

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 604 983**

51 Int. Cl.:

**G10L 19/00** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.01.2014 PCT/EP2014/050171**

87 Fecha y número de publicación internacional: **24.07.2014 WO14111290**

96 Fecha de presentación y número de la solicitud europea: **07.01.2014 E 14702195 (0)**

97 Fecha y número de publicación de la concesión europea: **02.11.2016 EP 2946384**

54 Título: **Ajuste de nivel en el dominio del tiempo para decodificación o codificación de señales de audio**

30 Prioridad:

**18.01.2013 EP 13151910**

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

**10.03.2017**

73 Titular/es:

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

72 Inventor/es:

**SCHREINER, STEPHAN;  
BORSUM, ARNE;  
NEUSINGER, MATTHIAS;  
JANDER, MANUEL;  
LOHWASSER, MARKUS y  
NEUGEBAUER, BERNHARD**

74 Agente/Representante:

**ARIZTI ACHA, Monica**

**ES 2 604 983 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).

## Ajuste de nivel en el dominio del tiempo para decodificación o codificación de señales de audio

**DESCRIPCIÓN**

5 La presente invención se refiere a codificación, decodificación, y procesamiento de señales de audio, y, en particular, al ajuste de un nivel de una señal que se va a convertir de frecuencia a tiempo (o que se va a convertir de tiempo a frecuencia) al intervalo dinámico de un convertidor de frecuencia a tiempo correspondiente (o convertidor de tiempo a frecuencia). Algunas realizaciones de la presente invención se refieren al ajuste del nivel de una señal que se va a convertir de frecuencia a tiempo (o que se va a convertir de tiempo a frecuencia) al intervalo dinámico de un convertidor correspondiente implementado en aritmética de punto fijo o enteros. Las realizaciones adicionales de la presente invención se refieren a la prevención de recortes para señales de audio decodificadas espectrales que utilizan ajuste de nivel en el dominio del tiempo en combinación con información adicional.

15 El procesamiento de señales de audio se hace más y más importante. Los desafíos aumentan conforme se necesitan códecs de audio perceptuales modernos para ofrecer una calidad de audio satisfactoria a tasas de bits cada vez más bajas.

En la producción actual de contenido de audio y las cadenas de suministro el contenido maestro digitalmente disponible (flujo PCM (flujo modulado por codificación de impulsos)) se codifica por ejemplo por un codificador AAC profesional (codificación de audio avanzada) en el lado de creación del contenido. El flujo de bits AAC resultante entonces se pone a disposición para la compra por ejemplo a través de una tienda de medios digitales en línea. Parece que en casos raros algunas de las muestras PCM decodificadas se "recortan", lo que significa que dos o más muestras consecutivas alcanzaron el nivel máximo que se puede representar por la resolución de bits subyacente (por ejemplo 16 bits) de una representación de punto fijo cuantificada de forma uniforme (por ejemplo modulada de acuerdo con PCM) para la forma de onda de salida. Esto puede conducir a distorsiones audibles (chasquidos o distorsión corta). Aunque típicamente se hará un esfuerzo en el lado del codificador para evitar la ocurrencia del recorte en el lado del decodificador, sin embargo el recorte puede presentarse en el lado del decodificador por varias razones, tal como diferentes implementaciones del decodificador, errores de redondeo, errores de transmisión, etc. Suponiendo que una señal de audio en la entrada del codificador está por debajo del umbral de recorte, son múltiples las razones del recorte en un codificador de audio perceptual moderno. En primer lugar, el codificador de audio aplica cuantificación a la señal transmitida que está disponible en una descomposición de frecuencia de la forma de onda de entrada para reducir la tasa de datos de transmisión. Los errores de cuantificación en el dominio de la frecuencia dan como resultado pequeñas desviaciones de la amplitud de señal y la fase con respecto a la forma de onda original. Si los errores de fase o amplitud se suman de forma constructiva, la posición resultante en el dominio del tiempo puede ser temporalmente más alta que la forma de onda original. En segundo lugar, los métodos de codificación paramétrica (por ejemplo replicación de banda espectral, SBR) parametrizan la potencia de señal de una forma bastante basta. Típicamente se omite la información de fase. En consecuencia, la señal en el lado del receptor se regenera solamente con la potencia correcta pero sin la conservación de la forma de onda. Las señales con una amplitud cerca a la escala completa son propensas al recorte.

