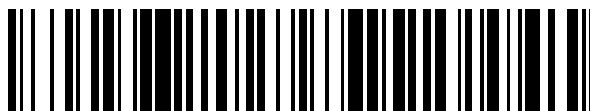


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 641 390**

51 Int. Cl.:

**G10L 19/008** (2013.01)

**G10L 19/032** (2013.01)

**H04S 3/00** (2006.01)

**G10L 19/24** (2013.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.05.2013 PCT/US2013/040919**

87 Fecha y número de publicación internacional: **21.11.2013 WO13173314**

96 Fecha de presentación y número de la solicitud europea: **14.05.2013 E 13726928 (8)**

97 Fecha y número de publicación de la concesión europea: **16.08.2017 EP 2850613**

54 Título: **Codificación y decodificación eficientes de una señal de audio multicanal con múltiples flujos secundarios**

30 Prioridad:

**15.05.2012 US 201261647226 P**

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

**08.11.2017**

73 Titular/es:

**DOLBY LABORATORIES LICENSING CORPORATION (50.0%)  
100 Potrero Avenue  
San Francisco, CA 94103-4813, US y  
DOLBY INTERNATIONAL AB (50.0%)**

72 Inventor/es:

**MUNDT, HARALD;  
RIEDMILLER, JEFFREY;  
ROEDEN, KARL J.;  
WARD, MICHAEL y  
WILLIAMS, PHILLIP**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

**ES 2 641 390 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

Codificación y decodificación eficientes de una señal de audio multicanal con múltiples flujos secundarios

### Referencia cruzada a solicitudes asociadas

5 Esta solicitud reivindica el beneficio de prioridad de la Solicitud de Patente Provisional de Estados Unidos Núm. Ser. 61/647.226 presentada el 15 de mayo de 2012

### Campo técnico de la invención

La presente invención está relacionada con codificación/decodificación de audio. En particular, el presente documento está relacionado con un método y un sistema para mejorar la calidad de señales de audio multicanal codificadas.

### 10 Antecedentes de la invención

Actualmente se encuentran en uso varios sistemas de renderización (conversión) de audio multicanal como, por ejemplo, sistemas de renderización de audio multicanal 5.1, 7.1 ó 9.1. Los sistemas de renderización de audio multicanal permiten la generación de un sonido envolvente que tenga su origen en localizaciones de altavoces 5+1, 7+1 ó 9+1, respectivamente. Para una transmisión eficiente o para un almacenamiento eficiente de señales de audio multicanal correspondientes, se utilizan sistemas codec (codificador/decodificador) de audio multicanal como, por ejemplo, Dolby Digital o Dolby Digital Plus. Estos sistemas codec de audio multicanal son típicamente compatibles hacia atrás con el fin de permitir un decodificador de audio multicanal N.1 (por ejemplo, N=5) para decodificar y renderizar al menos parte de una señal de audio multicanal M.1 (por ejemplo, M=7), siendo M mayor que N. Más en particular, los flujos de bits generados por los sistemas codec de audio multicanal son típicamente compatibles hacia atrás con el fin de permitir que un decodificador de audio multicanal N.1 (por ejemplo, N=5) decodifique y renderice al menos parte de una señal de audio multicanal M.1 (por ejemplo, M=7). A modo de ejemplo, un flujo de bits codificado de una señal de audio multicanal 7.1 debería poderse decodificar por un decodificador de audio multicanal 5.1. Una forma posible de implementar dicha compatibilidad hacia atrás es codificar una señal de audio multicanal M.1 en una pluralidad de flujos secundarios (por ejemplo, en un flujo independiente (denominado de aquí en adelante "IS") y en uno o más flujos secundarios dependientes (denominados de aquí en adelante "DS")). El IS puede comprender una señal de audio multicanal N.1 codificada básica (por ejemplo, una señal de audio 5.1 codificada) y uno o más DS pueden comprender unos canales de sustitución y/o ampliación para renderizar la señal de audio multicanal M.1 completa (tal como se resumirá con mayor detalle más abajo). Además, el flujo de bits puede comprender múltiples IS (esto es, una pluralidad de flujos secundarios independientes) teniendo cada uno uno o más DS asociados. La pluralidad de IS y los DS asociados pueden, por ejemplo, utilizarse para transportar una pluralidad de diferentes programas de difusión o una pluralidad de pistas de audio asociadas (por ejemplo, para diferentes idiomas o comentarios directores, etc.), respectivamente.

35 El presente documento aborda el aspecto de una codificación eficiente de una pluralidad de flujos secundarios (por ejemplo, un IS y uno o más DS asociados o una pluralidad de IS y uno o más DS asociados respectivos) de una señal de audio multicanal.

### Resumen de la invención

La invención aborda el problema declarado más arriba de una codificación eficiente de una pluralidad de flujos secundarios de una señal de audio multicanal tal como se define en las reivindicaciones independientes.

40 De acuerdo con un aspecto se describe un codificador de audio configurado para codificar una señal de audio multicanal en función de una tasa de datos total disponible. La señal de audio multicanal puede, por ejemplo, ser una señal de audio multicanal 9.1, 7.1 ó 5.1. El codificador de audio puede ser un codificador de audio por tramas configurado para codificar una secuencia de tramas de la señal de audio multicanal, con el fin de generar una secuencia correspondiente de tramas codificadas. En particular, el codificador se puede configurar para realizar una codificación de acuerdo con el estándar Dolby Digital Plus.

45 La señal de audio multicanal se puede representar como un grupo básico de canales para renderizar la señal de audio multicanal de acuerdo con una configuración básica de canales, y como un grupo de ampliación de canales, los cuales – en combinación con el grupo básico – se utilizan para una renderización de la señal de audio multicanal de acuerdo con una configuración ampliada de canales. Típicamente, la configuración básica de canales y la configuración ampliada de canales son diferentes entre sí. En particular, la configuración ampliada de canales comprende típicamente un mayor número de canales que la configuración básica de canales. A modo de ejemplo, la configuración básica de canales y el grupo básico de canales pueden comprender N canales. La configuración ampliada de canales puede comprender M canales, siendo M mayor que N. En dichos casos, el grupo de ampliación de canales puede comprender uno o más canales de ampliación con el fin de ampliar la

configuración básica de canales a la configuración ampliada de canales. Además, el grupo de ampliación de canales puede comprender uno o más canales de sustitución que sustituyan uno o más canales del grupo básico de canales cuando se renderizan en la configuración ampliada de canales.

5 El Informe de Búsqueda Internacional emitido junto con la presente solicitud citada con número de publicación de Solicitud de Patente Internacional número WO0187015, de aquí en adelante el "documento '015", y el número de patente de los Estados Unidos US5978762, de aquí en adelante el "documento '762", como "documentos de particular importancia". El documento '015 divulga un formato de audio multicanal que pretende proporcionar una mezcla discreta y compatible hacia atrás para sonido envolvente, frontal u otros canales de audio discretos en el cine, un cine en casa o entornos de música. Se mezclan señales de audio discretas adicionales con los canales de audio discretos existentes en un formato predeterminado como, por ejemplo, el formato de audio 5.1. Además, estos canales de audio discretos adicionales se codifican y añaden al formato predeterminado como bits de ampliación en el flujo de bits. El documento '762 divulga un codificador de audio de subbanda que utiliza filtros de reconstrucción perfecta/no perfecta, codificación de subbanda predictiva/no predictiva, análisis de transición, y asignación de bits psicoacústico/error cuadrático medio mínimo (mmse) sobre el tiempo, frecuencia y los canales múltiples de audio para codificar/decodificar un flujo de datos para generar audio de alta fidelidad reconstruido. El codificador de audio enmarca la señal de audio multicanal de modo que el tamaño de trama, esto es el número de bytes, se limita para que encaje en un rango deseado, y formatea los datos codificados de modo que las subtramas individuales se pueden reproducir según se reciben reduciendo de este modo la latencia.

20 En un modo de realización, la señal de audio multicanal es una señal de audio 7.1 que comprende un canal central, frontal izquierdo, frontal derecho, envolvente izquierdo, envolvente derecho, envolvente izquierdo posterior, envolvente derecho posterior y un canal de efectos de baja frecuencia. En dichos casos, el grupo básico de canales puede comprender los canales central, frontal izquierdo y frontal derecho, así como un canal envolvente izquierdo downmixed (mezclado para reducir canales) y un canal envolvente derecho downmixed, con el fin de permitir la renderización de una señal de audio multicanal en una configuración de canales 5.1 (la configuración básica). El canal envolvente izquierdo downmixed y el canal envolvente derecho downmixed se pueden obtener a partir de los canales envolvente izquierdo, envolvente derecho, envolvente izquierdo posterior y envolvente derecho posterior (por ejemplo, como una suma de algunos o todos los canales envolvente izquierdo, envolvente derecho, envolvente izquierdo posterior, envolvente derecho posterior). El grupo de ampliación de canales puede comprender los canales envolvente izquierdo, envolvente derecho, izquierdo posterior y derecho posterior, con el fin de permitir la renderización de los canales básicos y los canales de ampliación en una configuración de canales 7.1 (la configuración ampliada de canales). Se debe observar que la configuración de canales 7.1 mencionada más arriba es únicamente un ejemplo de configuraciones posibles de canales 7.1. A modo de ejemplo, los canales envolvente izquierdo y envolvente derecho se pueden etiquetar como canales laterales izquierdo y derecho (situados a +/-90 grados con respecto a una línea central frente a la cabeza del oyente). De una forma perecida, los canales posteriores se pueden denominar canales envolventes posteriores izquierdo y derecho.

40 El codificador de audio comprende un codificador básico configurado para codificar el grupo básico de canales de acuerdo con una tasa de datos IS (flujo secundario independiente), con el fin de producir un flujo secundario independiente. El flujo secundario independiente puede comprender una secuencia de tramas IS que comprende datos codificados representativos del grupo básico de canales. Además, el codificador de audio comprende un codificador de ampliación configurado para codificar el grupo de ampliación de canales de acuerdo con una tasa de datos DS (flujo secundario dependiente), con el fin de producir un flujo secundario dependiente. El flujo secundario dependiente puede comprender una secuencia de tramas DS que comprende datos codificados representativos del grupo de ampliación de canales. En un modo de realización el codificador básico y/o el codificador de ampliación están configurados para realizar una codificación Dolby Digital Plus.

50 Además, el codificador de audio comprende una unidad de control de tasa configurada para adaptar de forma regular la tasa de datos IS y la tasa de datos DS en función de un indicador de calidad de codificación IS instantáneo para el grupo básico de canales y/o en función de un indicador de calidad de codificación DS instantáneo para el grupo de ampliación de canales. La tasa de datos IS y la tasa de datos DS se pueden adaptar de modo que la suma de la tasa de datos IS y la tasa de datos DS se corresponda sustancialmente (por ejemplo, sea igual a) la tasa de datos total disponible. En particular, la unidad de control de tasa se puede configurar para determinar la tasa de datos IS y la tasa de datos DS de modo que se reduzca la diferencia entre el indicador de calidad de codificación IS instantáneo y el indicador de calidad de codificación DS instantáneo. Esto puede dar como resultado una calidad de audio mejorada para la combinación del grupo básico y el grupo ampliado de canales bajo la restricción de una tasa total de bits disponible.

60 El indicador de calidad de codificación IS instantáneo y/o el indicador de calidad de codificación DS instantáneo pueden ser indicativos de una complejidad de codificación de la señal de audio multicanal en un instante de tiempo concreto. A modo de ejemplo, la señal de audio multicanal se puede representar como una secuencia de tramas de audio. En dichos casos, el indicador de calidad de codificación IS instantáneo y/o el indicador de calidad de codificación DS instantáneo pueden ser indicativos de una complejidad para la codificación de una o

más tramas de audio de la señal de audio multicanal. Así pues, el indicador de calidad de codificación IS instantáneo y/o el indicador de calidad de codificación DS instantáneo pueden variar de trama a trama. Por lo tanto, la unidad de control de tasa puede estar configurada para adaptar la tasa de datos IS y la tasa de datos DS de trama a trama (en función de la variación del indicador de calidad de codificación IS instantáneo y/o el indicador de calidad de codificación DS instantáneo). En otras palabras, la unidad de control de tasa puede estar configurada para adaptar la tasa de datos IS y la tasa de datos DS para cada trama de la secuencia de tramas de la señal de audio multicanal.

El indicador de calidad de codificación IS instantáneo y/o el indicador de calidad de codificación DS instantáneo pueden comprender un parámetro de codificación del codificador básico y/o del codificador de ampliación, respectivamente. A modo de ejemplo, en el caso de codificación Dolby Digital Plus, el indicador de calidad de codificación IS instantáneo y/o el indicador de calidad de codificación DS instantáneo pueden comprender la variación de la SNR instantánea del codificador básico y/o el codificador de ampliación, respectivamente. Alternativamente o además, el indicador de calidad de codificación IS puede comprender uno o más de los siguientes: una entropía perceptual de una (primera) trama actual del grupo básico; una tonalidad de la primera trama del grupo básico; una característica transitoria de la primera trama del grupo básico; un ancho de banda espectral de la primera trama del grupo básico; una presencia de transitorios en la primera trama del grupo básico; un grado de correlación entre canales del grupo básico; y una energía de la primera trama del grupo básico. De forma parecida, el indicador de calidad de codificación DS puede comprender uno o más de los siguientes: una entropía perceptual de la primera trama del grupo de ampliación; una tonalidad de la primera trama del grupo de ampliación; una característica transitoria de la primera trama del grupo de ampliación; un ancho de banda espectral de la primera trama del grupo de ampliación; una presencia de transitorios en la primera trama del grupo de ampliación; un grado de correlación entre canales del grupo de ampliación; y una energía de la primera trama del grupo de ampliación.

En el caso de un codificador de audio basado en tramas, el codificador básico puede estar configurado para determinar una secuencia de tramas IS para la secuencia de tramas de la señal multicanal. De forma parecida, el codificador de ampliación puede estar configurado para determinar una secuencia de tramas DS para la secuencia de tramas de la señal multicanal. En dichos casos, el indicador de calidad de codificación IS puede comprender una secuencia de indicadores de calidad de codificación IS para la secuencia de tramas IS correspondiente. De forma parecida, el indicador de calidad de codificación DS puede comprender una secuencia de indicadores de calidad de codificación DS para la secuencia de tramas DS correspondiente. La unidad de control de tasa puede estar entonces configurada para determinar la tasa de datos IS para una trama IS de la secuencia de tramas IS y la tasa de datos DS para una trama DS de la secuencia de tramas DS basadas en al menos un indicador de la secuencia de indicadores de calidad de codificación IS y/o basada en al menos un indicador de la secuencia de indicadores de calidad de codificación DS. La tasa de datos IS para una trama IS y la tasa de datos DS para la trama DS correspondiente se pueden adaptar de modo que la suma de la tasa de datos IS para la trama IS y la tasa de datos DS para la trama DS correspondiente sea sustancialmente la tasa de datos total disponible para una trama de audio de la señal de audio multicanal.

El codificador puede comprender una unidad de determinación de la dificultad de codificación configurada para determinar el indicador de calidad de codificación IS en función de una primera trama del grupo básico de canales, y/o para determinar el indicador de calidad de codificación DS en función de una primera trama del grupo de ampliación de canales correspondiente. La primera trama puede ser la trama para la que se va a determinar la tasa de datos IS y la tasa de datos DS. Así pues, la unidad de determinación de la dificultad de codificación se puede configurar para analizar la trama a codificar del grupo básico de canales y/o del grupo de ampliación de canales y determinar los indicadores de calidad de codificación IS/DS que puede utilizar la unidad de control de tasa para adaptar la tasa de datos IS y la tasa de datos DS para la trama a codificar.

El codificador básico puede comprender una unidad de transformación configurada para determinar un bloque de coeficientes de transformación básico para la primera trama del grupo básico. De forma parecida, el codificador de ampliación puede comprender una unidad de transformación configurada para determinar un bloque de coeficientes de transformación de ampliación para la primera trama del grupo de ampliación correspondiente. Las unidades de transformación se pueden configurar para aplicar una transformación Tiempo-A-Frecuencia, por ejemplo, una Transformación de Coseno Discreta Modificada (MDCT). La primera trama se puede dividir en una pluralidad de bloques (por ejemplo, con solapamiento) y las unidades de transformación se pueden configurar para transformar un bloque de muestras obtenido a partir de las primeras tramas respectivas.

