representados por el canal rs downmix individual, y se puede utilizar otra instancia de la sección 300 de codificación paramétrica en paralelo con la primera sección de codificación para representar los tres canales RS, TBR y RB por el canal rs downmix individual y los parámetros de upmix seca y húmeda asociados. Es más, como por parte del decodificador son conocidas una matriz V predefinida y una clase de matrices predefinidas a la que pertenece una matriz intermedia H_R , ambas asociadas con la segunda instancia de la sección 300 de codificación paramétrica, se puede utilizar otra instancia de la sección 100 de reconstrucción paramétrica en paralelo con la primera sección de reconstrucción paramétrica para reconstruir los tres canales RS, TBR y RB a partir de la señal rs downmix y los parámetros de upmix seca y húmeda asociados. Otro grupo 503 de canales incluye solo dos canales L y TFL representados por un canal l downmix. La codificación de estos dos canales en el canal l downmix y los parámetros de upmix húmeda y seca asociados se puede realizar mediante secciones de codificación y una sección de reconstrucción análogas a las descritas en relación con las Fig. 3 y 1,

respectivamente, pero para $N = 2$. Otro grupo 504 de canales comprende únicamente un solo canal LFE representado por un canal lfe downmix. En este caso, no es necesaria ninguna señal downmix y el canal LFE downmix puede ser el propio canal LFE, transformado opcionalmente en un dominio MDCT y/o codificado utilizando un codec de audio perceptual.

5 El número total de canales downmix utilizados en las Fig. 5-11 para representar la señal de audio de 11.1 canales varía. Por ejemplo, el ejemplo que se ilustra en la Fig. 5 utiliza 6 canales downmix, mientras que el ejemplo de la Fig. 7 utiliza 10 canales downmix. Para diferentes situaciones pueden resultar adecuadas diferentes configuraciones de downmix, por ejemplo en función del ancho de banda disponible para la transmisión de las señales downmix y el parámetro de upmix asociado, y/o los requisitos sobre en qué medida debe ser fiel la reconstrucción de la señal de audio de 11.1 canales.

De acuerdo con los modos de realización de ejemplo, el sistema 400 de codificación de audio descrito en relación con la Fig. 4 puede comprender una pluralidad de secciones de codificación paramétrica, incluyendo la sección 300 de codificación paramétrica descrita en relación con la Fig. 3. El sistema 400 de codificación de audio puede comprender una sección de control (en la Fig. 4 no se muestra) configurada para determinar/seleccionar un formato de codificación para la señal de audio de 11.1 canales, a partir de una colección para formatos de codificación correspondientes a las respectivas particiones de la señal de audio de 11.1 canales que se ilustra en las Fig. 5-11. El formato de codificación determinado corresponde, además, a un conjunto de reglas predefinidas (al menos algunas de las cuales pueden coincidir) para calcular los respectivos canales downmix, un conjunto de clases de matrices predefinidas (al menos algunas de las cuales pueden coincidir) para matrices intermedias H_R , y un conjunto de matrices V predefinidas (al menos algunas de las cuales pueden coincidir) para obtener los coeficientes de upmix húmeda asociados con al menos algunos de los respectivos conjuntos de canales en función de los respectivos parámetros de upmix húmeda asociados. De acuerdo con los presentes modos de realización de ejemplo, el sistema de codificación de audio está configurado para codificar la señal de audio de 11.1 canales utilizando un subconjunto de la pluralidad de secciones de codificación apropiadas para el formato de codificación determinado. Si, por ejemplo, el formato de codificación determinado corresponde a la partición de los 11.1 canales que se ilustra en la Fig. 1, el sistema de codificación puede utilizar 2 secciones de codificación configuradas para representar los respectivos conjuntos de 3 canales por canales downmix individuales respectivos, 2 secciones de codificación configuradas para representar los respectivos conjuntos de 2 canales por canales downmix individuales respectivos, y 2 secciones de codificación configuradas para representar el respectivo canal individual como canales downmix individuales respectivos. Todas las señales downmix y los parámetros de upmix húmeda y seca asociados pueden codificarse en el mismo flujo B de bits para transmitirlos a un decodificador. Se debe observar que algunas de las secciones de codificación pueden utilizar el formato compacto de los metadatos que acompañan a los canales downmix, esto es, los parámetros de upmix húmeda y los parámetros de upmix húmeda, en tanto que, en al menos algunos modos de realización de ejemplo, se pueden utilizar otros formatos de metadatos. Por ejemplo, algunas de las secciones de codificación pueden generar el número total de coeficientes de upmix húmeda y seca en lugar de los parámetros de upmix húmeda y seca. También se conciben modos de realización en los que algunos canales se codifican para la reconstrucción utilizando menos de $N - 1$ canales descorrelacionados (o incluso sin ninguna descorrelación), y donde, por consiguiente, los metadatos para la reconstrucción paramétrica pueden tomar una forma diferente.

De acuerdo con los modos de realización de ejemplo, el sistema 200 de decodificación de audio descrito en relación con la Fig. 2 puede comprender una pluralidad correspondiente de secciones de reconstrucción, incluyendo la sección 100 de reconstrucción paramétrica descrita en relación con la Fig. 1, para reconstruir los respectivos conjuntos de canales de la señal de audio de 11.1 canales representados por las respectivas señales downmix. El sistema 200 de decodificación de audio puede comprender una sección de control (en la Fig. 2 no se muestra) configurada para recibir una señalización desde el codificador indicando el formato de codificación determinado, y el sistema 200 de decodificación de audio puede utilizar un subconjunto apropiado de la pluralidad de secciones de reconstrucción para reconstruir la señal de audio de 11.1 canales a partir de las señales downmix recibidas y los parámetros de upmix seca y húmeda asociados.

Las Fig. 12-13 ilustran formas alternativas de representación de una señal de audio de 13.1 canales por medio de canales downmix, de acuerdo con los modos de realización de ejemplo. La señal de audio de 13.1 canales incluye los canales: pantalla izquierdo (LSCRN), desplazado izquierdo (LW), pantalla derecho (RSCRN), desplazado derecho (RW), central (C), efectos de baja frecuencia (LFE), lateral izquierdo (LS), lateral derecho (RS), izquierdo posterior (LB), derecho posterior (RB), izquierdo frontal superior (TFL), derecho frontal superior (TFR), izquierdo posterior superior (TBL) y derecho posterior superior (TBR). La codificación de los respectivos grupos de canales como canales downmix respectivos se puede realizar mediante secciones de codificación respectivas operando de forma independiente en paralelo, tal como se ha descrito más arriba en relación con las Fig. 5-11. De modo análogo, la reconstrucción de los grupos respectivos de canales a partir de los respectivos canales downmix y los parámetros de upmix asociados se puede realizar mediante las respectivas secciones de reconstrucción operando independiente en paralelo.