40 Los sistemas de codificación de audio modernos ofrecen la posibilidad de transmitir un parámetro de nivel de sonoridad ( $g_1$ ) dándole a los decodificadores la posibilidad de ajustar la sonoridad para la reproducción con niveles unificados. En general, esto puede conducir al recorte, si la señal de audio se codifica a niveles suficientemente altos y las ganancias de normalización transmitidas sugieren el aumento de los niveles de sonoridad. Además, la práctica común en la masterización del contenido de audio (especialmente música) aumenta las señales de audio a los valores máximos posibles, produciendo el recorte de la señal de audio cuando se cuantifica de forma basta por los códecs de audio.

50 Para evitar el recorte de las señales de audio, se conocen los llamados limitadores como una herramienta apropiada para restringir los niveles de audio. Si una señal de audio entrante excede un cierto umbral, el limitador se activa y atenúa la señal de audio de una forma que la señal de audio no excede un nivel dado a la salida. Desafortunadamente, antes del limitador, se necesita un espacio suficiente (en términos del intervalo dinámico y/o resolución de bits).

55 Por lo general, se logra cualquier normalización de sonoridad en el dominio de la frecuencia junto con un llamado "control de intervalo dinámico" (TRC). Esto permite suavizar la mezcla de normalización de sonoridad e incluso si la ganancia de normalización varía de trama a trama debido a la superposición del banco de filtros.

60 Además, debido a una cuantificación o descripción paramétrica pobre, cualquier señal de audio codificada puede tener recorte si el audio original se masterizó a niveles cercanos al umbral de recorte.

La publicación técnica del ISO/IEC MPEG-2 Advanced Audio Coding por Bosi et al., Journal of the Audio Engineering Society, vol. 45, n. ° 10, octubre de 1997, págs. 789-811, describe las características principales del sistema AAC (ISO/IEC 13818-7). Esta tecnología combina la eficiencia de codificación de un banco de filtros de alta

resolución, técnicas de predicción y codificación Huffman con funcionalidades adicionales que tienen por objeto suministrar calidad de audio muy alta a una diversidad de tasas de datos.

Es típicamente deseable mantener la complejidad computacional, el uso de memoria y el consumo de potencia tan bajos como sea posible en dispositivos de procesamiento de señales de digitales altamente eficientes basándose en una aritmética de punto fijo. Por esta razón, también es deseable mantener la longitud de palabra de las muestras de audio tan baja como sea posible. Para tener en cuenta cualquier espacio potencial para recorte debido a la normalización de la sonoridad, un banco de filtros, que típicamente es una parte de un codificador o decodificador de audio, se tendría que diseñar con una longitud de palabra mayor.

10 Sería deseable permitir la limitación de señal sin perder precisión de datos y/o sin una necesidad de utilizar una longitud de palabra mayor para un banco de filtros del decodificador o un banco de filtros del codificador. Como alternativa o además, sería deseable si se pudiera determinar un intervalo dinámico relevante de la señal que se va a convertir de frecuencia a tiempo o viceversa de forma continua de una forma trama por trama para secciones de tiempo consecutivas o "tramas" de la señal de tal forma que el nivel de la señal se pueda ajustar de una forma que el intervalo dinámico relevante actual se ajuste en el intervalo dinámico proporcionado por el convertidor (convertidor del dominio de la frecuencia al dominio del tiempo o convertidor del dominio del tiempo al dominio de la frecuencia). También sería deseable hacer un cambio de nivel de este tipo para el fin de conversión de frecuencia a tiempo o conversión de tiempo a frecuencia sustancialmente "transparente" a otros componentes del decodificador o codificador.

20 Al menos uno de estos deseos y/o deseos adicionales posibles se tratan mediante un decodificador de señales de audio de acuerdo con la reivindicación 1, un codificador de señales de audio de acuerdo con la reivindicación 14, y un método para decodificar una representación de señal de audio codificada de acuerdo con la reivindicación 15.