Además, el codificador básico puede comprender una unidad de codificación de coma flotante configurada para determinar un bloque de exponentes básico y un bloque de mantisas básico a partir del bloque de coeficientes de transformación básico. De forma parecida, el codificador de ampliación puede comprender una unidad de codificación de coma flotante configurada para determinar un bloque de exponentes de ampliación y un bloque de mantisas de ampliación a partir del bloque de coeficientes de transformación de ampliación. La unidad de control de tasa puede estar configurada para determinar el número total de bits de mantisa disponibles para codificar el bloque de mantisas básico y el bloque de mantisas de ampliación, basados en la tasa de datos total disponible.

Con este propósito, la unidad de control de tasa puede considerar un número total de bits disponible derivado de la tasa de datos total disponible y restar el número de bits del número total de bits disponibles que se utilizan para la codificación de los exponentes y/u otros parámetros de codificación que no están relacionados con las mantisas. Los bits restantes pueden ser el número total de bits de mantisa disponibles. Además, la unidad de control de tasa puede estar configurada para distribuir el número total de bits de mantisa disponibles al bloque de mantisas básico y al bloque de mantisas de ampliación, basados en el indicador de calidad de codificación IS instantáneo y el indicador de calidad de codificación DS instantáneo, adaptando de este modo la tasa de datos IS y la tasa de datos DS.

En particular, la unidad de control de tasa puede estar configurada para determinar una distribución de densidad espectral de potencia (PSD) básica para el bloque de coeficientes de transformación básico. De forma parecida, la unidad de control de tasa puede determinar una distribución de la PSD de ampliación para el bloque de coeficientes de transformación de ampliación. Además, la unidad de control de tasa puede determinar una curva de enmascaramiento básica para el bloque de coeficientes de transformación básico y una curva de enmascaramiento de ampliación para el bloque de coeficientes de transformación de ampliación. La unidad de control de tasa puede utilizar la distribución de la PSD básica, la distribución de la PSD de ampliación, la curva de enmascaramiento básica y la curva de enmascaramiento de ampliación para distribuir el número total de bits de mantisa disponibles al bloque de mantisas básico y al bloque de mantisas de ampliación.

Incluso más en particular, la unidad de control de tasa puede estar configurada para determinar una curva de enmascaramiento básica de desplazamiento desplazando la curva de enmascaramiento básica utilizando un desplazamiento IS (también denominado "variación de la SNR del IS"). De forma parecida, la unidad de control de tasa puede estar configurada para determinar una curva de enmascaramiento de ampliación de desplazamiento desplazando la curva de enmascaramiento de ampliación utilizando un desplazamiento DS (también denominado "variación de la SNR del DS"). Además, la unidad de control de tasa puede estar configurada para comparar la distribución de la PSD básica y la curva de enmascaramiento básica de desplazamiento, y asignar el número de bits de mantisa básico al bloque de mantisas básico, en función del resultado de la comparación. Además, la unidad de control de tasa puede estar configurada para comparar la distribución de la PSD de ampliación y la curva de enmascaramiento de ampliación de desplazamiento, y asignar un número de bits de mantisa de ampliación al bloque de mantisas de ampliación, en función del resultado de la comparación.

Se puede determinar el número total de bits de mantisa asignados como la suma del número de bits de mantisa básico y el número de bits de mantisa de ampliación. La unidad de control de tasa puede entonces estar configurada para ajustar el desplazamiento IS y el desplazamiento DS de modo que la diferencia del número total de bits de mantisa asignados y el número total de bits de mantisa disponibles estén por debajo de un umbral de bits predeterminado. Con este propósito, la unidad de control de tasa puede hacer uso de un esquema de búsqueda iterativo, con el fin de determinar el desplazamiento IS y el desplazamiento DS que cumpla la condición mencionada anteriormente. En particular, la unidad de control de tasa puede estar configurada para ajustar el desplazamiento IS y el desplazamiento DS de modo que el desplazamiento IS y el desplazamiento DS sean iguales para la secuencia de tramas de la señal de audio multicanal, adaptando de este modo la tasa de datos IS y la tasa de datos DS para cada trama de la secuencia de tramas de la señal de audio multicanal. Como ya se ha indicado, el indicador de calidad de codificación IS instantáneo puede comprender el desplazamiento IS y/o el indicador de calidad de codificación DS instantáneo puede comprender el desplazamiento DS.

Así pues, el codificador de audio puede estar configurado para realizar un proceso de asignación conjunta de bits para el grupo básico de canales y para el grupo de ampliación de canales. En otras palabras, el codificador básico y el codificador de ampliación pueden utilizar un proceso de asignación de bits combinado, adaptando de este modo la tasa de datos IS y la tasa de datos DS de forma regular (por ejemplo, trama a trama).

La unidad de control de tasa puede estar configurada para determinar el desplazamiento IS y el desplazamiento DS para la primera trama de la señal de audio multicanal. A modo de ejemplo, el desplazamiento IS y el desplazamiento DS se pueden extraer de una trama IS y una trama DS, respectivamente, y la salida del codificador básico y el codificador de ampliación, respectivamente. Además, la unidad de control de tasa puede estar configurada para ajustar la tasa de datos IS y la tasa de datos DS para codificar una segunda trama de la señal de audio multicanal, en función del desplazamiento IS y del desplazamiento DS de la primera trama. Típicamente, la primera trama precede a la segunda trama. En particular, la segunda trama puede seguir directamente a la primera trama, sin ninguna trama intermedia entre la primera y segunda tramas. En otras palabras, el desplazamiento IS y el desplazamiento DS utilizado para una primera trama precedente, y posiblemente para una directamente precedente, se puede utilizar para determinar la tasa de datos IS y la tasa de datos DS para codificar la segunda trama actual. En aún otras palabras, se propone utilizar una indicación de la calidad de la codificación de la primera trama precedente para ajustar la tasa de datos IS y la tasa de datos DS para codificar la segunda trama actual.

En particular, la unidad de control de tasa puede estar configurada para ajustar la tasa de datos IS y la tasa de datos DS para codificar la segunda trama de la señal de audio multicanal, de modo que se reduce la diferencia entre el desplazamiento IS y el desplazamiento DS (por ejemplo, se reduce en promedio a lo largo de una pluralidad de tramas de audio). Con este propósito se puede utilizar un bucle de regulación, en donde el bucle de regulación se adapta para regular la diferencia entre el desplazamiento IS y el desplazamiento DS. A modo de ejemplo, la unidad de control de tasa puede estar configurada para determinar la diferencia entre el desplazamiento IS y el desplazamiento DS para la primera trama. Además, la unidad de control de tasa puede estar configurada para modificar la tasa de datos IS para la segunda trama en comparación con la tasa de datos IS de la primera trama mediante un ajuste de tasa, y modificar la tasa de datos DS para la segunda trama con respecto a la tasa de datos DS de la primera trama mediante un ajuste de tasa negativo. El ajuste de tasa (en particular el signo del ajuste de tasa) puede depender de la diferencia determinada.

El codificador de audio puede estar configurado para codificar una pluralidad de señales de audio multicanal (asociadas). Cada una de las señales de audio multicanal de la pluralidad de señales puede, por ejemplo, corresponderse con un programa de difusión diferente o con un idioma diferente. Esto puede ser beneficioso para los Discos de Vídeo Digital (DVD) que proporcionan una pluralidad de señales de audio multicanal diferentes (por ejemplo, idiomas diferentes) para una película. La pluralidad de señales de audio multicanal (asociadas) pueden tener tramas correspondientes (representando intervalos de tiempo correspondientes de la pluralidad de señales de audio multicanal asociadas). Cada una de la pluralidad de señales de audio multicanal se puede representar como un grupo básico de canales para renderizar la señal de audio multicanal respectiva en función de la configuración básica de canales, con el fin de producir una pluralidad de grupos básicos. Además, cada una de la pluralidad de señales de audio multicanal se puede representar como un grupo de ampliación de canales para renderizar la señal de audio multicanal respectiva en función de la configuración de ampliación de canales, los cuales – en combinación con el grupo básico – permiten renderizar la señal de audio multicanal respectiva en función de la configuración ampliada de canales, con el fin de producir una pluralidad de grupos de ampliación.

El codificador de audio puede comprender una pluralidad de codificadores básicos para codificar la pluralidad de grupos básicos de acuerdo con una pluralidad de tasas de datos IS, con el fin de producir una pluralidad de IS respectivos. Se debe observar que se puede configurar un codificador básico combinado para codificar la pluralidad de grupos básicos con el fin de producir la pluralidad de IS respectivos. De forma parecida, el codificador de audio puede comprender una pluralidad de codificadores de ampliación para codificar la pluralidad de grupos de ampliación de acuerdo con una pluralidad de tasas de datos DS, con el fin de producir una pluralidad de DS respectivos. Se debe observar que se puede configurar un codificador de ampliación combinado para codificar la pluralidad de grupos de ampliación con el fin de producir la pluralidad de DS respectivos.

La unidad de control de tasa puede entonces estar configurada para adaptar de forma regular la pluralidad de tasas de datos IS y la pluralidad de tasas de datos DS basándose en uno o más indicadores de calidad de codificación IS instantáneo para la pluralidad de grupos básicos de canales y/o basándose en uno o más indicadores de calidad de codificación DS instantáneo para la pluralidad de grupos de ampliación de canales, de modo que la suma de la pluralidad de tasas de datos IS y la pluralidad de tasas de datos DS se corresponde sustancialmente con la tasa de datos total disponible. Los indicadores de calidad de codificación instantáneos pueden ser, por ejemplo, las variaciones de SNR para codificar la pluralidad de grupos básicos/grupos de ampliación. En particular, la unidad de control de tasa puede estar configurada para aplicar los esquemas de asignación de tasa/asignación de bits descritos en el presente documento a una pluralidad de IS y una pluralidad de DS correspondientes. Así pues, cada IS y cada DS puede tener tasas de datos variables (por ejemplo, variables de trama a trama), mientras que se mantiene constante la tasa de bits global para la pluralidad de señales de audio multicanal codificadas (esto es, para la pluralidad de IS y DS).

De acuerdo con otro aspecto, se describe un método para codificar una señal de audio multicanal en función de una tasa de datos total disponible. La señal de audio multicanal se puede representar como un grupo básico de canales para renderizar la señal de audio multicanal de acuerdo con una configuración básica de canales, y como un grupo de ampliación de canales, el cual – en combinación con el grupo básico – se utiliza para renderizar la señal de audio multicanal de acuerdo con una configuración ampliada de canales. La configuración básica de canales y la configuración ampliada de canales pueden ser diferentes entre sí.

El método puede comprender codificar el grupo básico de canales en función de una tasa de datos IS, con el fin de producir un flujo secundario independiente. El método puede comprender, además, codificar el grupo de ampliación de canales en función de una tasa de datos DS, con el fin de producir un flujo secundario dependiente. Además, el método puede comprender adaptar de forma regular la tasa de datos IS y la tasa de datos DS en función de un indicador de calidad de codificación IS instantáneo para el grupo básico de canales y/o en función de un indicador de calidad de codificación DS instantáneo para el grupo de ampliación de canales, de modo que la suma de la tasa de datos IS y la tasa de datos DS se corresponde sustancialmente con la tasa de datos total disponible.

El método puede comprender, además, determinar el indicador de calidad de codificación IS en función de una muestra del grupo básico de canales, y/o determinar el indicador de calidad de codificación DS en función de una muestra del grupo de ampliación de canales. La muestra del grupo básico/grupo de ampliación puede ser, por ejemplo, una o más tramas del grupo básico/grupo de ampliación. Así pues, el indicador de calidad de codificación IS y/o el indicador de calidad de codificación DS se pueden determinar en función de la señal de entrada de un codificador de audio. A modo de ejemplo, los indicadores de calidad de codificación se pueden determinar en función de una entropía perceptual de la muestra del grupo básico/de ampliación; en función de una tonalidad de la muestra del grupo básico/de ampliación; en función de una característica transitoria de la muestra del grupo básico/de ampliación; en función del ancho de banda espectral de la muestra del grupo básico/de ampliación; una presencia de transitorios en la muestra del grupo básico/de ampliación; un grado de correlación entre canales de la muestra del grupo básico/de ampliación; y/o en función de una energía de la muestra del grupo básico/de ampliación.

Alternativamente o además, el indicador de calidad de codificación IS puede ser indicativo de una calidad perceptual de una muestra del flujo secundario independiente (esto es, de la calidad perceptual de la señal codificada). De forma parecida, el indicador de calidad de codificación DS puede ser indicativo de una calidad perceptual de una muestra del flujo secundario dependiente (esto es, de la calidad perceptual de la señal codificada).

En dichos casos, la adaptación de la tasa de datos IS y la tasa de datos DS puede comprender adaptar la tasa de datos IS y la tasa de datos DS para codificar la muestra del flujo secundario independiente y la muestra del flujo secundario dependiente, de modo que la diferencia absoluta entre el indicador de calidad de codificación IS y el indicador de calidad de codificación DS se encuentre por debajo de un umbral de diferencia. A modo de ejemplo, el umbral de diferencia puede ser sustancialmente cero. Como se ha resumido más arriba, la adaptación de la tasa de datos IS y la tasa de datos DS se puede conseguir utilizando una asignación de bits conjunta al codificar la muestra del flujo secundario independiente y la muestra del flujo secundario dependiente.

Alternativamente, la adaptación de la tasa de datos IS y la tasa de datos DS puede comprender adaptar la tasa de datos IS y la tasa de datos DS para codificar una muestra adicional del flujo secundario independiente y una muestra adicional correspondiente del flujo secundario dependiente, basado en la diferencia entre el indicador de calidad de codificación IS y el indicador de calidad de codificación DS. Las muestras adicionales de los grupos básico y de ampliación pueden ser posteriores a las muestras de los grupos básico y de ampliación. A modo de ejemplo, las muestras adicionales de los grupos básico y de ampliación pueden seguir directamente, sin muestras intermedias, las muestras de los grupos básico y de ampliación. Así pues, la tasa de datos IS y la tasa de datos DS se pueden adaptar muestra a muestra, basado en la realimentación del/de los indicador(es) de calidad de codificación IS/DS.

De acuerdo con un aspecto adicional, se describe un programa de software. El programa de software se puede adaptar para su ejecución en un procesador y para llevar a cabo los pasos del método resumido en el presente documento cuando se ejecuta en el procesador.

De acuerdo con otro aspecto, se describe un medio de almacenamiento. El medio de almacenamiento puede comprender un programa de software adaptado para su ejecución en un procesador y para llevar a cabo los pasos del método resumido en el presente documento cuando se ejecuta en el procesador.

De acuerdo un aspecto adicional, se describe un producto de programa informático. El producto de programa informático puede comprender instrucciones ejecutables para llevar a cabo los pasos del método resumido en el presente documento cuando se ejecuta en un ordenador.

Se debe observar que los métodos y sistemas que incluyen sus modos de realización preferidos tal como se han resumido en la presente solicitud de patente se pueden utilizar por sí mismos o en combinación con otros métodos y sistemas divulgados en este documento. Además, todos los aspectos de los métodos y sistemas resumidos en la presente solicitud de patente se pueden combinar de forma arbitraria. En particular, se pueden combinar entre sí de forma arbitraria las características de las reivindicaciones. Además, aunque los pasos de los métodos se pueden proporcionar en un orden concreto, los pasos se pueden combinar o llevar a cabo fuera del orden proporcionado.

## Descripción de las figuras

A continuación, se explica la invención de una forma ilustrativa haciendo referencia a los dibujos adjuntos, en donde

la Fig. 1a muestra un diagrama de bloques de alto nivel de un codificador de audio multicanal de ejemplo;

la Fig. 1b muestra una secuencia de ejemplo de tramas codificadas;

la Fig. 2a muestra un diagrama de bloques de alto nivel de decodificadores de audio multicanal de ejemplo;

la Fig. 2b muestra una disposición de ejemplo de altavoces para una señal de audio multicanal 7.1;

la Fig. 3 ilustra un diagrama de bloques de componentes de ejemplo de un codificador de audio multicanal;

las Fig. 4a a 4e ilustran aspectos particulares de un codificador de audio multicanal de ejemplo;

5 la Fig. 5a muestra un diagrama de bloques de un codificador de audio multicanal de ejemplo que comprende un control de tasa conjunto;

la Fig. 5b muestra un diagrama de flujo de un esquema de codificación multicanal de ejemplo;

la Fig. 5c muestra un diagrama de bloques de un codificador de audio multicanal de ejemplo adicional que comprende un control de tasa conjunto; y

10 la Fig. 6 muestra un diagrama de bloques de otro codificador de audio multicanal de ejemplo que comprende un control de tasa conjunto.