Las Fig. 14-16 ilustran formas alternativas de representación de una señal de audio de 22.2 canales por medio de señales downmix, de acuerdo con los modos de realización de ejemplo. La señal de audio de 22.2 canales incluye los canales: efectos de baja frecuencia 1 (LFE1), efectos de baja frecuencia 2 (LFE2), centro frontal inferior (BFC), centro (C), centro frontal superior (TFC), desplazado izquierdo (LW), izquierdo frontal inferior (BFL), izquierdo (L),

5 izquierdo frontal superior (TFL), izquierdo lateral superior (TSL), izquierdo posterior superior (TBL), lateral izquierdo (LS), posterior izquierdo (LB), centro superior (TC), centro posterior superior (TBC), centro posterior (CB), derecho frontal inferior (BFR), derecho (R), desplazado derecho (RW), derecho frontal superior (TFR), derecho lateral superior (TSR), derecho posterior superior (TBR), lateral derecho (RS) y posterior derecho (RB). La partición de la
 10 señal de audio de 22.2 canales que se ilustra en la Fig. 16 incluye un grupo 1601 de canales que incluye cuatro canales. La sección 300 de codificación paramétrica que se ha descrito en relación con la Fig. 3 se puede utilizar, pero implementada con $N = 4$, para codificar estos canales como una señal downmix junto con los parámetros de upmix húmeda y seca asociados. De modo análogo, la sección 100 de reconstrucción paramétrica que se ha descrito en relación con la Fig. 1 se puede utilizar, pero implementada con $N = 4$, para reconstruir estos canales a partir de la señal downmix y los parámetros de upmix húmeda y seca asociados.

III. Equivalentes, extensiones, alternativas y miscelánea

Cualesquiera signos de referencia que aparezcan en las reivindicaciones no debe entenderse como limitantes de su alcance.

15 En las reivindicaciones, el término "comprende" no excluye otros elementos o pasos, y el artículo indefinido "un" o "una" no excluye una pluralidad. El mero hecho de que ciertas medidas se enumeren en reivindicaciones dependientes mutuamente diferentes no indica que no se pueda utilizar de forma ventajosa una combinación de dichas medidas.

20 Los dispositivos y métodos descritos más arriba en la presente solicitud se pueden implementar mediante software, firmware, hardware o una combinación de los mismos. En una implementación de hardware, la división de tareas entre unidades funcionales a las que se hace referencia en la descripción que se ha realizado más arriba no corresponde necesariamente a la división en unidades físicas; en sentido contrario, un componente físico puede tener múltiples funcionalidades, y una tarea puede ser realizada por varios componentes físicos en cooperación. Ciertos componentes, o todos los componentes, se pueden implementar mediante un software ejecutado por un procesador o un microprocesador de señales digitales, o se puede implementar mediante un hardware o como un
 25 circuito integrado de aplicación específica. Dicho software se puede distribuir en medios legibles por un ordenador, que pueden comprender medios de almacenamiento en el ordenador (o medios no transitorios) y medios de comunicación (o medios transitorios). Como es bien conocido por una persona experimentada en la técnica, la expresión medios de almacenamiento en el ordenador incluye tanto medios volátiles como no volátiles, medios extraíbles y no extraíbles, implementados mediante cualquier método o tecnología para el almacenamiento de información tal como, por ejemplo, instrucciones legibles por un ordenador, estructuras de datos, módulos de programa u otros datos. Los medios de almacenamiento en el ordenador incluyen, pero no se limitan a, RAM, ROM, EEPROM, memoria flash u otra tecnología de memoria, CD-ROM, discos versátiles digitales (DVD) u otra forma de almacenamiento en disco óptico, casetes magnéticas, cinta magnética, almacenamiento en disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro medio que se pueda utilizar para almacenar la
 30 información deseada y al que se pueda acceder mediante un ordenador. Además, es bien conocido para la persona experimentada en la técnica que los medios de comunicación incorporan normalmente instrucciones legibles por un ordenador, estructuras de datos, módulos de programa u otros datos en una señal de datos modulada como, por ejemplo, una onda portadora u otro mecanismo de transporte e incluyen cualquier medio de entrega de información.

REIVINDICACIONES

1. Un método para reconstruir una señal (X) de audio de N canales, en donde $N \geq 3$, comprendiendo dicho método:
 - recibir una señal (Y) downmix (mezcla con reducción de canales) de un solo canal junto con los parámetros (\tilde{C} , \tilde{P}) de upmix seca (dry, original) y húmeda (wet, procesada) asociados;
 - 5 calcular una señal upmix seca mediante un mapeo lineal de la señal downmix, en donde a la señal downmix se le aplica un conjunto de coeficientes (C) de upmix seca;
 - generar una señal descorrelacionada (Z) a partir de la señal downmix, en donde la señal descorrelacionada tiene $N - 1$ canales;
 - calcular una señal upmix (mezcla con aumento de canales) húmeda mediante un mapeo lineal de la señal descorrelacionada de $N - 1$ canales, en donde a la señal descorrelacionada de $N - 1$ canales se le aplica un conjunto de coeficientes (P) de upmix húmeda; y
 - 10 combinar las señales upmix seca y húmeda para obtener una señal (\hat{X}) multidimensional reconstruida correspondiente a la señal de audio de N canales que se desea reconstruir,
 - en donde el método comprende, además:
 - 15 determinar el conjunto de coeficientes de upmix seca a partir de los parámetros de upmix seca recibidos;
 - rellenar una matriz intermedia que tiene un número de elementos mayor que el número de parámetros de upmix húmeda recibidos, a partir de los parámetros de upmix húmeda recibidos y sabiendo que la matriz intermedia pertenece a una clase de matrices predefinidas, en donde las propiedades conocidas de todas las matrices de la clase de matrices predefinidas incluyen relaciones conocidas entre los elementos de la matriz predefinida, o
 - 20 elementos de la matriz predefinida que son cero; y
 - obtener el conjunto de coeficientes de upmix húmeda multiplicando la matriz intermedia por una matriz predefinida, en donde el conjunto de coeficientes de upmix húmeda corresponde a la matriz resultante de la multiplicación y contiene un número de coeficientes mayor que el número de elementos en la matriz intermedia.
2. El método de la reivindicación 1, en donde la recepción de los parámetros de upmix húmeda incluye recibir $N(N - 1)/2$ parámetros de upmix húmeda, en donde rellenar la matriz intermedia incluye obtener valores para $(N - 1)^2$ elementos de la matriz a partir de los $N(N - 1)/2$ parámetros de upmix húmeda recibidos y sabiendo que la matriz intermedia pertenece a la clase de matrices predefinidas, en donde la matriz predefinida tiene $N(N - 1)$ elementos, y en donde el conjunto de coeficientes de upmix húmeda contiene $N(N - 1)$ coeficientes,
 - y en donde opcionalmente rellenar la matriz intermedia incluye utilizar los parámetros de upmix húmeda recibidos
 - 30 como elementos de la matriz intermedia.
3. El método de cualquiera de las reivindicaciones precedentes, en donde la recepción de los parámetros de upmix seca incluye recibir $N - 1$ parámetros de upmix seca, en donde el conjunto de coeficientes de upmix seca contiene N coeficientes, y en donde el conjunto de coeficientes de upmix seca se determina a partir de los $N - 1$ parámetros de upmix seca recibidos y de acuerdo con una relación predefinida entre los coeficientes del conjunto de coeficientes de upmix seca,
 - 35 y en donde opcionalmente la clase de matrices predefinidas es una de las siguientes:
 - matrices triangulares inferiores o superiores, en donde las propiedades conocidas de todas las matrices de la clase incluyen elementos de la matriz predefinida que son cero;
 - matrices simétricas, en donde las propiedades conocidas de todas las matrices de la clase incluyen elementos
 - 40 de la matriz predefinida que 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.
4. El método de cualquiera de las reivindicaciones precedentes, en donde la señal downmix se puede obtener, de acuerdo con una regla predefinida, mediante un mapeo lineal de la señal de audio de N canales que se desea reconstruir, en donde la regla predefinida define una operación de downmix predefinida, y en donde dicha matriz predefinida se basa en vectores que abarcan el espacio del núcleo de dicha operación de downmix predefinida.
5. El método de cualquiera de las reivindicaciones precedentes, en donde la recepción de la señal downmix de un solo canal junto con los parámetros de upmix seca y húmeda asociados incluye recibir un segmento de tiempo o un átomo de tiempo/frecuencia de la señal downmix junto con los parámetros de upmix seca y húmeda asociados, y en