25 Se proporciona un decodificador de señales de audio para proporcionar una representación de señal de audio decodificada basándose en una representación de señal de audio codificada. El decodificador de señales de audio comprende una etapa de pre-procesamiento de decodificador configurada para obtener una pluralidad de señales de banda de frecuencia a partir de la presentación de señal de audio codificada. El decodificador de señales de audio comprende además un estimador de recorte configurado para analizar al menos una de la representación de señal de audio codificada, la pluralidad de señales de frecuencia, y la información adicional con respecto a una ganancia de las señales de banda de frecuencia de la representación de señal de audio codificada en cuanto a si la información de señal de audio codificada, la pluralidad de señales de frecuencia, y/o la información adicional sugiere o sugieren un recorte potencial para determinar un factor de cambio de nivel actual para la representación de señal de audio codificada. Cuando la información adicional sugiere el recorte potencial, el factor de cambio de nivel actual provoca que la información de la pluralidad de señales de banda de frecuencia se cambie hacia un bit menos significativo por lo que se gana espacio en al menos un bit más significativo. El decodificador de señales de audio también comprende un cambiador de nivel configurado para cambiar los niveles de las señales de banda de frecuencia de acuerdo con el factor de cambio de nivel para obtener las señales de banda de frecuencia cambiadas de nivel. Además, el decodificador de señales de audio comprende un convertidor del dominio de la frecuencia al dominio del tiempo configurado para convertir las señales de banda de frecuencia cambiadas de nivel a una representación en el dominio del tiempo. El decodificador de señales de audio comprende además un compensador de cambio de nivel configurado para actuar sobre la representación en el dominio del tiempo para compensar al menos parcialmente un cambio de nivel aplicado a las señales de banda de frecuencia cambiadas de nivel por el cambiador de nivel y para obtener una representación en el dominio del tiempo sustancialmente compensada.

45 Las realizaciones adicionales de la presente invención proporcionan un codificador de señales de audio configurado para proporcionar una representación de señal de audio codificada basándose en una representación en el dominio del tiempo de una señal de audio de entrada. El codificador de señales de audio comprende un estimador de recorte configurado para analizar la representación en el dominio del tiempo de la señal de audio de entrada en cuanto a si se sugiere el recorte potencial para determinar un factor de cambio de nivel actual para la presentación de señal de entrada. Cuando se sugiere el recorte potencial, el factor de cambio de nivel actual provoca que la representación en el dominio del tiempo de la señal de audio de entrada se cambie hacia un bit menos significativo por lo que se gana espacio en al menos un bit más significativo. El codificador de señales de audio comprende además un cambiador de nivel configurado para cambiar un nivel de la representación del dominio del tiempo de la señal de audio de entrada de acuerdo con el factor de cambio de nivel para obtener una representación en el dominio del tiempo cambiada de nivel. Además, el codificador de señales de audio comprende un convertidor del dominio del tiempo al dominio de la frecuencia configurado para convertir la representación en el dominio del tiempo cambiada de nivel en una pluralidad de señales de banda de frecuencia. El codificador de señales de audio también comprende un compensador de cambio de nivel configurado para actuar sobre la pluralidad de señales de banda de frecuencia para compensar al menos parcialmente un cambio de nivel aplicado a la presentación en el dominio del tiempo cambiada de nivel mediante el cambiador de nivel y para obtener una pluralidad de señales de banda de frecuencia sustancialmente compensadas.

60 Las realizaciones adicionales de la presente invención proporcionan un método para decodificar la presentación de

señal de audio codificada para obtener una representación de señal de audio decodificada. El método comprende pre-procesar la representación de señal de audio codificada para obtener una pluralidad de señales de banda de frecuencia. El método comprende además analizar al menos una de la representación de señal de audio codificada, las señales de banda de frecuencia, y la información adicional con respecto a una ganancia en las señales de banda de frecuencia en cuanto si se sugiere el recorte potencial para determinar un factor de cambio de nivel actual para la presentación de señal de audio codificada. Cuando se sugiere el recorte potencial, el factor de cambio de nivel actual provoca que la representación en el dominio del tiempo de la señal de audio de entrada se cambie hacia un bit menos significativo por lo que se gana espacio en al menos un bit más significativo. Además, el método comprende cambiar los niveles de las señales de banda de frecuencia de acuerdo con el factor de cambio de nivel para obtener las señales de banda de frecuencia cambiadas de nivel. El método también comprende realizar una conversión del dominio frecuencia al dominio del tiempo de las señales de banda de frecuencia a una representación en el dominio del tiempo. El método comprende además actuar sobre la representación en el dominio del tiempo para compensar al menos parcialmente un cambio de nivel aplicado a las señales de banda de frecuencia cambiadas de nivel y para obtener una representación en el dominio del tiempo sustancialmente compensada.