### Descripción detallada de la invención

15 Tal como se ha resumido en la sección de introducción, es deseable proporcionar sistemas codec de audio multicanal que generen flujos de bits que sean compatibles hacia atrás en relación al número de canales que son decodificados por un decodificador de audio multicanal concreto. En particular, es deseable codificar una señal de audio multicanal M.1 de modo que pueda ser decodificada por un decodificador de audio multicanal N.1, siendo  $N < M$ . A modo de ejemplo, es deseable codificar una señal de audio 7.1 de modo que pueda ser decodificada por un decodificador de audio 5.1. Con el fin de permitir compatibilidad hacia atrás los sistemas codec de audio multicanal típicamente codifican una señal de audio multicanal M.1 en un flujo (secundario) independiente ("IS"), que comprende un número reducido de canales (por ejemplo, N.1 canales), y en uno o más flujos (secundarios) dependientes ("DS"), que comprenden canales de sustitución y/o ampliación con el fin de decodificar y renderizar la señal de audio M.1 completa.

25 En este contexto, es deseable permitir una codificación eficiente del IS y los uno o más DS. El presente documento describe métodos y sistemas que permiten la codificación eficiente de un IS y uno o más DS, mientras que al mismo tiempo se mantiene la independencia del IS y los uno o más DS con el fin de mantener la compatibilidad hacia atrás del sistema codec de audio multicanal. Los métodos y sistemas se describen basándose en el sistema codec Dolby Digital Plus (DD+) (también denominado AC-3 mejorado). El sistema codec DD+ se especifica en el "Digital Audio Compression Standard (Estándar de Compresión de Audio Digital) (AC-3, E-AC-3)" del Comité de Sistemas de Televisión Avanzada (ATSC), Documento A/52: 2010, de fecha 22 de noviembre de 2010. Sin embargo, se debe observar que los métodos y sistemas descritos en el presente documento son aplicables en general y se pueden aplicar a otros sistemas codec de audio que codifican señales de audio multicanal en una pluralidad de flujos secundarios.

35 Las configuraciones multicanal (y las señales de audio multicanal) utilizadas frecuentemente son la configuración 7.1 y la configuración 5.1. Una configuración multicanal 5.1 comprende típicamente un canal L (frontal izquierdo), un C (frontal central), un R (frontal derecho), un Ls (envolvente izquierdo), un Rs (envolvente derecho), y un LFE (Efectos de Baja Frecuencia). Una configuración multicanal 7.1 comprende, además, un canal Lb (envolvente izquierdo posterior) y un Rb (envolvente derecho posterior). En la Fig. 2b se ilustra un ejemplo de configuración multicanal 7.1. Con el fin de transmitir canales 7.1 en DD+, se utilizan dos flujos secundarios. El primer flujo secundario (denominado flujo secundario independiente, "IS") comprende una mezcla de canales 5.1, y el segundo flujo secundario (denominado flujo secundario dependiente, "DS") comprende canales de ampliación y canales de sustitución. Por ejemplo, con el fin de codificar y transmitir una señal de audio multicanal 7.1 con canales Lb y Rb envolventes posteriores, el flujo secundario independiente transporta los canales L (frontal izquierdo), C (frontal central), R (frontal derecho), Lst (envolvente izquierdo downmixed), Rst (envolvente derecho downmixed), LFE (Efectos de Baja Frecuencia), y el canal dependiente transporta los canales de ampliación Lb (envolvente izquierdo posterior), Rb (envolvente derecho posterior) y los canales de sustitución Ls (envolvente izquierdo) y Rs (envolvente derecho). Cuando se realiza una decodificación de señal 7.1 completa, los canales Ls y Rs del flujo secundario dependiente sustituyen los canales Lst y Rst del flujo secundario independiente.

50 La Fig. 1a muestra un diagrama de bloques de alto nivel de un codificador 100 de audio multicanal 7.1 DD+ de ejemplo que ilustra la relación entre canales 5.1 y 7.1. Los canales 101 de audio siete (7) más uno (1) (L, C, R, Ls, Lb, Rs y Rb más LFE) de la señal de audio multicanal se dividen en dos grupos de canales de audio. Un grupo básico 121 de canales comprende los canales de audio L, C, R y LFE, así como los canales envolventes downmixed Lst 102 y Rst 103 los cuales se obtienen típicamente a partir de los canales envolvente 7.1 Ls, Rs y los canales posteriores 7.1 Lb, Rb. A modo de ejemplo, los canales 102, 103 envolventes downmixed se obtienen añadiendo algunos o todos los canales Lb y Rb y los canales envolventes 7.1 Ls, Rs en una unidad 109 de



downmix. Se debe observar que los canales envolventes downmixed Lst 102 y Rst 103 se pueden determinar de otras formas. A modo de ejemplo, los canales envolventes downmixed Lst 102 y Rst 103 se pueden determinar directamente a partir de los dos canales 7.1, por ejemplo, los canales envolventes 7.1 Ls, Rs.

5 El grupo básico 121 de canales se codifica en un codificador 105 de audio 5.1 DD+, con el fin de producir el flujo secundario independiente ("IS") 110 que se transmite en una trama principal DD+ 151 (ver Fig. 1b). La trama principal 151 también se denomina trama IS. Un segundo grupo 122 de canales de audio comprende los canales envolventes 7.1 Ls, Rs y los canales envolventes posteriores 7.1 Lb, Rb. El segundo grupo 122 de canales se codifica en un codificador 106 de audio 4.0 DD+, con el fin de producir un flujo secundario dependiente ("DS") 120 que se transmite en una o más tramas 152, 153 de ampliación DD+ (ver Fig. 1b). El segundo grupo 122 de canales se denomina en la presente solicitud grupo 122 de ampliación de canales y las tramas 152, 153 se denominan tramas DS 152, 153.

15 La Fig. 1b ilustra una secuencia 150 de ejemplo de tramas 151, 152, 153, 161, 162 de audio codificadas. El ejemplo ilustrado comprende dos flujos secundarios independiente IS0 e IS1 que comprenden las tramas IS 151 y 161, respectivamente. Se pueden utilizar múltiples IS (y los DS respectivos) para proporcionar múltiples señales de audio asociadas (por ejemplo, para diferentes idiomas de una película o para diferentes programas). Cada uno de los flujos secundarios independientes comprende uno o más flujos secundarios dependientes DS0, DS1, respectivamente. Cada uno de los flujos secundarios dependientes comprende las respectivas tramas DS 152, 153 y 162. Además, la Fig. 1b indica la duración temporal 170 de una trama de audio completa de la señal de audio multicanal. La duración temporal 170 de la trama de audio puede ser 32ms (por ejemplo, a una tasa de muestreo  $f_s = 48\text{kHz}$ ). En otras palabras, la Fig. 1b indica la duración en tiempo 170 de una trama de audio que se codifica en una o más tramas IS 151, 161 y las respectivas tramas DS 152, 153, 162.

20 La Fig. 2a ilustra diagramas de bloques de alto nivel de ejemplo de sistemas decodificadores multicanal 200, 210 de ejemplo. En particular, la Fig. 2a muestra un sistema decodificador multicanal 5.1 200 de ejemplo que recibe el IS 201 codificado que comprende el grupo básico 121 de canales codificado. El IS 201 codificado se obtiene a partir de las tramas IS 151 de un flujo de bits recibido (por ejemplo, utilizando un desmultiplexador que no se muestra). Las tramas IS 151 comprenden el grupo básico 121 de canales codificado y se decodifican utilizando un decodificador 205 multicanal 5.1, con el fin de producir una señal de audio multicanal 5.1 decodificada que comprende el grupo básico 221 de canales decodificado. Además, la Fig. 2a muestra un sistema decodificador 210 multicanal 7.1 de ejemplo que recibe el IS 201 codificado que comprende el grupo básico 121 de canales codificado y el DS 202 codificado que comprende el grupo 122 de ampliación de canales codificado. Tal como se ha resumido más arriba, el IS 201 codificado puede obtenerse a partir de las tramas IS 151 y el DS 202 codificado se puede obtener a partir de las tramas DS 152, 153 del flujo de bits recibido (por ejemplo, utilizando un desmultiplexador que no se muestra). Después de la decodificación, se obtiene una señal de audio multicanal 7.1 decodificada que comprende el grupo básico 221 de canales decodificado y un grupo 222 de ampliación de canales decodificado. Se debe observar que los canales envolventes downmixed Lst, Rst 211 se pueden descartar, ya que el decodificador 215 multicanal 7.1 utiliza en su lugar el grupo 222 de ampliación de canales decodificado. En la configuración multicanal 230 de la Fig. 2b se muestran las posiciones 232 de renderización típicas de una señal de audio multicanal 7.1, la cual también ilustra una posición 231 de ejemplo de un oyente y una posición 233 de ejemplo de una pantalla para la renderización de vídeo.

40 Actualmente, la codificación de las señales de audio de canales 7.1 en DD+ se realiza mediante un primer codificador DD+ 105 de canales 5.1 principal y un segundo codificador DD+ 106. El primer codificador DD+ 105 codifica los canales 5.1 del grupo básico 121 (y puede por lo tanto denominarse codificador de canales 5.1) y el segundo codificador DD+ 106 codifica los canales 4.0 del grupo 122 de ampliación (y puede por lo tanto denominarse codificador de canales 4.0). Los codificadores 105, 106 para el grupo básico 121 y el grupo 122 de ampliación de canales típicamente son independientes. Cada uno de los dos codificadores 105, 106 se proporciona con una tasa de datos, la cual se corresponde con una porción fija de la tasa de datos total disponible. En otras palabras, el codificador 105 para el IS y el codificador 106 para el DS se proporcionan con una fracción fija de la tasa de datos total disponible (por ejemplo, X% de la tasa de datos total disponible para el codificador IS 105 (denominada "tasa de datos IS") y  $100\% - X\%$  de la tasa de datos total disponible para el codificador DS 106 (denominada "tasa de datos DS"), por ejemplo,  $X=50$ ). Utilizando las tasas de datos asignadas respectivamente (esto es, la tasa de datos IS y la tasa de datos DS), el codificador IS 105 y el codificador DS 106 realizan una codificación independiente del grupo básico 121 de canales y del grupo 122 de ampliación de canales, respectivamente.

55 En el presente documento se propone la creación de una dependencia entre el codificador IS 105 y el codificador DS 106 con el fin de aumentar la eficiencia del codificador 100 de audio multicanal conjunto. En particular, se propone proporcionar una asignación adaptativa de la tasa de datos IS y la tasa de datos DS en función de las características o condiciones el grupo básico 121 de canales y el grupo 122 de ampliación de canales.

A continuación, se describen detalles adicionales en relación con los componentes del codificador IS 105 y el codificador DS 106 en el contexto de la Fig. 3, la cual muestra un diagrama de bloques de un codificador

multicanal DD+ 300 de ejemplo. El codificador IS 105 y/o el codificador DS 106 se pueden materializar en el codificador multicanal DD+ 300 de la Fig. 3. Después de describir los componentes del decodificador 300, se describe cómo se puede adaptar el codificador multicanal 300 para permitir la asignación adaptativa de la tasa de datos IS y la tasa de datos DS mencionadas más arriba.

5 El codificador multicanal 300 recibe flujos 311 de muestras PCM correspondientes a los diferentes canales de la señal de entrada multicanal (por ejemplo, la señal de entrada 5.1). Los flujos 311 de muestras PCM se pueden disponer en tramas de muestras PCM. Cada una de las tramas puede comprender un número predeterminado de muestras PCM (por ejemplo, 1536 muestras) de un canal concreto de la señal de audio multicanal. Así pues, para cada segmento de tiempo de la señal de audio multicanal, se proporciona una trama de audio diferente para cada uno de los diferentes canales de la señal de audio multicanal. El codificador 300 de audio multicanal se describe a continuación para un canal concreto de la señal de audio multicanal. Sin embargo, se debe observar que la trama AC-3 resultante 318 comprende típicamente los datos codificados de todos los canales de la señal de audio multicanal.

15 Una trama de audio que comprende muestras PCM 311 se puede filtrar en una unidad 301 de acondicionamiento de la señal de entrada. Posteriormente, las muestras 311 (filtradas) se pueden transformar del dominio del tiempo al dominio de la frecuencia en una unidad 302 de Transformación Tiempo Frecuencia. Con este propósito, la trama de audio se puede subdividir en una pluralidad de bloques de muestras. Los bloques pueden tener un tamaño L predeterminado (por ejemplo, 256 muestras por bloque). Además, los bloques adyacentes pueden tener cierto grado de solapamiento (por ejemplo, 50% de solapamiento) de muestras a partir de la trama de audio. El número de bloques por trama de audio puede depender de la característica de la trama de audio (por ejemplo, la presencia de un transitorio). Típicamente, la unidad 302 de Transformación Tiempo Frecuencia aplica una Transformación Tiempo Frecuencia (por ejemplo, una transformada MDCT (Transformada de Coseno Discreta Modificada)) a cada bloque de muestras PCM obtenidas de la trama de audio. Así pues, en la unidad 302 de Transformación Tiempo Frecuencia se obtiene un bloque de coeficientes 312 de transformación para cada bloque de muestras.

Cada canal de la señal de audio multicanal se puede procesar por separado, con el fin de producir secuencias independientes de bloques de coeficientes 312 de transformación para los diferentes canales de la señal de entrada multicanal. En vista de las correlaciones entre algunos canales de la señal de entrada multicanal (por ejemplo, correlaciones entre las señales envolvente Ls y Rs), se puede realizar un canal conjunto en una unidad 303 de procesamiento conjunto de canales. En un modo de realización de ejemplo, la unidad 303 de procesamiento conjunto de canales realiza un acoplamiento de canal, con el fin de convertir un grupo de canales acoplados en un único canal compuesto más información lateral de acoplamiento la cual puede ser utilizada por un sistema decodificador 200, 210 correspondiente para reconstruir los canales individuales a partir del único canal compuesto. A modo de ejemplo, se pueden acoplar los canales Ls y Rs de una señal de audio 5.1 o se pueden acoplar los canales L, C, R, Ls y Rs. Si se utiliza el acoplamiento en la unidad 303, únicamente se envía el canal compuesto único a las unidades de procesamiento adicionales que se muestran en la Fig. 3. En caso contrario, los canales individuales (esto es, las secuencias individuales de bloques de coeficientes 312 de transformación) se pasan a las unidades de procesamiento adicionales del codificador 300.

40 A continuación, se describen las unidades de procesamiento adicionales para una secuencia de ejemplo de bloques de coeficientes 312 de transformación. La descripción es aplicable a cada uno de los canales que se van a codificar (por ejemplo, a los canales individuales de la señal de entrada multicanal o a uno o más canales compuestos resultantes del acoplamiento de canales).

45 La unidad 304 de codificación de coma flotante de bloques está configurada para renderizar los coeficientes 312 de transformación de un canal (aplicable a todos los canales, incluyendo los canales de ancho de banda completa (por ejemplo, los canales L, C y R), el canal LFE (Efectos de Baja Frecuencia), y el canal de acoplamiento) en un formato de exponente/mantisa. Convirtiendo los coeficientes 312 de transformación a un formato exponente/mantisa, se puede independizar el ruido de cuantificación que resulta de la cuantificación de los coeficientes 312 de transformación del nivel absoluto de la señal de entrada.

50 Típicamente, la codificación de coma flotante de bloques en la unidad 304 puede convertir cada uno de los coeficientes 312 de transformación en un exponente y una mantisa. Los exponentes se van a codificar tan eficientemente como sea posible con el fin de reducir la sobrecarga de tasa de datos requerido para transmitir los exponentes 313 codificados. Al mismo tiempo, los exponentes se deberían codificar de forma tan precisa como sea posible con el fin de evitar la pérdida de resolución espectral de los coeficientes 312 de transformación. A continuación, se describe brevemente un esquema de codificación de coma flotante de bloques de ejemplo que se utiliza en DD+ para conseguir los objetivos mencionados anteriormente. Para obtener detalles adicionales en relación con el esquema de codificación DD+ (y en particular el esquema de codificación de coma flotante de bloques utilizado por DD+) se puede consultar el documento Fielder, L.D. y otros "Introduction to Dolby Digital Plus, and Enhancement to the Dolby Digital Coding System (Introducción a Dolby Digital Plus y Mejora del Sistema de Codificación Dolby Digital)", Convención AEC, 28-31 de octubre de 2004.