donde dicha señal multidimensional reconstruida corresponde a un segmento de tiempo o a un átomo de tiempo/frecuencia de la señal de audio de N canales que se desea reconstruir.

5 6. Un sistema (200) de decodificación de audio que comprende una primera sección (100) de reconstrucción paramétrica configurada para reconstruir una señal (X) de audio de N canales a partir de una primera señal (Y) downmix de un solo canal y los parámetros (\tilde{C} , \tilde{P}) de upmix seca y húmeda asociados, en donde $N \geq 3$, comprendiendo dicha primera sección de reconstrucción paramétrica:

una primera sección (101) de descorrelación configurada para recibir la primera señal downmix y generar, a partir de la misma, una primera señal (Z) descorrelacionada con $N - 1$ canales;

una primera sección seca (102) de upmix configurada para

10 recibir los parámetros \tilde{C} de upmix seca y la señal downmix,

determinar un primer conjunto de coeficientes (C) de upmix seca a partir de los parámetros de upmix seca, y

generar una primera señal upmix seca calculada mapeando linealmente la primera señal downmix de acuerdo con el primer conjunto de coeficientes de upmix seca;

una primera sección (103) de upmix húmeda configurada para

15 recibir los parámetros \tilde{P} de upmix húmeda y la primera señal descorrelacionada,

rellenar una primera matriz intermedia cuyo número de elementos es mayor que el número de parámetros de upmix húmeda recibidos, basándose en los parámetros de upmix húmeda recibidos y sabiendo que la primera matriz intermedia pertenece a una primera clase de matrices predefinidas, en donde las propiedades conocidas de todas las matrices de la primera clase de matrices predefinidas incluye relaciones conocidas entre los elementos de la matriz predefinida o elementos de la matriz predefinida que son cero,

20 obtener un primer conjunto de coeficientes (P) de upmix húmeda multiplicando la primera matriz intermedia por una primera matriz predefinida, en donde el primer conjunto de coeficientes de upmix húmeda corresponde a la matriz resultante de la multiplicación y contiene más coeficientes que el número de elementos de la primera matriz intermedia, y

25 generar una primera señal upmix húmeda calculada mapeando linealmente los $N - 1$ canales de la primera señal descorrelacionada de acuerdo con el primer conjunto de coeficientes de upmix húmeda; y

una primera sección (104) de combinación configurada para recibir la primera señal upmix seca y la primera señal upmix húmeda y combinar estas señales para obtener una primera señal (\hat{X}) multidimensional reconstruida correspondiente a la señal de audio de N canales que se desea reconstruir.

30 7. El sistema de decodificación de audio de la reivindicación 6, que comprende además una segunda sección de reconstrucción paramétrica que se utiliza independientemente de la primera sección de reconstrucción paramétrica y está configurada para reconstruir una señal de audio de N_2 canales a partir de una segunda señal downmix de un solo canal y los parámetros de upmix seca y húmeda asociados, en donde $N_2 \geq 2$, comprendiendo dicha segunda sección de reconstrucción paramétrica una segunda sección de descorrelación, una segunda sección de upmix seca, una segunda sección de upmix húmeda y una segunda sección de combinación, estando configuradas dichas secciones de la segunda sección de reconstrucción paramétrica de forma análoga a las secciones correspondientes de la primera sección de reconstrucción paramétrica, en donde la segunda sección de upmix húmeda está configurada para utilizar una segunda matriz intermedia perteneciente a una segunda clase de matrices predefinidas y una segunda matriz predefinida.

40 8. El sistema de decodificación de audio de la reivindicación 6 ó 7, en donde el sistema de decodificación de audio está adaptado para reconstruir una señal de audio multicanal a partir de una pluralidad de canales downmix y los parámetros de upmix seca y húmeda asociados, en donde el sistema de decodificación de audio comprende:

45 una pluralidad de secciones de reconstrucción, que incluye secciones de reconstrucción paramétrica que se utilizan para reconstruir de forma independiente conjuntos respectivos de canales de señal de audio a partir de los respectivos canales downmix y de los respectivos parámetros de upmix seca y húmeda asociados; y

50 una sección de control configurada para recibir una señalización indicando un formato de codificación de la señal de audio multicanal correspondiente a una partición de los canales de la señal de audio multicanal en conjuntos (501-504) de canales representados por los canales downmix respectivos y, para al menos algunos de los canales downmix, por los respectivos parámetros de upmix seca y húmeda asociados, correspondiendo además dicho formato de codificación a un conjunto de matrices predefinidas para obtener los coeficientes de upmix húmeda asociados con al menos algunos de los respectivos conjuntos de canales de acuerdo con los respectivos parámetros de upmix húmeda asociados,

en donde el sistema de decodificación está configurado para reconstruir la señal de audio multicanal utilizando un primer subconjunto de la pluralidad de secciones de reconstrucción, en respuesta a la señalización recibida que indica un primer formato de codificación, en donde el sistema de decodificación está configurado para reconstruir la señal de audio multicanal utilizando un segundo subconjunto de la pluralidad de secciones de reconstrucción, en respuesta a la señalización recibida que indica un segundo formato de codificación, y en donde al menos uno entre el primer y segundo subconjuntos de secciones de reconstrucción comprende dicha primera sección de reconstrucción paramétrica.

9. El sistema de decodificación de audio de la reivindicación 8, en donde la pluralidad de secciones de reconstrucción incluye una sección de reconstrucción de un solo canal que se utiliza para reconstruir de forma independiente un único canal de audio a partir de un canal downmix en el que únicamente se ha codificado un solo canal de audio, y en donde al menos uno entre el primer y segundo subconjuntos de secciones de reconstrucción comprende la sección de reconstrucción de un solo canal, y/o

en donde el primer formato de codificación corresponde a la reconstrucción de dicha señal de audio multicanal a partir de un número de canales downmix menor que el segundo formato de codificación.