Además, se proporciona un programa informático para implementar los métodos descritos anteriormente cuando se ejecutan en una computadora o procesador de señales.

Las realizaciones adicionales proporcionan un decodificador de señales de audio para proporcionar una representación de señal de audio decodificada basándose en una representación de señal de audio codificada. El decodificador de señales de audio comprende una etapa de pre-procesamiento de decodificador configurada para obtener una pluralidad de señales de banda de frecuencia a partir de la presentación de señal de audio codificada. El decodificador de señales de audio comprende además un estimador de recorte configurado para analizar al menos una de la representación de señal de audio codificada, la pluralidad de señales de banda de frecuencia, y la información adicional con respecto a una ganancia de las señales de banda de frecuencia de la representación de señal de audio codificada para determinar un factor de cambio de nivel actual para la representación de señal de audio codificada. El decodificador de señales de audio también comprende un cambiador de nivel configurado para cambiar los niveles de las señales de banda de frecuencia de acuerdo con el factor de cambio de nivel para obtener las señales de banda de frecuencia cambiadas de nivel. Además, el decodificador de señales de audio comprende un convertidor del dominio de la frecuencia al dominio del tiempo configurado para convertir las señales de banda de frecuencia cambiadas de nivel a una representación en el dominio del tiempo. El decodificador de señales de audio comprende además un compensador de cambio de nivel configurado para actuar sobre la representación en el dominio del tiempo para compensar al menos parcialmente un cambio de nivel aplicado a las señales de banda de frecuencia cambiadas de nivel por el cambiador de nivel y para obtener una representación en el dominio del tiempo sustancialmente compensada.

Las realizaciones adicionales de la presente invención proporcionan un codificador de señales de audio configurado para proporcionar una representación de señal de audio codificada basándose en una representación en el dominio del tiempo de una señal de audio de entrada. El codificador de señales de audio comprende un estimador de recorte configurado para analizar la representación en el dominio del tiempo de la señal de audio de entrada para determinar un factor de cambio de nivel actual para la representación de señal de entrada. El codificador de señales de audio comprende además un cambiador de nivel configurado para cambiar un nivel de la representación en el dominio del tiempo de la señal de audio de entrada de acuerdo con el factor de cambio de nivel para obtener una representación en el dominio del tiempo cambiada de nivel. Además, el codificador de señales de audio comprende un convertidor del dominio del tiempo al dominio de la frecuencia configurado para convertir la representación en el dominio del tiempo cambiada de nivel en una pluralidad de señales de banda de frecuencia. El codificador de señales de audio también comprende un compensador de cambio de nivel configurado para actuar sobre la pluralidad de señales de banda de frecuencia para compensar al menos parcialmente un cambio de nivel aplicado a la presentación en el dominio del tiempo cambiada de nivel por el cambiador de nivel y para obtener una pluralidad de señales de banda de frecuencia sustancialmente compensadas.

Las realizaciones adicionales de la presente invención proporcionan un método para decodificar la presentación de señal de audio codificada para obtener una representación de señal de audio decodificada. El método comprende pre-procesar la representación de señal de audio codificada para obtener una pluralidad de señales de banda de frecuencia. El método comprende además analizar al menos una de la representación de señal de audio codificada, las señales de banda de frecuencia, y se sugiere información adicional con respecto a una ganancia de las señales de banda de frecuencia para determinar un factor de cambio actual para la presentación de señal de audio codificada. Además, el método comprende cambiar los niveles de las señales de banda de frecuencia de acuerdo con el factor de cambio de nivel para obtener las señales de banda de frecuencia cambiadas de nivel. El método también comprende realizar una conversión del dominio de la frecuencia al dominio del tiempo de las señales de banda de frecuencia a una representación en el dominio del tiempo. El método comprende además actuar sobre la representación en el dominio del tiempo para compensar al menos parcialmente un cambio de nivel aplicado a las señales de banda de frecuencia cambiadas de nivel y para obtener una representación en el dominio del tiempo sustancialmente compensada.