En un primer paso de codificación de coma flotante de bloques, los exponentes originales se pueden determinar para un bloque de coeficientes 312 de transformación. Esto se ilustra en la Fig. 4a, donde se ilustra un bloque de exponentes originales 401 para un bloque de ejemplo de coeficientes 402 de transformación. Se supone que un coeficiente 402 de transformación tiene un valor  $X$ , en donde el coeficiente 402 de transformación se puede normalizar de modo que  $X$  sea menor o igual que 1. El valor  $X$  se puede representar en un formato mantisa/exponente  $X=m*2(-e)$ , siendo  $m$  la mantisa ( $m \leq 1$ ) y siendo  $e$  el exponente. En un modo de realización, el exponente originales 401 puede tomar valores ente 0 y 24, con el fin de cubrir un rango dinámico de más de 144dB (esto es,  $2(-0)$  a  $2(-24)$ ).

Con el fin de reducir aún más el número de bits necesarios para codificar los exponentes (originales) 401, se pueden aplicar varios esquemas como, por ejemplo, la compartición de tiempo de exponentes a través de bloques de coeficientes 312 de transformación de una trama de audio completa (típicamente seis bloques por trama de audio). Además, se pueden compartir exponentes a través de frecuencias (esto es, a través de contenedores de frecuencia adyacentes en el dominio transformado/de frecuencia). A modo de ejemplo, se puede compartir un exponente a través de dos o más contenedores de frecuencia. Además, los exponentes de un bloque de coeficientes 312 de transformación se pueden acotar con el fin de asegurar que las diferencias entre exponentes adyacentes no exceden un valor máximo predeterminado, por ejemplo,  $\pm 2$ . Esto permite una codificación diferencial eficiente de los exponentes de un bloque de coeficientes 312 de transformación (por ejemplo, utilizando cinco diferenciales). Los esquemas mencionados anteriormente para reducir la tasa de datos requerida para codificar los exponentes (esto es, codificación de compartición de tiempo, de compartición de frecuencia, de acotación y diferencial) se pueden combinar de diferentes maneras para definir modos de codificación de exponentes para obtener diferentes tasas de datos utilizadas para codificar los exponentes. Como resultado de la codificación de exponentes mencionada anteriormente, se obtiene una secuencia de exponentes 313 codificados para los bloques de coeficientes 312 de transformación de una trama de audio (por ejemplo, seis bloques por trama de audio).

Como paso adicional del esquema de Codificación de Coma Flotante de Bloques realizado en la unidad 304, las mantisas  $m'$  de los coeficientes 402 de transformación originales se normalizan mediante el correspondiente exponente  $e'$  codificado resultante. El exponente  $e'$  codificado resultante puede ser diferente del exponente  $e$  original mencionado más arriba (debido a la compartición de tiempo, compartición de frecuencia y/o pasos de acotación). Para cada uno de los coeficientes 402 de transformación de la Fig. 4a, la mantisa  $m'$  normalizada se puede determinar como  $X=m'*2(-e')$ , en donde  $X$  es el valor del coeficiente 402 de transformación original. Las mantisas  $m'$  normalizadas 314 para los bloques de la trama de audio se pasan a la unidad 306 de cuantificación para la cuantificación de las mantisas 314. La cuantificación de las mantisas 314, esto es, la precisión de las mantisas 317 cuantificadas, depende de la tasa de datos que se encuentra disponible para la cuantificación de las mantisas. La tasa de datos disponible se determina en la unidad 305 de asignación de bits.

El proceso de asignación de bits realizado en la unidad 305 determina el número de bits que pueden ser asignados a cada una de las mantisas 314 normalizadas de acuerdo con principios psicoacústicos. El proceso de asignación de bits comprende el paso de determinar el número de bits disponibles para cuantificar las mantisas normalizadas de una trama de audio. Además, el proceso de asignación de bits determina una distribución de densidad espectral de potencia (PSD) y una curva de enmascaramiento en el dominio de la frecuencia (basada en un modelo psicoacústico) para cada canal. La distribución de la PSD y la curva de enmascaramiento en el dominio de la frecuencia se utilizan para determinar una distribución sustancialmente óptima de los bits disponibles para las diferentes mantisas 314 normalizadas de la trama de audio.

El primer paso del proceso de asignación de bits es determinar cuántos bits de mantisa hay disponibles para codificar las mantisas 314 normalizadas. La tasa de datos objetivo se convierte en un número total de bits que están disponibles para codificar una trama de audio actual. En particular, la tasa de datos objetivo especifica un número  $k$  bits/s para la señal de audio multicanal codificada. Considerando una longitud de trama de  $T$  segundos, el número total de bits se puede determinar como  $T*k$ . El número disponible de bits de mantisa se puede determinar a partir del número total de bits restando los bits que ya se han utilizado para codificar la trama de audio como, por ejemplo, metadatos, indicadores de cambio de bloque (para indicar transitorios detectados y tamaños de bloque seleccionados), factores de escala de acoplamiento, exponentes, etc. El proceso de asignación de bits también puede sustraer bits que todavía puede ser necesario asignar a otros aspectos como, por ejemplo, parámetros 315 de asignación de bits (ver más abajo). Como resultado, se puede determinar el número total de bits de mantisa disponibles. El número total de bits de mantisa disponibles se puede distribuir a continuación entre todos los canales (por ejemplo, los canales principales, el canal LFE y el canal de acoplamiento) sobre todos (por ejemplo, uno, dos, tres o seis) bloques de la trama de audio.

Como paso adicional, se puede determinar la distribución de densidad espectral de potencia ("PSD") del bloque de coeficientes 312 de transformación. La PSD es una medida de la energía de la señal en cada contenedor de frecuencia del coeficiente de transformación de la señal de entrada. La PSD se puede determinar basándose en los exponentes 313 codificados, con el fin de permitir que el sistema 200, 210 decodificador de audio multicanal correspondiente determine la PSD del mismo modo que el codificador multicanal 300 de audio. La Fig. 4b ilustra

la distribución 410 de la PSD de un bloque de coeficientes 312 de transformación que se ha obtenido a partir de los exponentes 313 codificados. La distribución 410 de la PSD se puede utilizar para calcular la curva 431 de enmascaramiento en el dominio de la frecuencia (ver Fig. 4d) para el bloque de coeficientes 312 de transformación. La curva 431 de enmascaramiento en el dominio de la frecuencia tiene en cuenta los efectos de enmascaramiento psicoacústico que describen el fenómeno de que una frecuencia enmascaradora enmascara frecuencias en la vecindad directa de la frecuencia enmascaradora, convirtiendo de este modo en inaudibles las frecuencias en la vecindad directa de la frecuencia enmascaradora si su energía se encuentra por debajo de cierto umbral de enmascaramiento. La Fig. 4c muestra una frecuencia enmascaradora 421 y la curva 422 de umbral de enmascaramiento para frecuencias vecinas. La curva 422 de umbral de enmascaramiento real se puede modelar mediante una plantilla 423 de enmascaramiento (lineal definida por tramos) (de dos segmentos) utilizada en el codificador DD+.

Se ha observado que la forma de la curva 422 de umbral de enmascaramiento (y en consecuencia también de la plantilla 423 de enmascaramiento) se mantiene sustancialmente sin cambios para diferentes frecuencias de enmascaramiento sobre una escala de banda crítica tal como ha sido definida por, por ejemplo, Zwicker (o sobre una escala logarítmica). Basándose en esta observación, el codificador DD+ aplica la plantilla 423 de enmascaramiento sobre una distribución de la PSD en bandas (en donde la distribución de la PSD en bandas se corresponde con la distribución de la PSD sobre la escala de banda crítica en la que las bandas tienen aproximadamente una anchura mitad de las bandas críticas). En el caso de una distribución de la PSD en bandas se determina un valor de PSD único para cada una de la pluralidad de bandas en la escala de bandas crítica (o en la escala logarítmica). La Fig. 4d ilustra un ejemplo de distribución 430 de la PSD en bandas para la distribución 410 de la PSD espaciada linealmente de la Fig. 4b. La distribución 430 de la PSD en bandas se puede determinar a partir de la distribución 410 de la PSD espaciada linealmente combinando (por ejemplo, utilizando una operación suma de logaritmos) valores de PSD de la distribución 410 de la PSD espaciada linealmente que se encuentran dentro de la misma banda sobre la escala de bandas crítica (o en la escala logarítmica). La plantilla 423 de enmascaramiento se puede aplicar a cada valor de la PSD de la distribución 430 de la PSD en bandas, con el fin de producir una curva 431 de enmascaramiento en el dominio de la frecuencia global para el bloque de coeficientes 402 de transformación sobre la escala de bandas crítica (o sobre la escala logarítmica) (ver Fig. 4d).

La curva 431 de enmascaramiento en el dominio de la frecuencia global de la Fig. 4d se puede retrotraer a la resolución de la frecuencia lineal y se puede comparar con la distribución 410 de la PSD lineal de un bloque de coeficientes 402 de transformación que se muestran en la Fig. 4b. Esto se ilustra en la Fig. 4e que muestra la curva 441 de enmascaramiento en el dominio de la frecuencia sobre una resolución lineal, así como la distribución 410 de la PSD sobre una resolución lineal. Se debería observar que la curva 441 de enmascaramiento en el dominio de la frecuencia también puede tener en cuenta el umbral absoluto de la curva de escucha. El número de bits para codificar la mantisa de los coeficientes 402 de transformación de un contenedor de frecuencia concreto se puede determinar en función de la distribución 410 de la PSD y en función de la curva 441 de enmascaramiento. En particular, los valores de PSD de la distribución 410 de la PSD que se encuentran por debajo de la curva 441 de enmascaramiento se corresponden con las mantisas que son perceptualmente irrelevantes (porque el componente de frecuencia de la señal de audio en dichos contenedores de frecuencia es enmascarada por una frecuencia de enmascaramiento en su vecindad). En consecuencia, no es necesario en absoluto asignar ningún bit a las mantisas de dichos coeficientes 402 de transformación. Por otro lado, los valores de PSD de la distribución 410 de la PSD que se encuentran por encima de la curva 441 de enmascaramiento indican que se deberían asignar bits para codificar las mantisas de los coeficientes 402 de transformación en dichos contenedores de frecuencia. El número de bits asignados a dichas mantisas debería aumentar con el aumento de la diferencia entre el valor de PSD de la distribución 410 de la PSD y el valor de la curva 441 de enmascaramiento. El proceso de asignación de bits mencionado más arriba produce una asignación 442 de bits a los diferentes coeficientes 402 de transformación tal como se muestra en la Fig. 4e.

El proceso de asignación de bits mencionado más arriba se ejecuta para todos los canales (por ejemplo, los canales directos, el canal LFE y el canal de acoplamiento) y para todos los bloques de la trama de audio, con el fin de generar un número (preliminar) total de bits asignados. Es poco probable que este número preliminar total de bits asignados se corresponda con (por ejemplo, se igual a) el número total de bits de mantisa disponibles. En algunos casos (por ejemplo, para señales de audio complejas), el número preliminar total de bits asignados puede exceder el número de bits de mantisa disponibles (falta de bits). En otros casos (por ejemplo, en el caso de señales de audio simples), el número preliminar total de bits asignados puede encontrarse por debajo del número de bits de mantisa disponibles (exceso de bits). El codificador 300 intenta típicamente ajustar el número total (final) de bits asignados tanto como sea posible al número de bits de mantisa disponibles. Para este propósito, el codificador 300 puede utilizar un parámetro denominado variación de SNR. La variación de SNR permite un ajuste de la curva 441 de enmascaramiento, moviendo la curva 441 de enmascaramiento hacia arriba o hacia abajo con respecto a la distribución 410 de la PSD. Desplazando hacia arriba o hacia abajo la curva 441 de enmascaramiento, el número (preliminar) de bits asignados puede disminuir o aumentar, respectivamente. Así pues, la variación de SNR se puede ajustar de forma iterativa hasta que se cumpla un criterio de terminación (por

ejemplo, el criterio de que el número preliminar de bits asignados se encuentre tan cerca como sea posible (pero por debajo) del número de bits asignados; o el criterio de que se ha realizado un número máximo de iteraciones).

Tal como se ha indicado más arriba, la búsqueda iterativa para una variación de SNR que permite un mejor ajuste entre el número final de bits asignados y el número de bits disponibles hace uso de una búsqueda binaria. En cada iteración, se determina si el número preliminar de bits asignados excede o no el número de bits disponibles. A partir de este paso de determinación se modifica la variación de SNR y se realiza el paso siguiente de la iteración. La búsqueda binaria se configura para determinar el mejor ajuste (y la variación de SNR correspondiente) utilizando iteraciones  $(\log_2(K)+1)$ , en donde K es el número de variaciones de SNR posibles. Después de terminar la búsqueda iterativa se obtiene un número de bits asignados final (el cual se corresponde típicamente con uno de los números de bits asignados preliminares determinados previamente). Se debería observar que el número final de bits asignados puede ser (ligeramente) inferior al número de bits disponibles. En dichos casos, se pueden utilizar bits de salto para alinear completamente el número final de bits asignados con el número de bits disponibles.

La variación de SNR se puede definir de modo que una variación de SNR cero produce mantisas codificadas que dan lugar a una condición de codificación conocida como "diferencia casi imperceptible" entre la señal de audio original y la señal codificada. En otras palabras, en una variación de SNR cero el codificador 300 opera de acuerdo con el modelo de percepción. Un valor positivo de la variación de SNR puede desplazar hacia abajo la curva 441 de enmascaramiento, aumentando de este modo el número de bits asignados (típicamente sin ninguna mejora de calidad perceptible). Un valor negativo de la variación de SNR puede desplazar hacia arriba la curva 441 de enmascaramiento, disminuyendo de este modo el número de bits asignados (e incrementando así típicamente el ruido de cuantificación audible). La variación de SNR puede ser, por ejemplo, ser un parámetro de 10 bits con un rango válido de -48 a +144 dB. Con el fin de encontrar el valor de variación de SNR óptimo, el codificador 300 puede realizar una búsqueda binaria iterativa. La búsqueda binaria iterativa puede entonces requerir hasta 11 iteraciones (en el caso de un parámetro de 10 bits) de comparaciones de distribución 410 de la PSD/curva 441 de enmascaramiento. El valor de variación de SNR utilizado realmente se puede transmitir al decodificador correspondiente como un parámetro 315 de asignación de bits. Además, las mantisas se codifican de acuerdo con los bits asignados finales, con el fin de producir un conjunto de mantisas codificadas 317.

Así pues, el parámetro de variación de SNR (Relación Señal a Ruido) se puede utilizar como indicador de la calidad de codificación de la señal de audio multicanal codificada. De acuerdo con el convenio mencionado más arriba de la variación de SNR, una variación de SNR cero indica una señal de audio multicanal codificada con una "diferencia casi imperceptible" con respecto a la señal de audio multicanal original. Una variación de SNR positiva indica una señal de audio multicanal codificada que tiene una calidad de al menos la "diferencia casi imperceptible" con respecto a la señal de audio multicanal original. Una variación de SNR negativa indica una señal de audio multicanal codificada que tiene una calidad menor que la "diferencia casi imperceptible" con respecto a la señal de audio multicanal original. Se debería observar que pueden ser posibles otros convenios del parámetro de variación de SNR (por ejemplo, una definición inversa).

El codificador 300 comprende, además, una unidad 307 de empaquetado de flujos de bits, la cual está configurada para organizar los exponentes 313 codificados, las mantisas 317 codificadas, los parámetros 315 de asignación de bits, así como otros datos de codificación (por ejemplo, indicadores de cambio de bloque, metadatos, factores de escala de acoplamiento, etc.) en una estructura de trama predeterminada (por ejemplo, la estructura de trama AC-3), con el fin de generar una trama codificada 318 para una trama de audio de la señal de audio multicanal.

Tal como ya se ha resumido más arriba, y tal como se muestra en la Fig. 1a, los flujos DD+ 7.1 se codifican típicamente mediante una codificación independiente del grupo básico 121 de canales utilizando un codificador IS 105, con el fin de generar el IS 110 y un grupo 122 de ampliación de canales utilizando un codificador DS 106, con el fin de generar el DS 120. El codificador IS 105 y el codificador DS 106 se configuran típicamente con un porcentaje fijo de la tasa de datos total, esto es, cada codificador 105, 106 realiza un proceso de asignación de bits independiente sin ninguna interacción entre ambos codificadores 105, 106. Típicamente, al codificador IS 105 se le asigna el X% de la tasa de datos total y al codificador DS 106 se le asigna el 100-X% de la tasa de datos total, en donde X es un valor fijo, por ejemplo, X=50.