10. Un método para codificar una señal (X) de audio de N canales como una señal (Y) downmix de un solo canal junto con los metadatos apropiados para la reconstrucción paramétrica de dicha señal de audio a partir de la señal downmix y una señal (Z) descorrelacionada determinada a partir de la señal downmix, en donde $N \geq 3$, y en donde la señal descorrelacionada tiene $N - 1$ canales, comprendiendo dicho método:

recibir dicha señal de audio;

calcular, de acuerdo con una regla predefinida, la señal downmix de un solo canal mediante un mapeo lineal de dicha señal de audio;

determinar un conjunto (C) de coeficientes de upmix seca con el fin de definir un mapeo lineal de la señal downmix que aproxime dicha señal de audio;

determinar una matriz intermedia a partir de una diferencia entre una covarianza de dicha señal de audio recibida y una covarianza de dicha señal de audio aproximada mediante el mapeo lineal de la señal downmix, en donde la matriz intermedia, al multiplicarse por una matriz predefinida, corresponde a un conjunto (P) de coeficientes de upmix húmeda que define un mapeo lineal de los $N - 1$ canales de dicha señal descorrelacionada como parte de la reconstrucción paramétrica de dicha señal de audio, en donde el conjunto de coeficientes de upmix húmeda contiene más coeficientes que el número de elementos de la matriz intermedia; y

generar la señal downmix junto con los parámetros \tilde{C} de upmix seca, a partir de los cuales se puede derivar el conjunto de coeficientes de upmix seca, y los parámetros \tilde{P} de upmix húmeda, en donde la matriz intermedia tiene más elementos que el número de parámetros de upmix húmeda generados, y en donde la matriz intermedia está definida de forma unívoca por los parámetros de upmix húmeda generados siempre que la matriz intermedia pertenezca a una clase de matrices predefinidas, en donde las propiedades conocidas de todas las matrices de la clase de matrices predefinidas incluyen relaciones conocidas entre los elementos de la matriz predefinida o elementos de la matriz predefinida que son cero.

11. El método de la reivindicación 10, en donde la determinación de la matriz intermedia incluye determinar la matriz intermedia de tal modo que una covarianza de la señal obtenida mediante el mapeo lineal de dicha señal descorrelacionada, definida por el conjunto de coeficientes de upmix húmeda, aproxime la diferencia entre la covarianza de dicha señal de audio recibida y la covarianza de dicha señal de audio aproximada por el mapeo lineal de la señal downmix, y/o

en donde la generación de los parámetros de upmix húmeda incluye generar no más de $N(N - 1)/2$ parámetros de upmix húmeda, en donde la matriz intermedia tiene $(N - 1)^2$ elementos y está definida de forma unívoca por los parámetros de upmix húmeda generados siempre que la matriz intermedia pertenezca a la clase de matrices predefinidas, y en donde el conjunto de coeficientes de upmix húmeda contiene $N(N - 1)$ coeficientes, y/o en donde el conjunto de coeficientes de upmix seca incluye N coeficientes, y en donde la generación de los parámetros de upmix seca incluye generar no más de $N - 1$ parámetros de upmix seca, pudiéndose obtener el conjunto de coeficientes de upmix seca a partir de los $N - 1$ parámetros de upmix seca utilizando dicha regla predefinida, y/o

en donde el conjunto de coeficientes de upmix seca determinados define un mapeo lineal de la señal downmix correspondiente a una aproximación del mínimo error cuadrático medio de dicha señal de audio.

12. Un sistema (400) de codificación de audio que comprende una sección (300) de codificación paramétrica configurada para codificar una señal (X) de audio de N canales como una señal (Y) downmix de un solo canal junto con los metadatos apropiados para la reconstrucción paramétrica de dicha señal de audio a partir de la señal downmix y una señal (Z) descorrelacionada determinada a partir de la señal downmix, en donde $N \geq 3$, y en donde la señal descorrelacionada tiene $N - 1$ canales, comprendiendo dicha sección de codificación paramétrica:

una sección (301) de downmix configurada para recibir dicha señal de audio y calcular, de acuerdo con una regla predefinida, la señal downmix de un solo canal mediante un mapeo lineal de dicha señal de audio;

una primera sección (302) de análisis configurada para determinar un conjunto (C) de coeficientes de upmix seca con el fin de definir un mapeo lineal de la señal downmix que aproxime dicha señal de audio; y

5 una segunda sección (303) de análisis configurada para determinar una matriz intermedia basada en una diferencia entre una covarianza de dicha señal de audio recibida y una covarianza de dicha señal de audio aproximada mediante el mapeo lineal de la señal downmix, en donde la matriz intermedia, al multiplicarse por una matriz predefinida, corresponde a un conjunto (P) de coeficientes de upmix húmeda que define un mapeo lineal de los $N - 1$ canales de dicha señal descorrelacionada como parte de la reconstrucción paramétrica de dicha señal de audio, en donde el conjunto de coeficientes de upmix húmeda contiene más coeficientes que el número de elementos de la matriz intermedia,

15 en donde la sección de codificación paramétrica está configurada para generar la señal downmix junto con los parámetros \tilde{C} de upmix seca, a partir de los cuales se puede derivar el conjunto de coeficientes de upmix seca, y los parámetros \tilde{P} de upmix húmeda, en donde la matriz intermedia tiene más elementos que el número de parámetros de upmix húmeda generados, y en donde la matriz intermedia está definida de forma unívoca por los parámetros de upmix húmeda generados siempre que la matriz intermedia pertenezca a una clase de matrices predefinidas, en donde las propiedades conocidas de todas las matrices de la clase de matrices predefinidas incluyen relaciones conocidas entre los elementos de la matriz predefinida o elementos de la matriz predefinida que son cero.

20 13. El sistema de codificación de audio de la reivindicación 12, en donde el sistema de codificación de audio está adaptado para proporcionar una representación de una señal de audio multicanal en forma de una pluralidad de canales downmix y los parámetros de upmix seca y húmeda asociados, en donde el sistema de codificación de audio comprende:

25 una pluralidad de secciones de codificación, que incluye secciones de codificación paramétrica que se utilizan para calcular de forma independiente los respectivos canales downmix y los respectivos parámetros de upmix asociados en función de los conjuntos de canales de señal de audio respectivos;

30 una sección de control configurada para determinar un formato de codificación para dicha señal de audio multicanal correspondiente a una partición de los canales de dicha señal de audio multicanal en conjuntos (501-504) de canales para ser representados por los respectivos canales downmix y, para al menos algunos de los canales downmix, por los respectivos parámetros de upmix asociados, correspondiendo además dicho formato de codificación a un conjunto de reglas predefinidas para calcular al menos algunos de los respectivos canales downmix,

35 en donde el sistema de codificación de audio está configurado para codificar la señal de audio multicanal utilizando un primer subconjunto de la pluralidad de secciones de codificación, en respuesta a que el formato de codificación determinado es un primer formato de codificación, en donde el sistema de codificación de audio está configurado para codificar la señal de audio multicanal utilizando un segundo subconjunto de la pluralidad de secciones de codificación, en respuesta a que el formato de codificación determinado es un segundo formato de codificación, y en donde al menos uno entre el primer y el segundo subconjuntos de las secciones de codificación comprende dicha primera sección de codificación paramétrica.