Al menos algunas de las realizaciones están basadas en la idea de que es posible, sin perder información relevante, cambiar la pluralidad de señales de banda de frecuencia de una representación en el dominio de la frecuencia mediante un cierto factor de cambio de nivel durante intervalos de tiempo, en los cuales es relativamente alto un nivel general de sonoridad de la señal de audio. En su lugar, la información relevante se cambia a bits que son susceptibles de contener ruido, de cualquier modo. De esta forma, un convertidor del dominio de la frecuencia al dominio del tiempo que tiene una longitud de palabra limitada se puede utilizar a pesar de que un intervalo dinámico de las señales de banda de frecuencia puede ser más grande que el soportado por la longitud de palabra limitada del convertidor del dominio de la frecuencia al dominio del tiempo. En otras palabras, al menos algunas realizaciones de la presente invención aprovechan el hecho de que el bit o bits menos significativos típicamente no transportan alguna información relevante mientras la señal de audio es relativamente sonora, es decir, mientras la información relevante es más probable que esté contenida en el bit o bits más significativos. El cambio de nivel aplicado a las señales de banda de frecuencia cambiadas de nivel también puede tener el beneficio de reducir que se presente una probabilidad de recorte dentro de la representación en el dominio del tiempo, donde dicho recorte puede resultar a partir de una superposición constructiva de una o más señales de banda de frecuencia de la pluralidad de señales de banda de frecuencia.

Estas ideas y hallazgos también se aplican de una forma análoga al codificador de señales de audio y el método para codificar una señal de audio original para obtener una presentación de señal de audio codificada.

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

La Figura 1 ilustra un codificador de acuerdo con el estado de la técnica;

La Figura 2 representa un decodificador de acuerdo con el estado de la técnica;

La Figura 3 ilustra otro codificador de acuerdo con el estado de la técnica;

La Figura 4 representa un decodificador adicional de acuerdo con el estado de la técnica;

La Figura 5 muestra un diagrama de bloques esquemático de un decodificador de señales de audio de acuerdo con al menos una realización;

La Figura 6 muestra un diagrama de bloques esquemático de un decodificador de señales de audio de acuerdo con al menos una realización adicional;

La Figura 7 muestra un diagrama de bloques que ilustra un concepto del decodificador de señales de audio propuesto y el método propuesto para decodificar una representación de señal de audio codificada de acuerdo con las realizaciones;

La Figura 8 es una visualización esquemática del cambio de nivel para ganar espacio;

La Figura 9 muestra un diagrama de bloques esquemático de un ajuste de forma de transición posible que puede ser un componente del decodificador o codificador de señales de audio de acuerdo con al menos a algunas realizaciones;

La Figura 10 representa una unidad de estimación de acuerdo con una realización adicional que comprende un ajustador de filtro de predicción;

La Figura 11 ilustra un aparato para generar un flujo de datos de regreso;

La Figura 12 ilustra un codificador de acuerdo con el estado de la técnica;

La Figura 13 representa un decodificador de acuerdo con el estado de la técnica;

La Figura 14 ilustra otro codificador de acuerdo con el estado de la técnica; y

La Figura 15 muestra un diagrama de bloques esquemático de un codificador de señales de audio de acuerdo con al menos una realización; y

La Figura 16 muestra un diagrama de flujo esquemático de un método para decodificar la representación de señal de audio codificada de acuerdo con al menos una realización.

El procesamiento de audio ha avanzado en muchos aspectos y ha sido objeto de varios estudios, la forma de codificar y decodificar de forma eficiente de una señal de datos de audio. La codificación eficiente, por ejemplo, se proporciona mediante MPEG AAC (MPEG = grupo de expertos en imágenes en movimiento; AAC = codificación de audio avanzada). Algunos aspectos de MPEG AAC se explican en más detalle más adelante, como una introducción a la codificación y decodificación de audio. La descripción de MPEG AAC solamente se ha de entender como un ejemplo, puesto que los conceptos descritos se pueden aplicar a otros esquemas de codificación y decodificación de audio, también.

De acuerdo con MPEG AAC, los valores espectrales de una señal de audio se codifican empleando factores de escala, cuantificación y libros de códigos, en particular libros de código Huffman.

Antes que se lleve a cabo la codificación Huffman, el codificador agrupa la pluralidad de coeficientes espectrales que se van a codificar en diferentes secciones (los coeficientes espectrales que se han obtenido a partir de componentes corriente arriba, tal como un banco de filtros, un modelo psico-acústico, y un cuantificador controlado por el modelo psico-acústico con respecto a las resoluciones de cuantificación y los umbrales de cuantificación). Para cada sección de coeficientes espectrales, el codificador elige un libro de códigos Huffman para codificación Huffman. La MPEG AAC proporciona once libros de códigos Huffman de espectros diferentes para codificar los datos espectrales a partir de los cuales el codificador selecciona el libro de códigos que mejor se ajusta para codificar los coeficientes espectrales de la sección. El codificador proporciona un identificador del libro de códigos que identifica el libro de códigos utilizado para la codificación Huffman de los coeficientes espectrales de la sección al decodificador como información adicional.