Tal como se ha descrito más arriba, el codificador multicanal 300 ajusta la variación de SNR de modo que el número total (final) de bits asignados se aproxima (tanto como sea posible) al número total de bits disponibles. En el contexto de este proceso de asignación de bits, la variación de SNR se puede ajustar (por ejemplo, aumentar/reducir) de modo que el número de bits asignados se aumenta/reduce. Sin embargo, si el codificador 300 asigna más bits de los necesarios con el fin de conseguir la "diferencia casi imperceptible", los bits asignados adicionalmente realmente se desperdician, debido a que los bits asignados adicionalmente típicamente no producen ninguna mejora de la calidad perceptual de la señal de audio codificada. Teniendo esto en cuenta, se propone proporcionar un proceso de asignación de bits flexible y combinado para el codificador IS 105 y para el codificador DS 106, con el fin de permitir que los dos codificadores 105, 106 ajusten de forma dinámica la

fracción de la tasa de datos total para el codificador IS 105 (denominada "tasa de datos IS") y la fracción de la tasa de datos total para el codificador DS 106 (denominada "tasa de datos DS") a lo largo del tiempo (en función de los requisitos de la señal de audio multicanal). La tasa de datos IS y la tasa de datos DS se ajustan preferiblemente de modo que su suma se corresponde siempre con la tasa de datos total. El proceso de asignación de bits combinado se ilustra en la Fig. 5a. La Fig. 5a muestra el codificador IS 105 y el codificador DS 106. Además, la Fig. 5a muestra una unidad 501 de control de tasa configurada para determinar la tasa de datos IS y la tasa de datos DS en función de datos 505 de salida realimentados desde el codificador IS 105 y en función de datos 506 de salida realimentados desde el codificador DS 106. Los datos 505, 506 de salida pueden ser, por ejemplo, ser el IS codificado 110 y el DS codificado 120, respectivamente; y/o la variación de SNR del codificador 105, 106 respectivo. Así pues, la unidad 501 de control de tasa puede tener en cuenta los datos 505, 506 de salida de los dos codificadores 105, 106 para determinar de forma dinámica la tasa de datos IS y la tasa de datos DS. En un modo de realización preferido, la asignación variable de la tasa de datos IS y la tasa de datos DS se realiza de modo que la asignación variable no tiene impacto en el sistema 200, 210 decodificador de audio multicanal correspondiente. En otras palabras, la asignación variable debería ser transparente al sistema 200, 210 decodificador de audio multicanal correspondiente.

Una posible forma de implementar una asignación variable de las tasas de datos IS/DS es implementar un proceso de asignación de bits compartido para asignar los bits de mantisa. El codificador IS 105 y el codificador DS 106 pueden realizar de forma independiente pasos de codificación que preceden al proceso de asignación de bits de mantisa (realizado en la unidad 305 de asignación de bits). En particular, la codificación de los indicadores de cambio de bloque, factores de escala de acoplamiento, exponentes, extensión espectral, etc. se pueden realizar de forma independiente en el codificador IS 105 y en el codificador DS 106. Por otro lado, el proceso de asignación de bits realizado en las respectivas unidades 305 del codificador IS 105 y el codificador DS 106 se pueden realizar conjuntamente. Típicamente, alrededor del 80% de los bits del IS y el DS se utilizan para codificar las mantisas. En consecuencia, incluso aunque los codificadores IS y DS 105, 106 trabajen de forma independiente para una codificación distinta de la asignación de bits de mantisa, la parte significativa de la codificación (esto es, la asignación de bits de mantisa) se realiza conjuntamente.

En otras palabras, se propone codificar independientemente los datos "fijos" de cada grupo de canales (por ejemplo, los exponentes, coordenadas de acoplamiento, extensión espectral, etc.). Posteriormente, se realiza un proceso de asignación de bits único para el grupo básico 121 y el grupo 122 de ampliación utilizando el total de bits restantes. A continuación, se cuantifican y empaquetan las mantisas de ambos flujos para generar las tramas 151 del IS codificadas (denominadas tramas IS 151) y las tramas 152 del DS codificadas (denominadas tramas DS 152). Como resultado del proceso de asignación de bits combinado, las tramas IS 151 pueden variar de tamaño a lo largo del tiempo (debido a una variación de la tasa de datos del IS). De forma parecida, las tramas DS 152 pueden variar de tamaño a lo largo del tiempo (debido a una variación de la tasa de datos del IS). Sin embargo, para cada intervalo de tiempo 170 (esto es, para cada trama de audio de la señal de audio multicanal) la suma del tamaño de la(s) trama(s) IS 151 y la(s) trama(s) DS 152 debería ser sustancialmente constante (debido a la tasa de datos total constante). Además, como resultado del proceso de asignación de bits combinado, la variación de SNR del IS y el DS deberían ser idénticas, debido a que el proceso de asignación de bits conjunto realizado en la unidad 305 de asignación de bits conjunta ajusta una variación de SNR conjunta con el fin de hacer coincidir el número de bits de mantisa asignados (conjunto para el IS y el DS) y el número de bits de mantisa disponibles (conjunto para el IS y el DS). El hecho de tener variaciones de SNR idénticas para el IS y el DS debería mejorar la calidad total permitiendo que el flujo secundario más falto de bits (por ejemplo, el IS) utilice los bits extra si y cuando el otro flujo secundario (por ejemplo, el DS) tenga en exceso.

La Fig. 5b ilustra el diagrama de flujo de un método 510 de codificación IS/DS combinado de ejemplo. El método comprende los pasos 521, 531 independientes de acondicionamiento de la señal para las tramas de señal del grupo básico 121 y el grupo 122 de ampliación, respectivamente. El método 510 continúa con los pasos 522, 532 independientes de Transformación de Tiempo a Frecuencia para los bloques del grupo básico 121 y para los bloques del grupo 122 de ampliación, respectivamente. A continuación, se pueden realizar los pasos 523, 533 conjuntos de procesamiento de canales para el grupo básico 121 y el grupo 122 de ampliación, respectivamente. A modo de ejemplo, en el caso del grupo básico 121, se pueden acoplar los canales Lst y Rst o todos los canales (excepto el canal LFE) (paso 523), en donde para el grupo 122 de ampliación, se pueden acoplar los canales Ls y Rs y/o Lb y Rb (paso 533), con el fin de generar los canales acoplados y los parámetros de acoplamiento respectivos. Además, se puede realizar la Codificación 524, 534 de Bloques en Coma Flotante para los bloques del grupo básico 121 y para los bloques del grupo 122 de ampliación, respectivamente. Como resultado, se obtienen los exponentes 313 codificados para el grupo básico 121 y para el grupo 122 de ampliación, respectivamente. Los pasos de procesamiento mencionados más arriba se pueden realizar tal como se ha resumido en el contexto de la Fig. 3.

El método 510 comprende un paso 540 conjunto de asignación de bits. La asignación conjunta 540 de bits comprende un paso 541 conjunto para determinar los bits de mantisa disponibles, esto es, para determinar el número total de bits que se encuentran disponibles para codificar las mantisas del grupo básico 121 y del grupo 122 de ampliación. Además, el método 510 comprende los pasos 525, 535 de determinación de la distribución de

la PSD para los bloques del grupo básico 121 y para los bloques del grupo 122 de ampliación, respectivamente. Además, el método 510 comprende los pasos 526, 536 de determinación de la curva de enmascaramiento para el grupo básico 121 y el grupo 122 de ampliación, respectivamente. Tal como se ha resumido más arriba, las distribuciones de la PSD y las curvas de enmascaramiento se determinan para cada uno de los canales de la señal multicanal y para cada uno de los bloques de la trama de señal. En el contexto de los pasos 527, 537 de comparación de PSD/enmascaramiento (para el grupo básico 121 y el grupo 122 de ampliación, respectivamente) se comparan las distribuciones de la PSD y las curvas de enmascaramiento y se asignan los bits a las mantisas del grupo básico 121 y el grupo 122 de ampliación, respectivamente. Estos pasos se ejecutan para cada uno de los canales y para cada uno de los bloques. Además, estos pasos se ejecutan para una variación de SNR dada (que es la misma para los pasos 527 y 537 de comparación de PSD/enmascaramiento).

Después de la asignación de bits a las mantisas utilizando una variación de SNR dada, el método 510 continúa con el paso 542 conjunto de ajuste de determinación del número total de bits de mantisa disponibles. Además, en el contexto del paso 542 se determina si el número total de bits de mantisa asignados coincide con el número total de bits de mantisa disponibles (determinado en el paso 541). Si se ha determinado una correspondencia óptima, el método 510 continúa con la cuantificación 528, 538 de las mantisas del grupo básico 121 y el grupo 122 de ampliación, respectivamente, en función de los bits de mantisa disponibles determinados en los pasos 527, 537. Además, en los pasos 529, 539 de empaquetamiento del flujo de bits se determinan las tramas IS 151 y las tramas DS 152, respectivamente. Por otro lado, si todavía no se ha conseguido una correspondencia óptima, se modifica la variación de SNR y se repiten los pasos 527, 537 de comparación de PSD/enmascaramiento y el paso 542 de ajuste. Los pasos 527, 537 y 542 se iteran hasta que se determina una correspondencia óptima y/o hasta que se alcanza una condición de terminación (por ejemplo, un número máximo de iteraciones).

Se debería observar que los pasos 525, 535 de determinación de la PSD, los pasos 526, 536 de determinación de la curva de enmascaramiento y los pasos 527, 537 de comparación de PSD/enmascaramiento se realizan para cada canal de la señal multicanal y para cada bloque de una trama de señal. En consecuencia, estos pasos se realizan (por definición) por separado para el grupo básico 121 y el grupo 122 de ampliación. De hecho, estos pasos se realizan por separado para cada canal de la señal multicanal.

En conjunto, el método 510 de codificación permite una asignación mejorada de las tasas de datos al IS y al DS (en comparación con un proceso de asignación de bits independiente). Como consecuencia, se mejora la calidad perceptual de la señal multicanal codificada (que comprende un IS y al menos un DS) (en comparación con una señal multicanal codificada utilizando codificadores IS y DS 105, 106 independientes).

Se debería observar que las tramas IS 151 y las tramas DS 152 que genera el método 510 se pueden disponer de forma que sean compatibles con las tramas IS y las tramas DS generadas por los codificadores IS y DS 105, 106 independientes, respectivamente. En particular, cada una de las tramas IS y DS 151, 152 puede comprender parámetros de asignación de bits que permitan que un sistema 200, 210 decodificador multicanal convencional decodifique por separado las tramas IS y DS 151, 152. En particular, en la trama IS 151 y en la trama DS 152 se puede insertar el (mismo) valor de variación de SNR. Por lo tanto, se puede utilizar un codificador multicanal basado en el método 510 junto con los sistemas 200, 210 decodificadores multicanal convencionales.

Puede ser deseable utilizar un codificador IS 105 estándar y un codificador DS 106 estándar para codificar el grupo básico 121 y el grupo 122 de ampliación, respectivamente. Esto puede ser beneficioso por razones de coste. Además, en algunas situaciones puede no ser posible implementar un proceso 540 de asignación conjunta de conjunto tal como se ha descrito en el contexto de la Fig. 5b. Sin embargo, es deseable permitir la adaptación de la tasa de datos IS y la tasa de datos DS a la señal de audio multicanal y mejorar de este modo la calidad global de la señal de audio multicanal codificada.

Con el fin de permitir la adaptación de la tasa de datos IS y la tasa de datos DS sin modificar el codificador IS 105 y el codificador DS 106, la tasa de datos IS y la tasa de datos DS se pueden controlar externamente a los codificadores IS/DS 105, 106, por ejemplo, en función de la dificultad de codificación de flujo relativa estimada para una trama concreta. La dificultad de codificación relativa para una trama concreta se puede estimar, por ejemplo, en función de la entropía perceptual, en función de la tonalidad o en función de la energía. La dificultad de codificación se puede calcular en función de las muestras PCM de entrada del codificador relevantes para la trama actual a codificar. Esto puede requerir una alineación de tiempo correcta de las muestras PCM en función de cualquier retardo de tiempo de codificación posterior (por ejemplo, provocado por un filtro LFE, un filtro HP, un desplazamiento de fase de 90° de los canales Envolventes Izquierdo y Derecho y/o Pre-Procesamiento de Ruido Temporal (TPNP)). Ejemplos de indicadores de la dificultad de codificación pueden ser la potencia de la señal, la uniformidad espectral, las estimaciones de tonalidad, las estimaciones transitorias y/o la entropía perceptual. La entropía perceptual mide el número de bits necesarios para codificar un espectro de señal con un ruido de cuantificación justo por debajo del umbral de enmascaramiento. Un valor mayor para la entropía perceptual indica una dificultad de codificación mayor. Sonidos con un carácter tonal (esto es, sonidos con una estimación de tonalidad alta) son típicamente más difíciles de codificar tal como se refleja, por ejemplo, en el cálculo de la curva de enmascaramiento del Modelo Psicoacústico MPEG-1 ISO/IEC 11172-3. Así pues, una estimación de tonalidad

alta puede indicar una dificultad de codificación alta (y viceversa). Un indicador simple para la dificultad de codificación se puede basar en la potencia promedio de la señal del grupo básico de canales y/o de los grupos de ampliación de canales.

5 Se puede comparar la dificultad de codificación estimada de una trama actual del grupo básico y la trama actual correspondiente del grupo de ampliación y se puede distribuir consecuentemente la tasa de datos IS / tasa de datos DS (y los bits de mantisa respectivos). Una posible fórmula para determinar la tasa de datos IS / tasa de datos DS puede ser:

$$R_{IS} = R_T \left( \frac{(D_{IS} N_{IS})}{(D_{IS} N_{IS} + D_{DS} N_{DS})} \right) \quad \text{y} \quad R_{DS} = R_T \left( \frac{(D_{DS} N_{DS})}{(D_{IS} N_{IS} + D_{DS} N_{DS})} \right)$$

10 en donde  $R_{DS}$  es la tasa de datos DS,  $R_T$  es la tasa de datos total,  $R_{IS}$  es la tasa de datos IS,  $D_{IS}$  es la dificultad de codificación de un canal del grupo básico (por ejemplo, la dificultad promedio de codificación de los canales del grupo básico,  $D_{DS}$  es la dificultad de codificación de un canal del grupo de ampliación (por ejemplo, la dificultad promedio de codificación de los canales del grupo de ampliación),  $N_{IS}$  es el número de canales en el grupo básico, y  $N_{DS}$  es el número de canales en el grupo de ampliación.

15 Las tasas de datos DS e IS determinadas se pueden calcular de modo que el número de bits para el IS y/o el DS no se encuentre por debajo de un mínimo número de bits fijo para una trama IS y/o una trama DS. Así pues, se puede asegurar una calidad mínima para el IS y/o el DS. En concreto, el mínimo número de bits fijo para una trama IS y/o para una trama DS puede estar limitado por el número de bits necesarios para codificar todos los datos distintos de las mantisas (por ejemplo, los exponentes, etc.).

20 En otra estrategia, se puede determinar la diferencia (IS vs. DS) de dificultad de codificación mediana (o media) sobre un conjunto grande de contenido multicanal relevante. El control de la distribución de tasa de datos puede ser de modo que para las tramas típicas (con una diferencia de dificultad de codificación dentro de un rango predeterminado de la diferencia de dificultad de codificación mediana) se utilice una distribución de tasa de datos por defecto (por ejemplo, X% y 100%-X%). En caso contrario, la distribución de tasa de datos puede desviarse del valor por defecto en función de la desviación de la diferencia de dificultad de codificación real con respecto a la diferencia de dificultad de codificación mediana.

25 En la Fig. 5c se ilustra un codificador 550 que adapta la tasa de datos IS y la tasa de datos DS en función de la dificultad de codificación. El codificador 550 comprende una unidad 551 de determinación de dificultad de codificación que recibe la señal 552 de audio multicanal (y/o el grupo básico 121 de canales y el grupo 122 de ampliación de canales). La unidad 551 de determinación de dificultad de codificación analiza las tramas de señal respectivas del grupo básico 121 y el grupo 122 de ampliación y determina una dificultad de codificación relativa de las tramas del grupo básico 121 y el grupo 122 de ampliación. La dificultad de codificación relativa se pasa a la unidad 553 de control de tasa, la cual está configurada para determinar la tasa de datos IS 561 y la tasa de datos DS 562 en función de la dificultad de codificación relativa. A modo de ejemplo, si la dificultad de codificación relativa indica una dificultad de codificación mayor para el grupo básico 121 con respecto al grupo 122 de ampliación, se aumenta la tasa de datos IS 561 y se reduce la tasa de datos DS 562 (y viceversa).