40 14. El sistema de codificación de audio de la reivindicación 13, en donde la pluralidad de secciones de codificación incluye una sección de codificación de un solo canal que se utiliza para codificar de forma independiente no más de un canal de audio individual en un canal downmix, y en donde al menos uno entre el primer y el segundo subconjuntos de las secciones de codificación comprende la sección de codificación de un solo canal.

45 15. Un producto de programa de ordenador que comprende un medio legible por ordenador con instrucciones para poner en práctica el método de cualquiera de las reivindicaciones 1-5, 10 u 11.

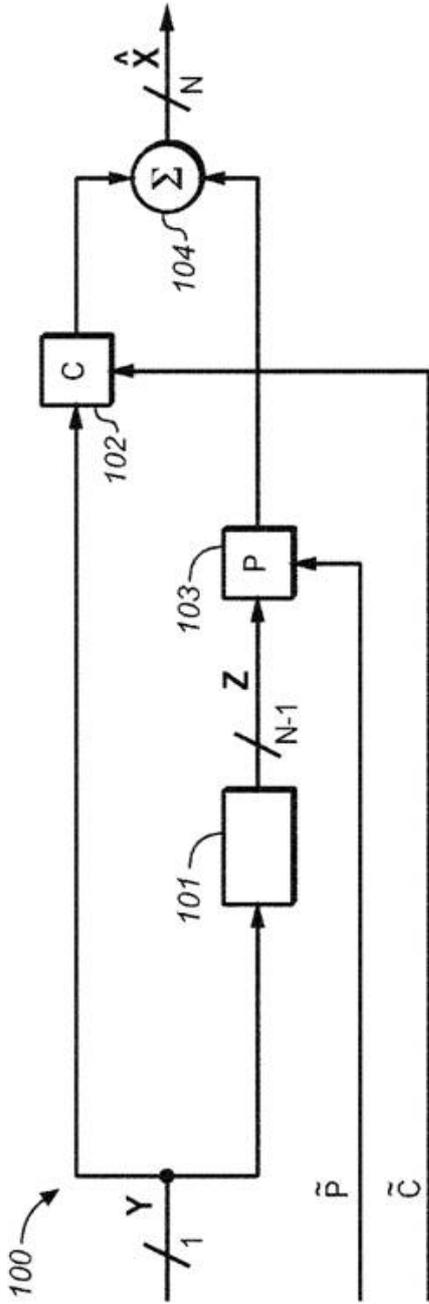


FIG. 1

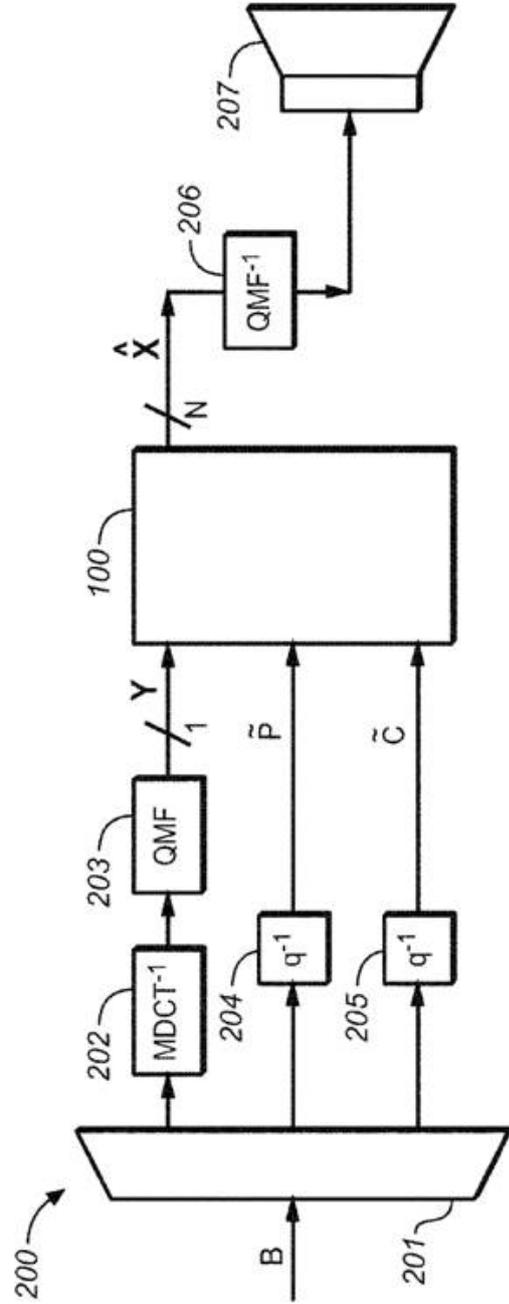
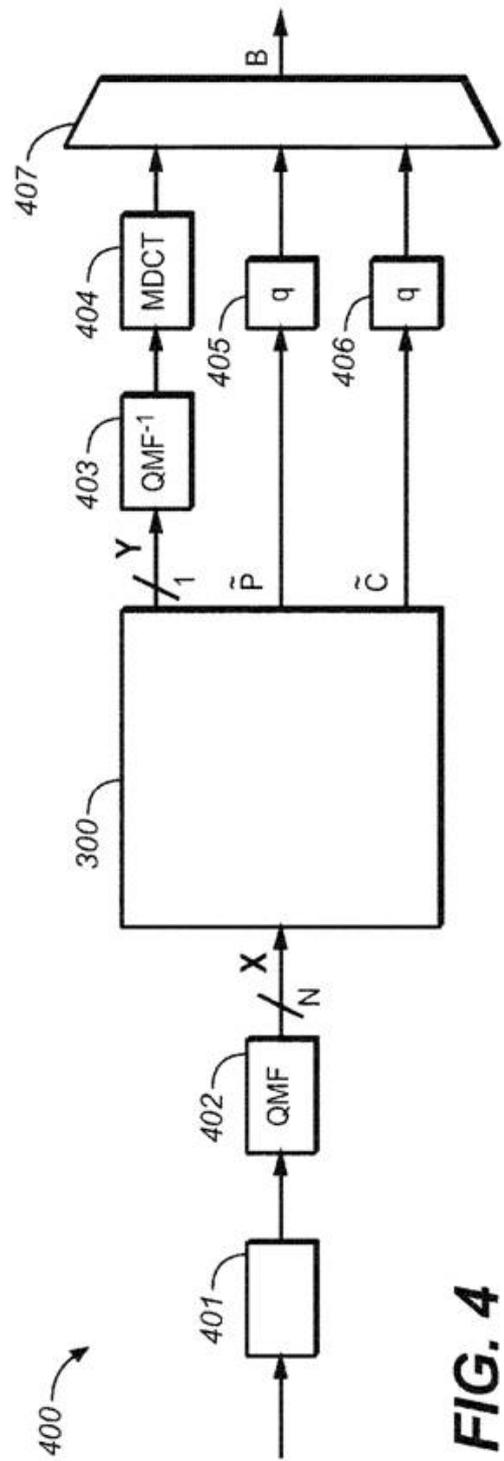