En un lado del decodificador, el decodificador analiza la información adicional recibida para determinar cuál de la pluralidad de libros de códigos Huffman de espectro se ha utilizado para codificar los valores espectrales de una sección. El decodificador lleva a cabo la decodificación Huffman basándose en la información adicional acerca del libro de códigos Huffman empleado para codificar los coeficientes espectrales de la sección que se va a decodificar por el decodificador.

Después de la decodificación Huffman, se obtiene una pluralidad de valores espectrales cuantificados en el decodificador. El decodificador entonces puede llevar a cabo cuantificación inversa para invertir una cuantificación no uniforme que pudo haber sido llevada a cabo por el codificador. Mediante esto, se obtienen los valores espectrales cuantificados de forma inversa en el decodificador.

Sin embargo, los valores espectrales cuantificados de forma inversa pueden aún no estar modificados a escala. Los valores espectrales no modificados a escala, derivados se han agrupado en bandas de factor de escala, cada banda de factor de escala que tiene un factor de escala común. El factor de escala para cada banda de factor de escala está disponible para el decodificador como información adicional, que se ha proporcionado por el codificador. Al usar esta información, el decodificador multiplica los valores espectrales no modificados a escala de una banda de factor de escala por su factor de escala. Mediante esto, se obtienen los valores espectrales modificados a escala.

La codificación y decodificación de los valores espectrales de acuerdo con el estado de la técnica se explica ahora con referencia a las Figuras 1-4.

La Figura 1 ilustra un codificador de acuerdo con el estado de la técnica. El codificador comprende un banco de filtros T/F (tiempo a frecuencia) 10 para transformar una señal de audio AS, que será codificada, del dominio del tiempo a un dominio de la frecuencia para obtener una señal de audio en el dominio de la frecuencia. La señal de audio en el dominio de la frecuencia se alimenta en una unidad de factor de escala 20 para determinar los factores de escala. La unidad de factor de escala 20 se adapta para dividir los coeficientes espectrales de la señal de audio en el dominio de la frecuencia en varios grupos de coeficientes espectrales llamados bandas de factor de escala, que comparten un factor de escala. Un factor de escala representa un valor de ganancia utilizado para cambiar la amplitud de todos los coeficientes espectrales en la banda de factor de escala respectiva. La unidad de factor de escala 20 se adapta además para generar y emitir coeficientes espectrales no modificados a escala de la señal de audio en el dominio de la frecuencia.

Además, el codificador en la Figura 1 comprende un cuantificador para cuantificar los coeficientes espectrales no modificados a escala de la señal de audio en el dominio de la frecuencia. El cuantificador 30 puede ser un cuantificador no uniforme.

Después de la cuantificación, los espectros no modificados a escala, cuantificados de la señal de audio se alimentan en un codificador Huffman 40 para someterse a codificación Huffman. La codificación Huffman se utiliza para reducir la redundancia del espectro cuantificado de la señal de audio. La pluralidad de coeficientes espectrales cuantificados no modificados a escala se agrupa en secciones. Mientras en la MPEG AAC se proporcionan once libros de códigos posibles, todos los coeficientes espectrales de una sección se codifican por el mismo libro de códigos Huffman.

El codificador elegirá uno de los once libros de códigos Huffman posibles que es particularmente adecuado para codificar los coeficientes espectrales de la sección. Mediante esto, la selección del libro de códigos Huffman del codificador para una sección particular depende de los valores espectrales de la sección particular. Los coeficientes espectrales sometidos a codificación Huffman entonces se pueden transmitir al decodificador junto con información adicional que comprende, por ejemplo, información acerca del libro de códigos Huffman que se ha utilizado para codificar una sección de coeficientes espectrales, un factor de escala que se ha utilizado para una banda de factor de escala particular, etc.

Se codifican dos o cuatro coeficientes espectrales por una palabra de código del libro de códigos Huffman empleado para codificación Huffman de los coeficientes espectrales de la sección. El codificador transmite las palabras de código que representan los coeficientes espectrales codificados al decodificador junto con información adicional que comprende la longitud de una sección así como información acerca del libro de códigos Huffman utilizado para codificar los coeficientes espectrales de la sección.