30 Otra estrategia para una adaptación de la tasa de datos IS y la tasa de datos DS sin modificar el codificador IS 105 y el codificador DS 106 es extraer uno o más parámetros de codificación de las tramas IS/DS 151, 152 y utilizar el uno o más parámetros de codificación para modificar la tasa de datos IS y la tasa de datos DS. A modo de ejemplo, el uno o más parámetros de codificación extraídos de las tramas IS/DS 151, 152 de una trama de señal (n-1) se puede tener en cuenta para determinar las tasas de datos IS/DS para codificar la trama de señal (n) siguiente. El uno o más parámetros de codificación pueden estar relacionados con la calidad de percepción del IS 110 codificado y el DS 120 codificado. A modo de ejemplo, el uno o más parámetros de codificación puede ser la variación de SNR DD/DD+ utilizada en el codificador IS 105 (denominada variación de SNR IS) y la variación de SNR utilizada en el codificador DS 106 (denominada variación de SNR DS). Así pues, se pueden utilizar las variaciones de SNR IS/DS tomadas de las tramas IS/DS 151, 152 anteriores (en el instante de tiempo (n-1)) para controlar de forma adaptativa las tasas de datos IS/DS para la trama de señal siguiente (en el instante de tiempo (n)), de modo que las variaciones de SNR IS/DS se ecualizan a lo largo del flujo de señal de audio multicanal. En términos más generales, se puede establecer que se puede utilizar el uno o más parámetros de codificación tomados de las tramas IS/DS 151, 152 (en el instante de tiempo (n-1)) para controlar de forma adaptativa las tasas de datos IS/DS para la trama de señal siguiente (en el instante de tiempo (n)), de modo que el uno o más parámetros de codificación se ecualizan a lo largo del flujo de señal de audio multicanal. Por lo tanto, el objetivo es proporcionar la misma calidad para diferentes grupos de la señal multicanal codificada. En otras palabras, el objetivo es asegurar que la calidad de los flujos secundarios codificados es tan parecida como sea posible para todos los flujos secundarios de un flujo de señal de audio multicanal. El objetivo se debería



conseguir para cada trama de la señal de audio, esto es, para todos los instantes de tiempo o para todas las tramas de la señal.

La Fig. 6 muestra un diagrama de bloques de un codificador 600 de ejemplo que comprende un esquema de adaptación de tasa de datos IS/DS externo. El codificador 600 comprende un codificador IS 105 y un codificador DS 106 los cuales pueden estar configurados de acuerdo con el codificador 300 ilustrado en la Fig. 3. Para una trama de señal (n-1) y para una tasa de datos IS (n-1) y una tasa de datos DS (n-1) asignadas en un instante temporal o número de trama (n-1), los codificadores IS/DS 105, 106 proporcionan una trama IS(n-1) codificada y una trama DS(n-1) codificada, respectivamente. El codificador IS 105 utiliza la variación de SNR IS(n-1) y el codificador DS 106 utiliza la variación de SNR DS(n-1) para asignar a las mantisas la tasa de datos IS(n-1) y la tasa de datos DS(n-1), respectivamente. La variación de SNR IS(n-1) y la variación de SNR DS(n-1) se pueden extraer de la trama IS(n-1) y la trama DS(n-1), respectivamente. Con el fin de asegurar una alineación entre la variación de SNR IS y la variación de SNR DS a lo largo del flujo (esto es, a lo largo de los números (n) de trama), la variación de SNR IS(n-1) y la variación de SNR DS(n-1) pueden realimentar la entrada de los codificadores IS/DS 105, 106 con el fin de adaptar la tasa de datos IS(n) y la tasa de datos DS(n) para codificar la trama (n) de señal siguiente.

En particular, el codificador 600 comprende una unidad 601 de desviación de variación de SNR configurada para determinar la diferencia entre la variación de SNR IS(n-1) y la variación de SNR DS(n-1). La diferencia se puede utilizar para controlar las tasas de datos IS/DS(n) (para la trama de señal siguiente). En un modo de realización, si una variación de SNR IS(n-1) es menor que una variación de SNR DS(n-1) (esto es, la diferencia es negativa) indica que la calidad perceptual del IS es más bien menor que la calidad perceptual del DS. En consecuencia, la tasa de datos DS(n) debería reducirse con respecto a la tasa de datos DS(n-1), con el fin de reducir la calidad perceptual del IS (o posiblemente dejarla sin afectar) en la trama de señal (n) siguiente. Al mismo tiempo, la tasa de datos IS(n) debería aumentarse con respecto a la tasa de datos IS(n-1), con el fin de aumentar la calidad perceptual del IS en la trama de señal (n) siguiente y también para satisfacer el requisito de tasa de datos total. La modificación de la tasa de datos IS(n) en función de la variación de SNR IS(n-1) se basa en la suposición de que la dificultad de codificación tal como se refleja en el parámetro variación de SNR IS(n-1) no cambia significativamente entre dos tramas consecutivas. De forma parecida, una variación de SNR IS(n-1) que es mayor que la variación de SNR DS(n-1) (esto es, una diferencia positiva) puede indicar que la calidad perceptual del IS es mayor que la calidad perceptual del DS. La tasa de datos IS(n) y la tasa de datos DS(n) se pueden modificar con respecto a la tasa de datos IS(n-1) y la tasa de datos DS(n-1) de modo que se reduce la calidad perceptual del IS (o se deja sin afectar) y la calidad perceptual del DS aumenta.

El mecanismo de control mencionado más arriba se puede implementar de varias formas. El codificador 600 comprende una unidad 602 de determinación de signo que está configurada para determinar el signo de la diferencia entre la variación de SNR IS(n-1) y la variación de SNR DS(n-1). Además, el codificador 600 utiliza una variación 603 de la tasa de datos predeterminada (por ejemplo, un porcentaje de la tasa de datos total disponible, por ejemplo, aproximadamente el 0,5%, 1%, 2%, 3%, 4%, 5% o 10% de la tasa de datos total disponible) que se puede aplicar para modificar la tasa de datos IS(n) y la tasa de datos DS(n) con respecto a la tasa de datos IS(n-1) con regularidad y la tasa de datos DS(n-1) en la unidad 605 de modificación de la tasa de IS y en la unidad 606 de modificación de la tasa de DS. A modo de ejemplo, si la diferencia es negativa, la unidad 605 de modificación de la tasa de IS determina que la tasa de datos IS(n) = la tasa de datos IS(n-1) + la variación de tasa de datos, y la unidad 606 de modificación de la tasa de DS determina que la tasa de datos DS(n) = la tasa de datos DS(n-1) – la variación de tasa de datos (y viceversa en el caso de una diferencia positiva).

El esquema de control externo mencionado más arriba para adaptar la asignación de la tasa de datos total a la tasa de datos IS y la tasa de datos DS se encamina a reducir la diferencia entre la variación de SNR IS y la variación de SNR DS. En otras palabras, el esquema de control mencionado más arriba intenta alinear la variación de SNR IS y la variación de SNR DS, con el fin de alinear la calidad percibida del IS codificado y el DS codificado. Como resultado, se mejora la calidad percibida global de la señal multicanal codificada (que comprende el IS codificado y el DS codificado) (comparada con la del codificador 100 que utiliza tasas de datos IS/DS fijas).

En el presente documento, se han descrito métodos y sistemas para codificar una señal de audio multicanal. Los métodos y sistemas codifican una señal de audio multicanal en una pluralidad de flujos secundarios, en donde la pluralidad de flujos secundarios permite una decodificación eficiente de diferentes combinaciones de canales de la señal de audio multicanal. Además, los métodos y sistemas permiten una asignación conjunta de los bits de mantisa a lo largo de la pluralidad de flujos, aumentando de este modo la calidad percibida de la señal de audio multicanal codificada (y posteriormente decodificada). Los métodos y sistemas se pueden configurar de modo que los flujos codificados sean compatibles con decodificadores de audio multicanal antiguos.

En particular, el presente documento describe la transmisión de canales 7.1 en DD+ dentro de dos flujos secundarios, en donde un primer flujo secundario "independiente" comprende una mezcla de canales 5.1, y un segundo flujo secundario "dependiente" comprende canales de "ampliación" y/o "sustitución". Actualmente, la

5 codificación de flujos 7.1 se realiza típicamente mediante dos codificadores 5.1 principales que no se relacionan entre sí. Los dos codificadores 5.1 principales disponen de una tasa de datos – una parte fija de la tasa de datos total disponible – y realizan de forma independiente la codificación de los dos flujos secundarios. En el presente documento, se ha propuesto compartir los bits de mantisa entre (al menos) dos flujos secundarios. En un modo de realización, los datos 'fijos' de cada flujo se codifican de forma independiente (exponentes, coordenadas de acoplamiento, etc.). A continuación, se realiza un proceso de asignación de bits único para ambos flujos con el resto de bits. Por último, se pueden cuantificar y empaquetar las mantisas de ambos flujos. Haciendo esto, cada intervalo de tiempo de una señal codificada es idéntico en tamaño, pero las tramas individuales codificadas (por ejemplo, la trama IS y/o las tramas DS) pueden variar. Además, la Variación de SNR de los flujos independiente y dependientes puede ser igual (o se puede reducir su diferencia). De este modo, la calidad de codificación global se puede mejorar permitiendo que el flujo secundario más falto de bits utilice bits extra si/cuando el otro flujo secundario tenga en exceso.

15 Se debería observar que mientras que los métodos y sistemas se han descrito en el contexto de un codificador de audio DD+ 7.1, los métodos y sistemas son aplicables a otros codificadores que crean flujos de bits DD+ con múltiples flujos secundarios. Además, los métodos y sistemas son aplicables a otros codec de audio/vídeo que utilicen el concepto de un grupo de bits, múltiples flujos secundarios y que tengan una restricción sobre la tasa de datos total (por ejemplo, que requieran una tasa de datos constante). Los codec de audio/vídeo que operan sobre los flujos secundarios relacionados pueden aplicar un grupo de bits compartido para asignar bits a los flujos secundarios relacionados según sea necesario, y variar las tasas de datos de los flujos secundarios al tiempo que mantienen constante la tasa de datos total.

20 Los métodos y sistemas descritos en el presente documento se pueden implementar como software, firmware y/o hardware. Ciertos componentes se pueden implementar, por ejemplo, como un software ejecutándose sobre un procesador de señales digitales o un microprocesador. Otros componentes se pueden implementar, por ejemplo, como hardware o como circuitos integrados de aplicación específica. Las señales que se encuentran en los métodos y sistemas descritos se pueden almacenar en medios como, por ejemplo, una memoria de acceso aleatorio o medios de almacenamiento óptico. Se pueden transferir a través de redes como, por ejemplo, redes de radio, redes de satélites, redes inalámbricas o redes de cable como, por ejemplo, Internet. Dispositivos típicos que utilizan los métodos y sistemas descritos en el presente documento son dispositivos electrónicos portátiles u otro equipo de consumo que se utilice para almacenar y/o renderizar señales de audio.

25

**REIVINDICACIONES**

- 5 1) Un codificador de audio configurado para codificar una señal de audio multicanal de acuerdo con una tasa de datos total disponible; en donde la señal de audio multicanal se puede representar como un grupo básico (121) de canales para renderizar (convertir) la señal de audio multicanal de acuerdo con una configuración básica de canales, y como un grupo (122) de ampliación de canales que – en combinación con el grupo básico (122) – se utiliza para renderizar la señal de audio multicanal de acuerdo con una configuración ampliada de canales; en donde la configuración básica de canales y la configuración ampliada de canales son diferentes entre sí; comprendiendo el codificador de audio:
- 10 – un codificador básico (105) configurado para codificar el grupo básico (121) de canales de acuerdo con una tasa de datos IS, con el fin de generar un flujo secundario independiente (110), denominado IS; y
- un codificador ampliado (106) configurado para codificar el grupo (122) de ampliación de canales de acuerdo con una tasa de datos DS, con el fin de generar un flujo secundario dependiente (120), denominado DS,
- 15 – caracterizado por que el codificador de audio comprende: una unidad (501) de control de tasa configurada para adaptar con regularidad la tasa de datos IS y la tasa de datos DS en función de un indicador de calidad de codificación IS instantáneo para el grupo básico (121) de canales y/o en función de un indicador de calidad de codificación DS instantáneo para el grupo (122) de ampliación de canales, de modo que la suma de la tasa de datos IS y la tasa de datos DS se corresponde sustancialmente con la tasa de datos total disponible.
- 20 2) El codificador de la reivindicación 1, en donde la unidad (501) de control de tasa está configurada para determinar la tasa de datos IS y la tasa de datos DS con el fin de reducir la diferencia entre el indicador de calidad de codificación IS instantáneo y el indicador de calidad de codificación DS instantáneo.
- 25 3) El codificador de cualquiera de las reivindicaciones anteriores, en donde el codificador básico (105) y el codificador (106) de ampliación son codificadores de audio basados en tramas configurados para codificar una secuencia de tramas de la señal de audio multicanal, con el fin de generar las secuencias correspondientes de tramas IS (151) y tramas DS (152) del flujo secundario independiente (110) y el flujo secundario dependiente (120), respectivamente.
- 30 4) El codificador de la reivindicación 3, en donde la unidad (501) de control de tasa está configurada para adaptar la tasa de datos IS y la tasa de datos DS para cada trama de la secuencia de tramas de la señal de audio multicanal.
- 5) El codificador de cualquiera de las reivindicaciones 3 a 4, en donde
- el indicador de calidad de codificación IS comprende una secuencia de indicadores de calidad de codificación IS para la secuencia de tramas IS (151) correspondientes;
- 35 – el indicador de calidad de codificación DS comprende una secuencia de indicadores de calidad de codificación DS para la secuencia de tramas DS (152) correspondientes;
- la unidad (501) de control de tasa está configurada para determinar la tasa de datos IS para una trama IS (151) de la secuencia de tramas IS (151) y la tasa de datos DS para una trama DS de la secuencia de tramas DS (152) en función de la secuencia de indicadores de calidad de codificación IS y la secuencia de indicadores de calidad de codificación DS, de modo que la suma de la tasa de datos IS para la trama IS (151) y la tasa de datos DS para la trama DS es sustancialmente la tasa de datos total disponible.
- 40 6) El codificador de la reivindicación 5, que comprende, además
- una unidad (551) de determinación de dificultad de codificación configurada para determinar el indicador de calidad de codificación IS en función de una primera trama del grupo básico (121) de canales, y/o para determinar el indicador de calidad de codificación DS en función de una primera trama correspondiente del grupo (121) de ampliación de canales.
- 45 7) El codificador de la reivindicación 6, en donde
- el indicador de calidad de codificación IS es uno o más entre: una entropía perceptual de la primera trama del grupo básico (121); una tonalidad de la primera trama del grupo básico (121); un ancho de banda espectral de la primera trama del grupo básico (121); una presencia de
- 50

transitorios en la primera trama del grupo básico (121); un grado de correlación entre canales del grupo básico (121); y una energía de la primera trama del grupo básico (121); y

- 5
  - el indicador de calidad de codificación DS es uno o más entre: una entropía perceptual de la primera trama del grupo (122) de ampliación; una tonalidad de la primera trama del grupo (122) de ampliación; un ancho de banda espectral de la primera trama del grupo (122) de ampliación; una presencia de transitorios en la primera trama del grupo (122) de ampliación; un grado de correlación entre canales del grupo (122) de ampliación; y una energía de la primera trama del grupo (122) de ampliación.
  
- 10 8) Un método para codificar una señal de audio multicanal de acuerdo con una tasa de datos total disponible; en donde la señal de audio multicanal se puede representar como un grupo básico (121) de canales para regularizar la señal de audio multicanal de acuerdo con una configuración básica de canales, y con regularidad como un grupo (122) de ampliación de canales, que – en combinación con el grupo básico (122) – se utiliza para regularizar la señal de audio multicanal de acuerdo con una configuración ampliada de canales; en donde la configuración básica de canales y la configuración ampliada de canales son diferentes entre sí; comprendiendo el método:
  - 15
    - codificar el grupo básico (121) de canales de acuerdo con una tasa de datos IS, con el fin de generar un flujo secundario independiente (110), denominado IS; y
    - codificar el grupo (122) de ampliación de canales de acuerdo con una tasa de datos DS, con el fin de generar un flujo secundario dependiente (120), denominado DS,
  - 20 caracterizado por que el codificador de audio comprende:
    - adaptar con regularidad la tasa de datos IS y la tasa de datos DS en función de un indicador de calidad de codificación IS instantáneo para el grupo básico (121) de canales y/o en función de un indicador de calidad de codificación DS instantáneo para el grupo (122) de ampliación de canales, de modo que la suma de la tasa de datos IS y la tasa de datos DS se corresponde sustancialmente con la tasa de datos total disponible.
  