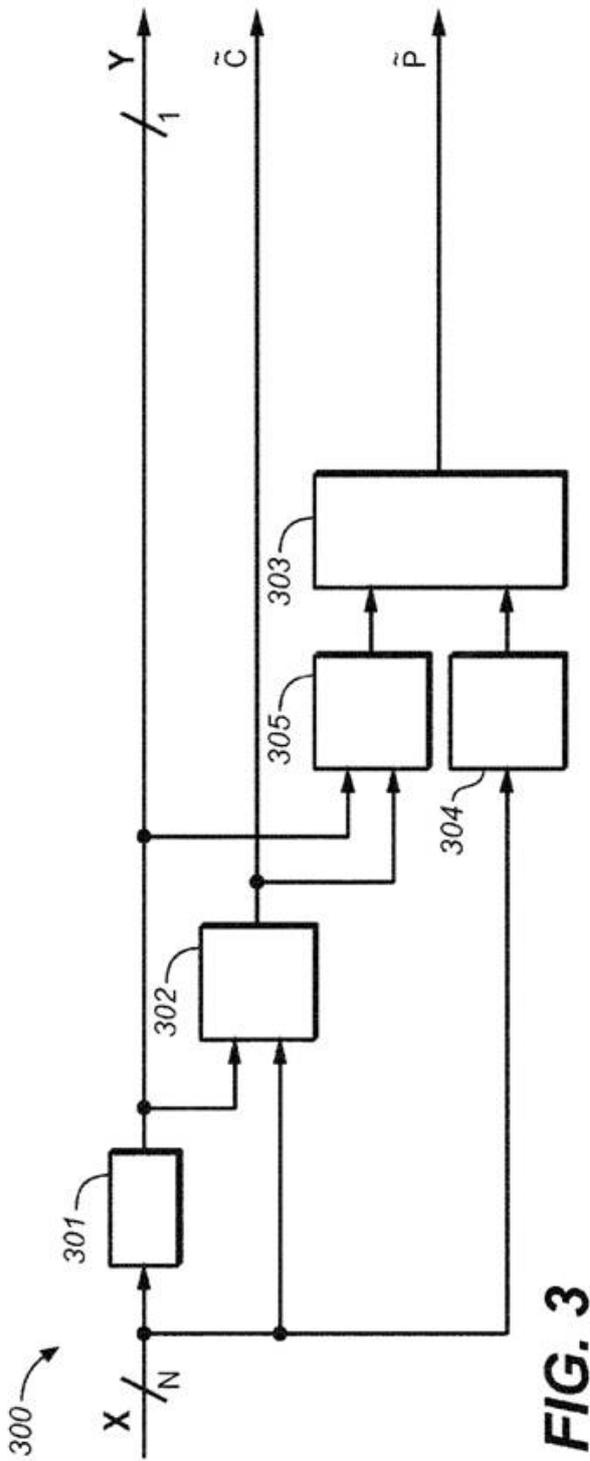
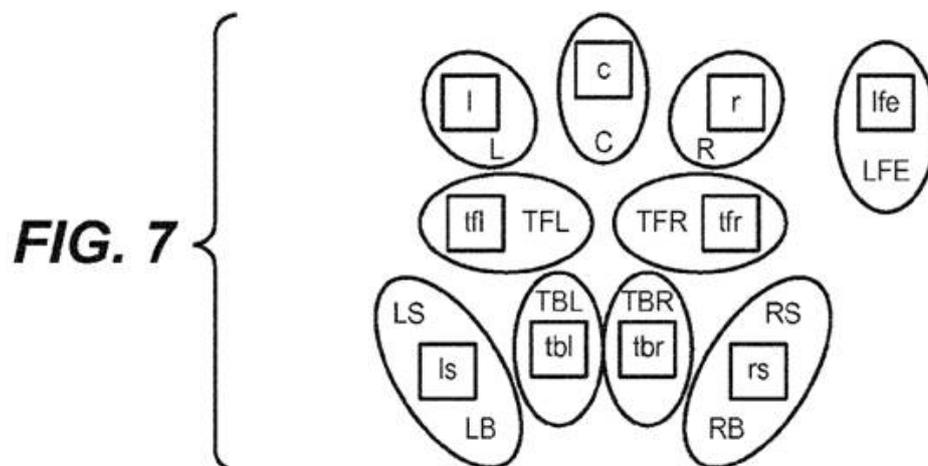
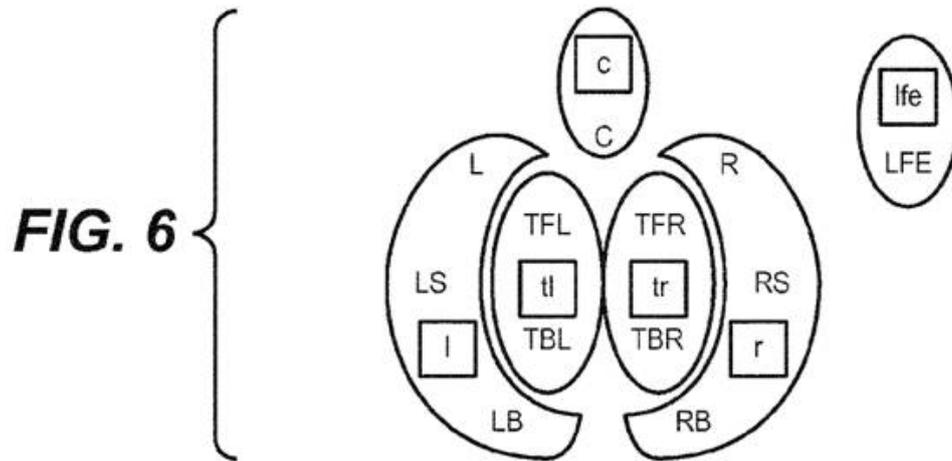
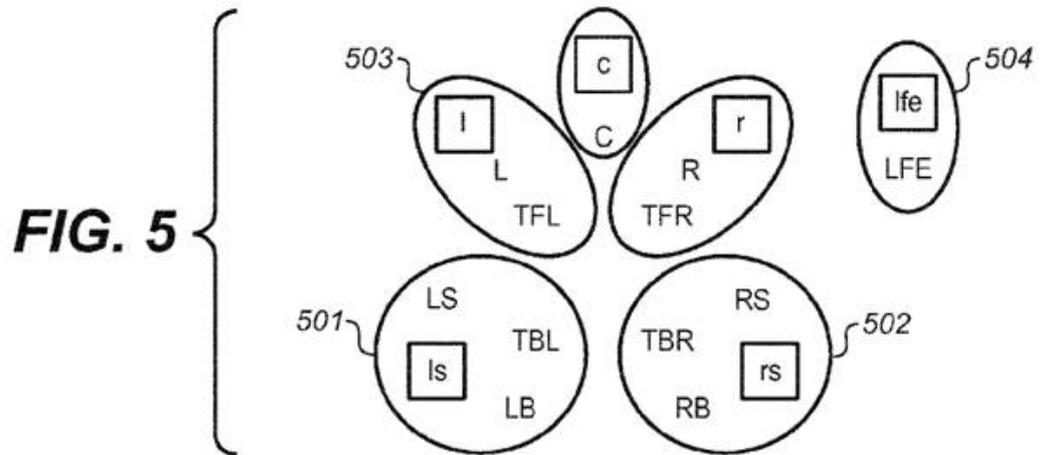
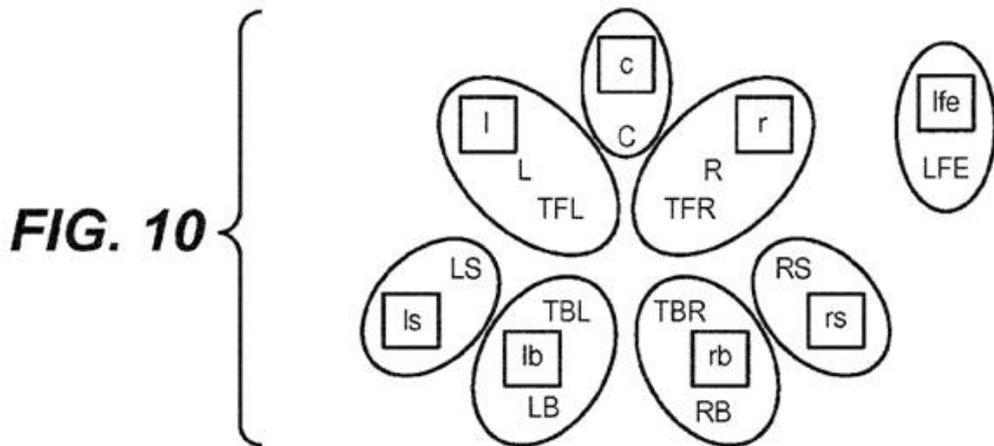
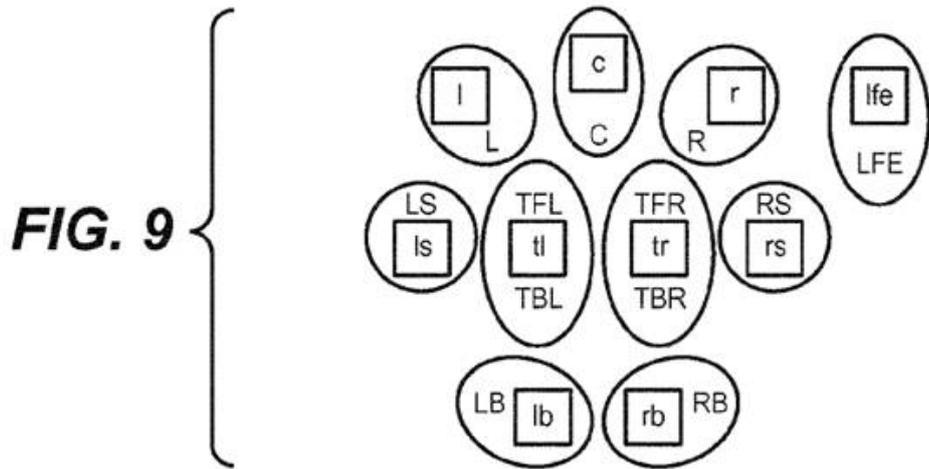
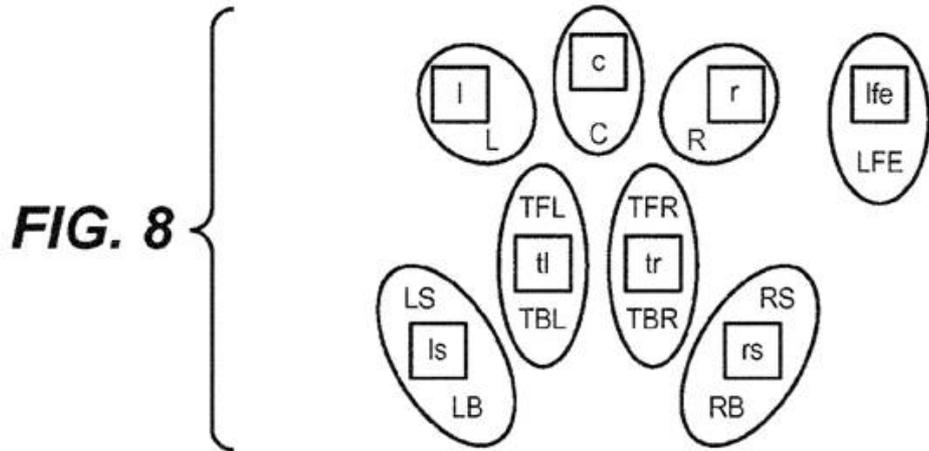
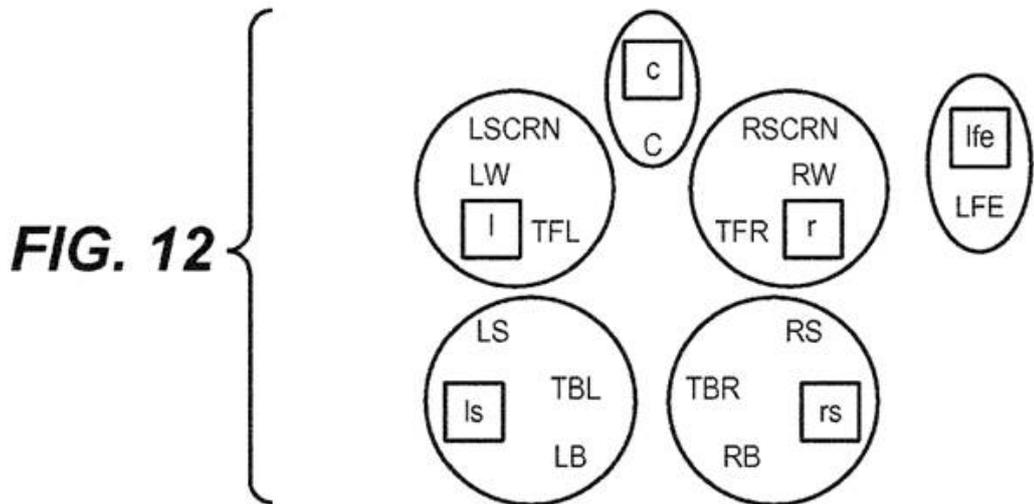
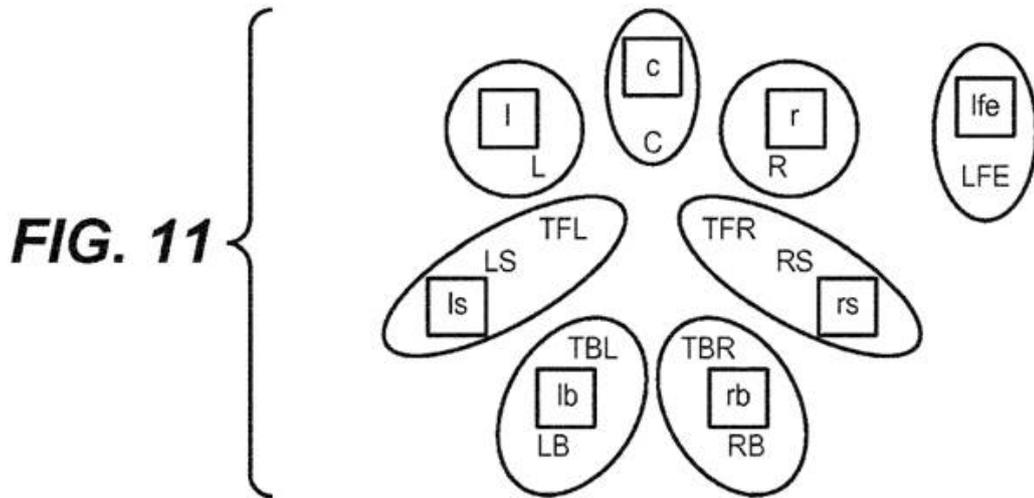