- 25 9) El método de la reivindicación 8, que comprende, además
  - determinar el indicador de calidad de codificación IS en función de una o más tramas del grupo básico (121) de canales, y/o determinar el indicador de calidad de codificación DS en función de una o más tramas correspondientes del grupo (121) de ampliación de canales.
  
- 30 10) El método de cualquiera de las reivindicaciones 8 a 9, en donde
  - el indicador de calidad de codificación IS es indicativo de una calidad perceptual de una o más tramas del flujo secundario independiente; y
  - el indicador de calidad de codificación DS es indicativo de una calidad perceptual de una o más tramas del flujo secundario dependiente.
  
- 35 11) El método de la reivindicación 10, en donde la adaptación de la tasa de datos IS y la tasa de datos DS comprende
  - 40
    - adaptar la tasa de datos IS y la tasa de datos DS para codificar la una o más tramas del flujo secundario independiente y la una o más tramas del flujo secundario dependiente, de modo que la diferencia absoluta entre el indicador de calidad de codificación IS y el indicador de calidad de codificación DS se encuentre por debajo de un umbral de diferencia.
  
- 45 12) El método de la reivindicación 10, en donde la adaptación de la tasa de datos IS y la tasa de datos DS comprende
  - adaptar la tasa de datos IS y la tasa de datos DS para codificar una o más tramas adicionales del flujo secundario independiente y una o más tramas adicionales correspondientes del flujo secundario dependiente, basándose en que una diferencia entre el indicador de calidad de codificación IS y el indicador de calidad de codificación DS se encuentre por debajo de un umbral de diferencia; en donde la una o más tramas adicionales son posteriores a la una o más tramas.
  
- 13) Un método para decodificar datos de audio codificados, comprendiendo el método:

- generar datos de audio codificados codificando una señal de audio multicanal utilizando un método de acuerdo con cualquiera de las reivindicaciones 8-12;
- recibir los datos de audio codificados en un decodificador; y
- decodificar los datos de audio codificados recibidos para generar una señal indicativa de los datos de audio.

5

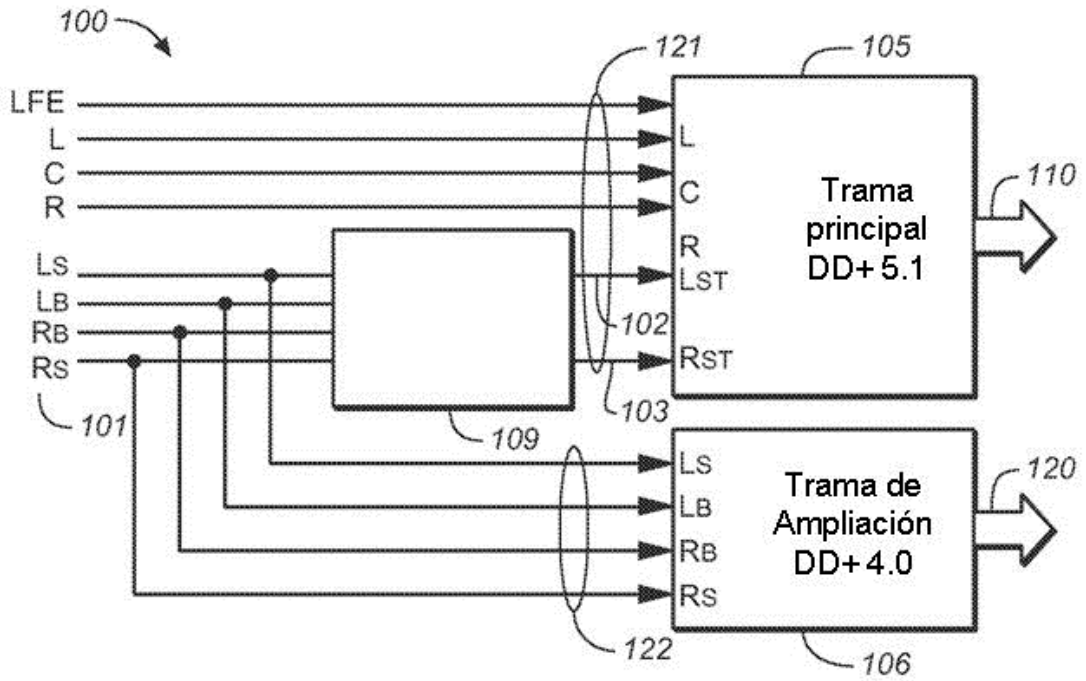
14) Un programa de software adaptado para su ejecución en un procesador y para realizar los pasos del método de cualquiera de las reivindicaciones 8 a 12 cuando se llevan a cabo en el procesador.

15) Un medio de almacenamiento que comprende un programa de software adaptado para su ejecución en un procesador y para realizar los pasos del método de cualquiera de las reivindicaciones 8 a 12 cuando se llevan a cabo en el procesador.

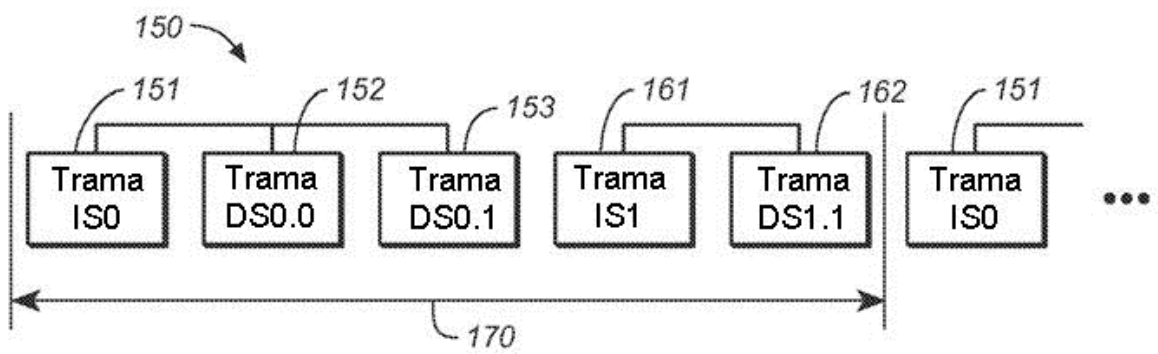
10

16) Un sistema codec de audio, que comprende:

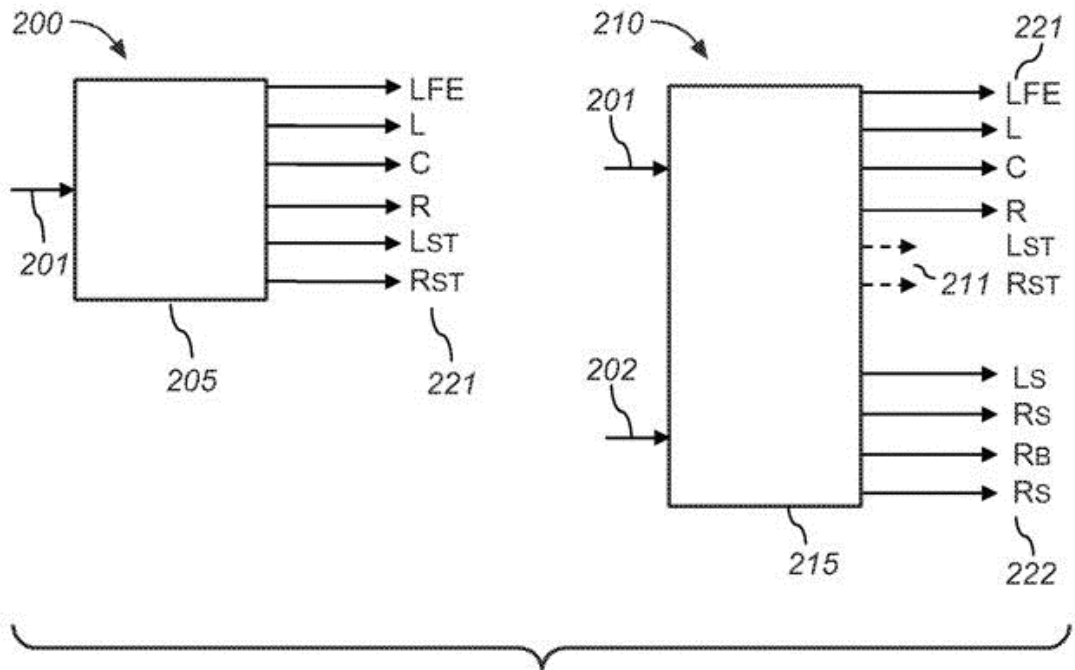
- un codificador de acuerdo con cualquiera de las reivindicaciones 1-7; y
- un decodificador para decodificar datos de audio codificados por el codificador.



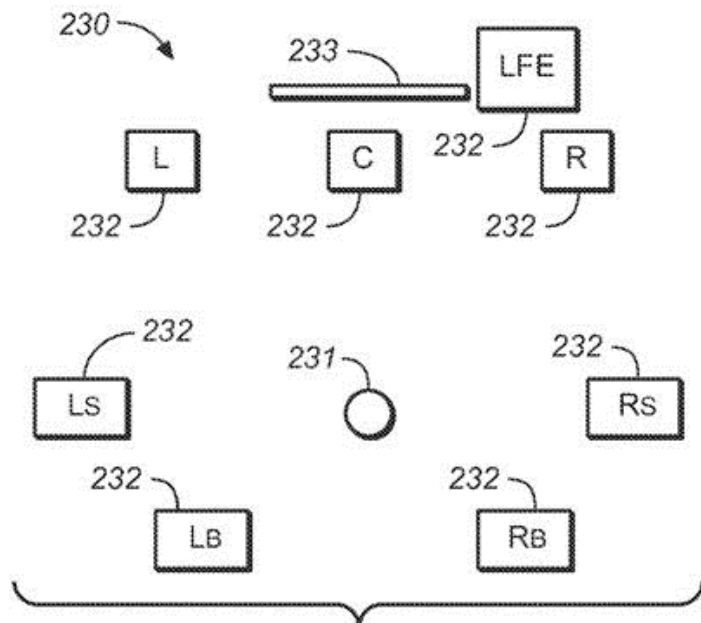
**FIG. 1A**



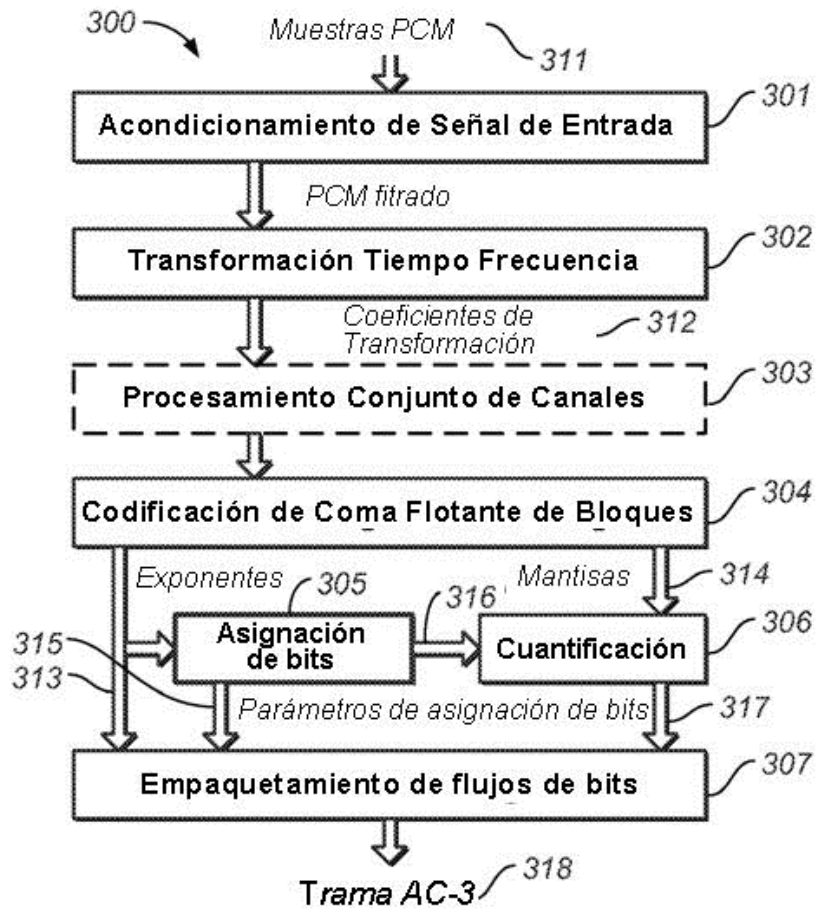
**FIG. 1B**



**FIG. 2A**

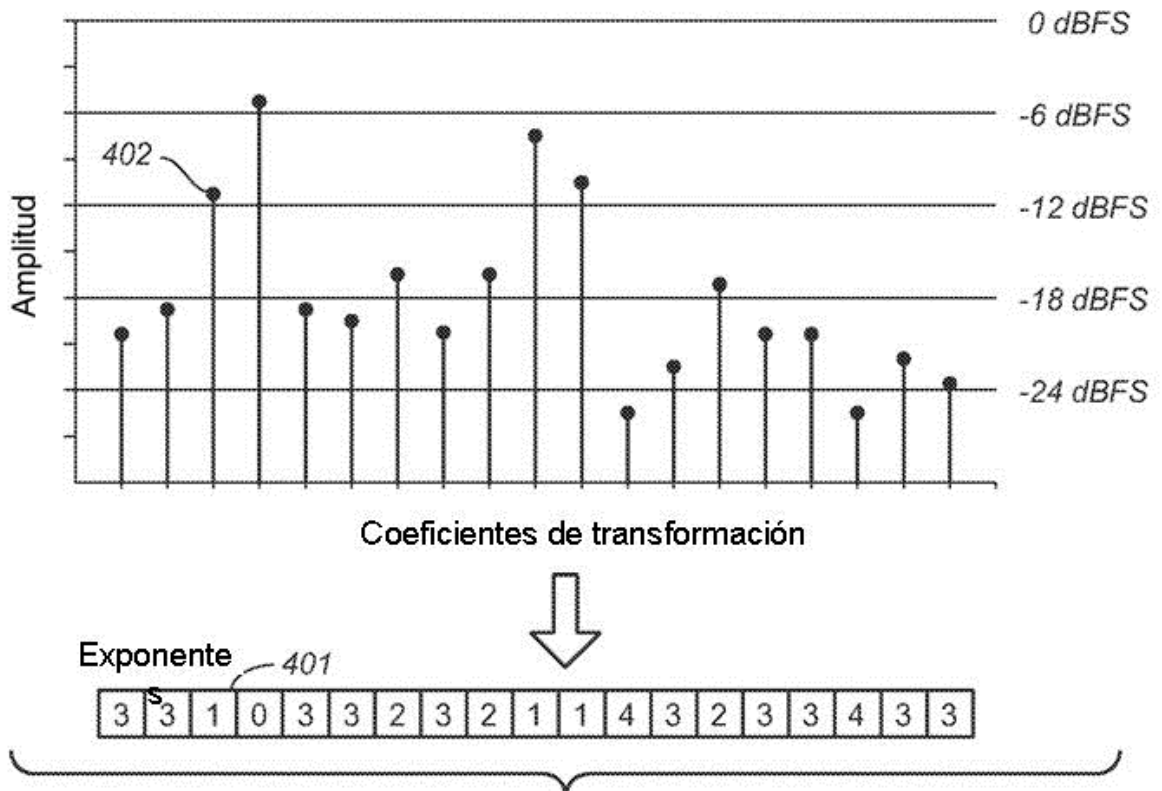


**FIG. 2B**

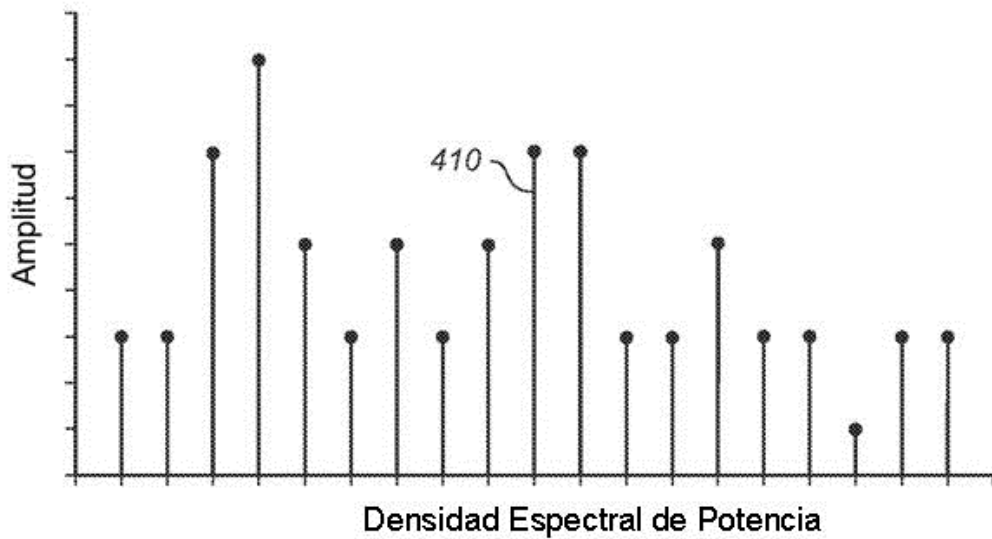


**FIG. 3**

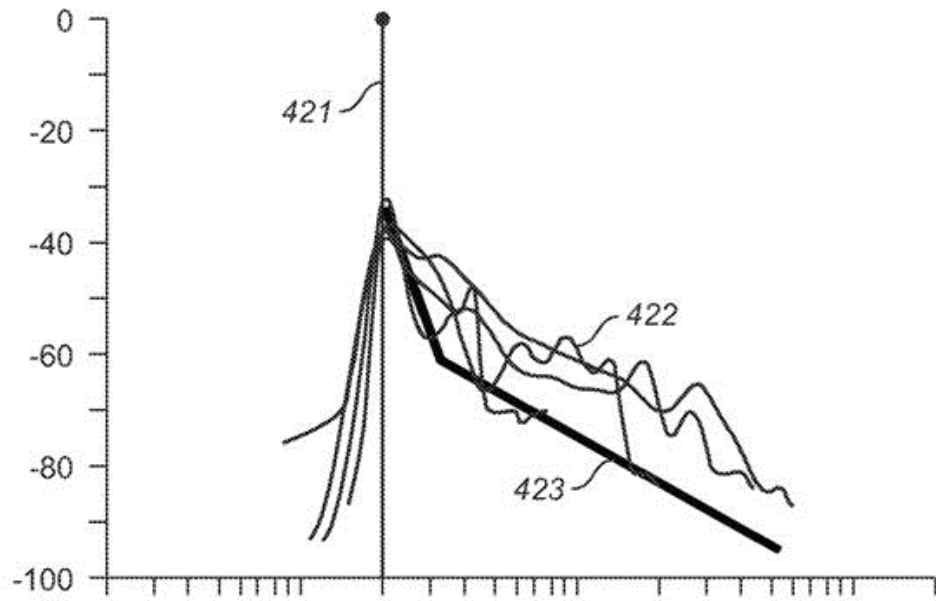




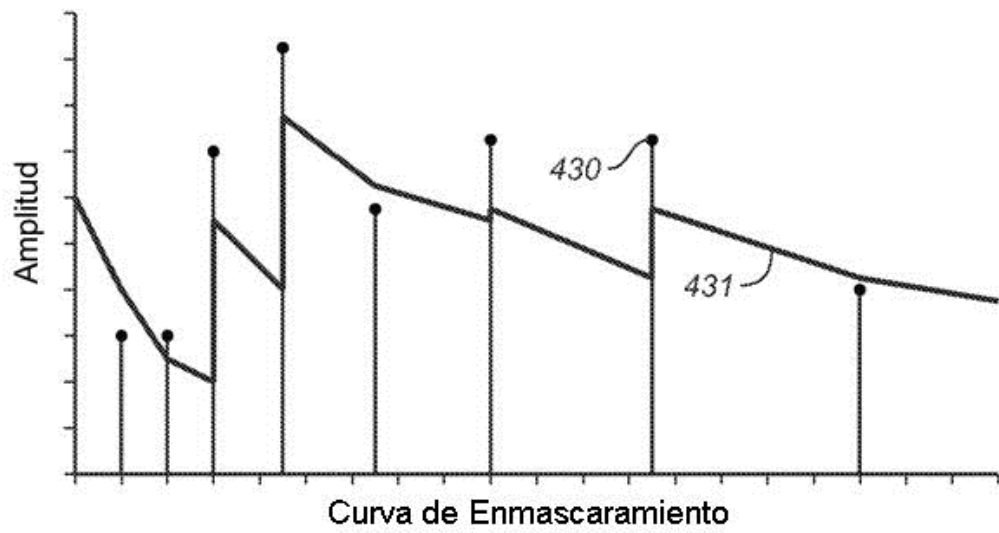
**FIG. 4A**



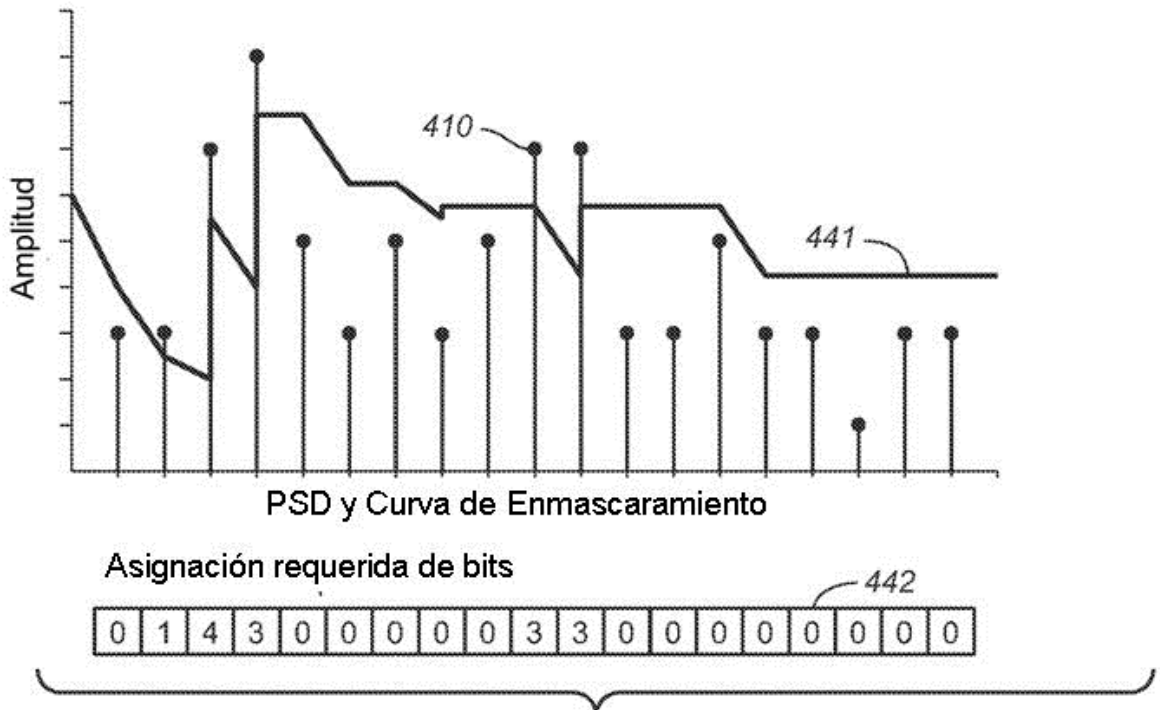
**FIG. 4B**



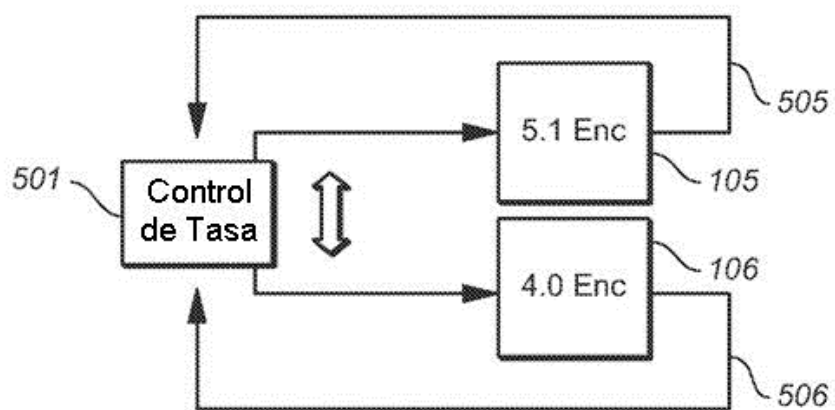
**FIG. 4C**



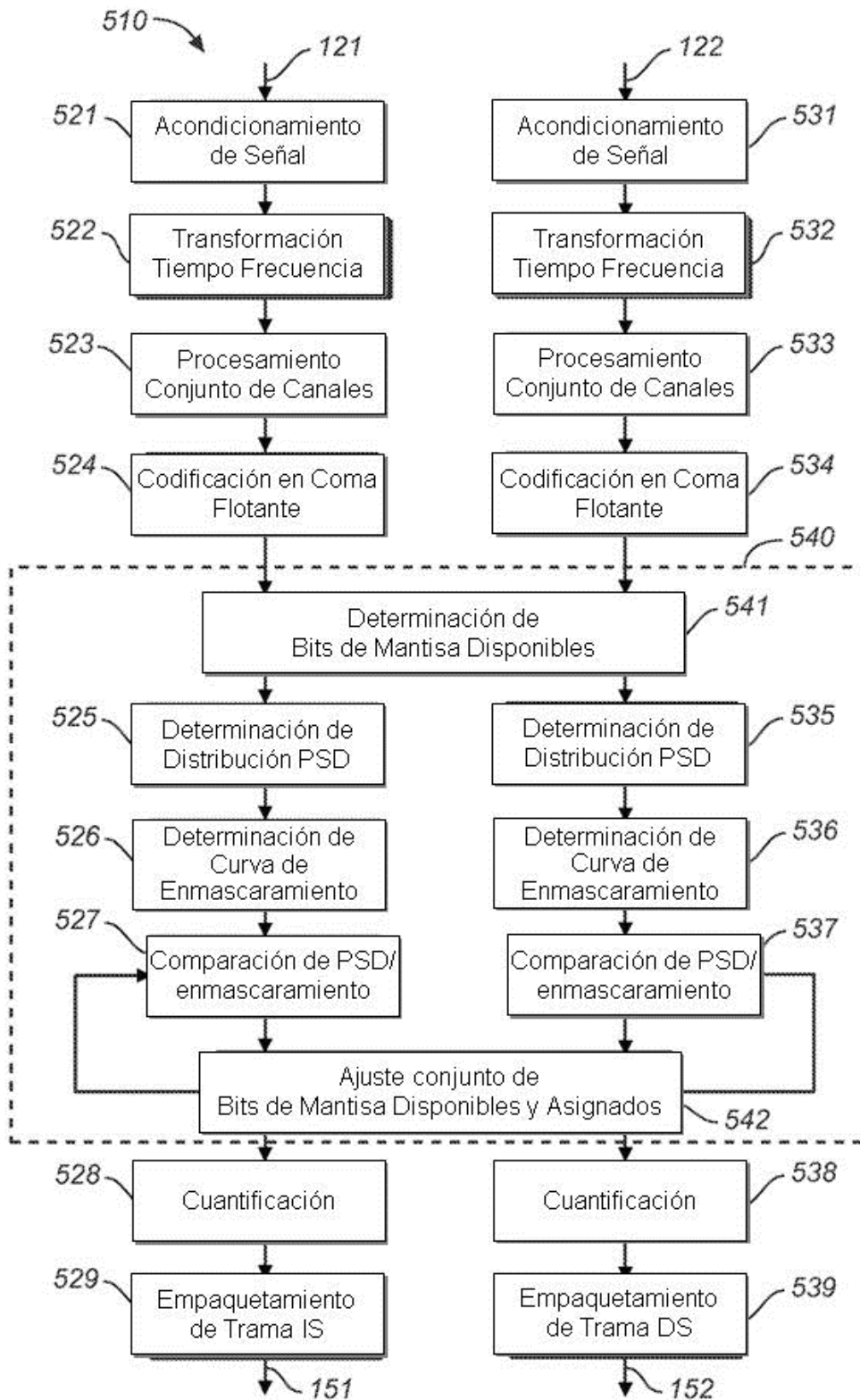
**FIG. 4D**



**FIG. 4E**



**FIG. 5A**



**FIG. 5B**

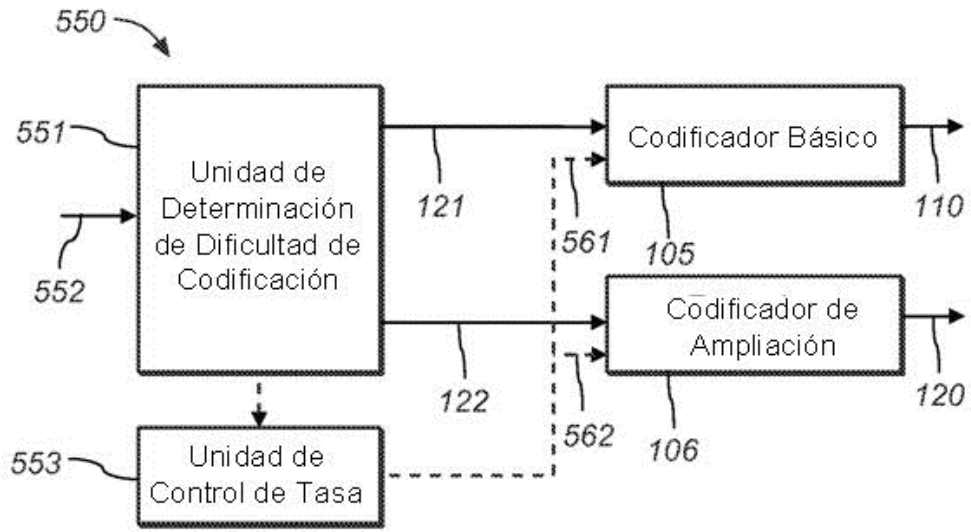


FIG. 5C

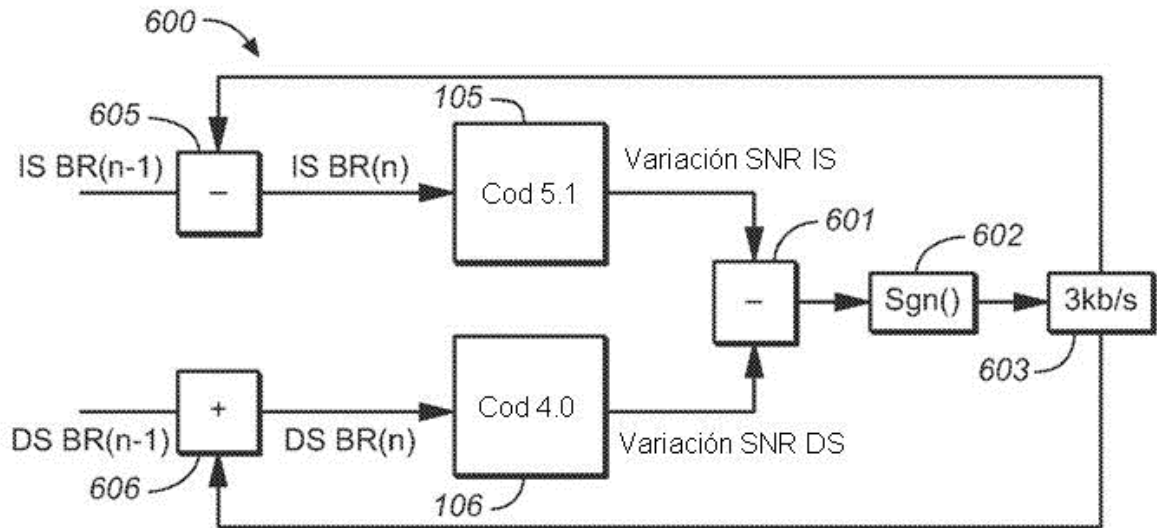


FIG. 6