FIG. 2









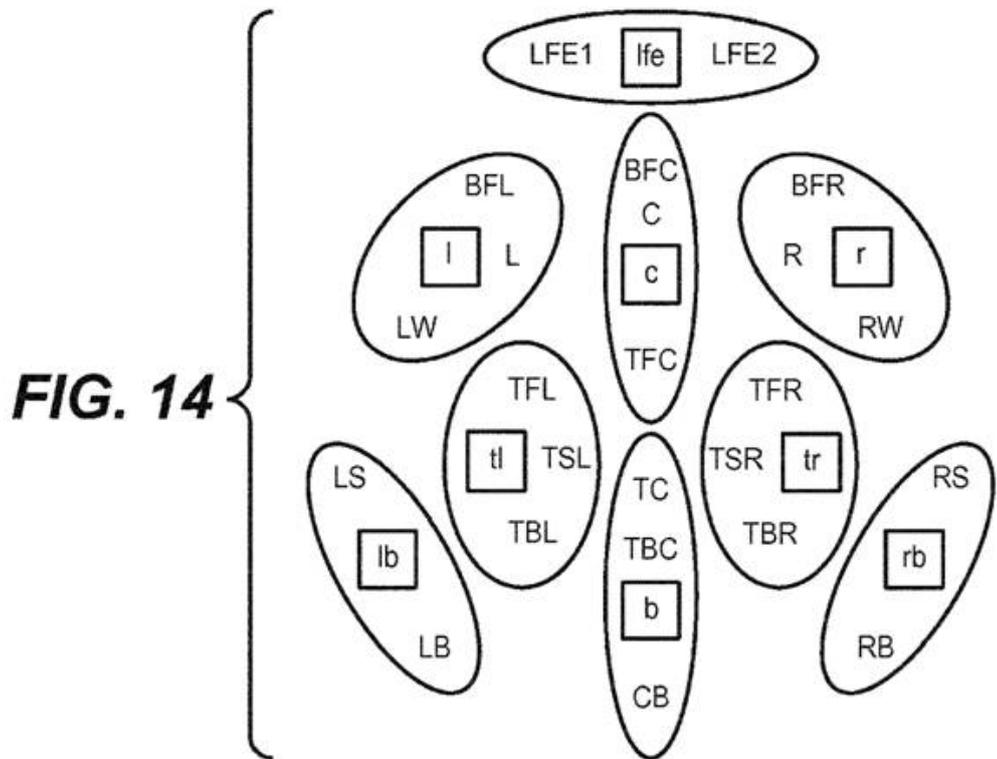
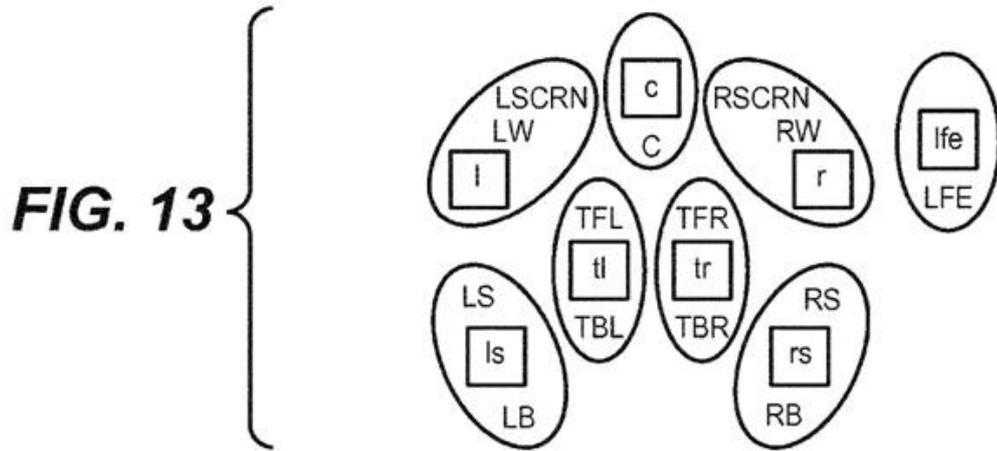


FIG. 15

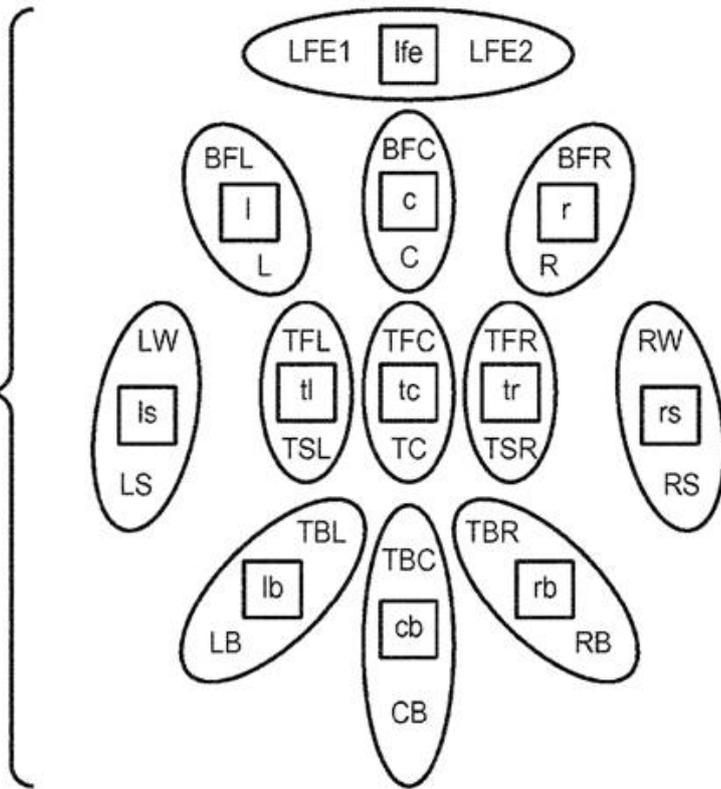


FIG. 16