En la MPEG AAC, se proporcionan once libros de códigos Huffman de espectro para codificar los datos espectrales de la señal de audio. El libro de códigos Huffman de espectro diferente se puede identificar por su índice el libro de códigos (un valor entre 1 y 11). La dimensión del libro de códigos Huffman indica cuántos coeficientes espectrales se codifican por una palabra de código del libro de códigos Huffman considerado. En la MPEG AAC, la dimensión de un libro de códigos Huffman indica, ya sea 2 o 4, que una palabra de código codifica ya sea dos o cuatro valores espectrales de la señal de audio.

Sin embargo, los libros de códigos Huffman diferentes también difieren con respecto a otras propiedades. Por ejemplo, el valor absoluto máximo de un coeficiente espectral que se puede codificar por el libro de códigos Huffman varía de libro de códigos a libro de códigos y puede ser, por ejemplo, 1, 2, 4, 7, 12 o mayor. Además, un libro de códigos Huffman considerado se puede adaptar para codificar valores con signo o no.

Al emplear codificación Huffman, los coeficientes espectrales se codifican por palabras de códigos de diferentes longitudes. La MPEG AAC proporciona dos libros de códigos Huffman diferentes que tienen un valor máximo absoluto de 1, dos libros de códigos Huffman diferentes que tienen un valor máximo absoluto de 2, dos libros de códigos Huffman diferentes que tienen un valor máximo absoluto de 4, dos libros de códigos Huffman diferentes que tienen un valor máximo absoluto de 7 y dos libros de códigos Huffman diferentes que tienen un valor máximo absoluto de 12, en donde cada libro de códigos Huffman representa una función de distribución de probabilidad diferente. El codificador Huffman siempre elegirá el libro de códigos Huffman que mejor se ajusta para la codificación de los coeficientes espectrales.

La Figura 2 ilustra un decodificador de acuerdo con el estado de la técnica. Los valores espectrales sometidos a codificación Huffman se reciben por un decodificador Huffman 50. El decodificador Huffman 50 también recibe, como información adicional, información acerca del libro de códigos Huffman utilizado para codificar los valores espectrales para cada sección de valores espectrales. El decodificador Huffman 50 entonces realiza decodificación Huffman para obtener los valores espectrales cuantificados no modificados a escala. Los valores espectrales cuantificados, no modificados a escala se alimentan en un cuantificador inverso 60. El cuantificador inverso realiza cuantificación inversa para obtener los valores espectrales no modificados a escala, cuantificados de forma inversa, que se alimentan a un modificador de escala 70. El modificador de escala 70 también recibe factores de escala como información adicional para cada banda de factor de escala. Basándose en los factores de escala recibidos, el modificador de escala 70 modifica la escala de los valores espectrales cuantificados de forma inversa no modificados a escala para obtener los valores espectrales cuantificados de forma inversa, modificados a escala. Un banco de filtros F/T 80 entonces transforma los valores espectrales cuantificados de forma inversa, modificados a escala de la señal de audio en el dominio de la frecuencia a partir del dominio de la frecuencia al dominio del tiempo para obtener valores de muestra de una señal de audio en el dominio del tiempo.

La Figura 3 ilustra un codificador de acuerdo con el estado de la técnica que difiere del codificador de la Figura 1 en que el codificador de la Figura 3 comprende además una unidad TNS del lado del codificador (TNS = conformación de ruido temporal). La conformación de ruido temporal se puede emplear para controlar la forma temporal del ruido de cuantificación al llevar a cabo un proceso de filtrado con respecto a las porciones de los datos espectrales de la señal de audio. La unidad TNS del lado del codificador 15 lleva a cabo un cálculo de codificación predictiva lineal (LPC) con respecto a los coeficientes espectrales de la señal de audio en el dominio de la frecuencia que se va a codificar. Entre otros, los que resultan a partir del cálculo LPC son los coeficientes de reflexión, también llamados como coeficientes PARCOR. La conformación de ruido temporal no se utiliza si la ganancia de predicción, que también se deriva por el cálculo LPC, no excede un cierto valor umbral. Sin embargo, si la ganancia de predicción es mayor que el valor umbral, se emplea la conformación de ruido temporal. La unidad TNS del lado del codificador elimina todos los coeficientes de reflexión que son menores que un cierto valor umbral. Los coeficientes de reflexión restantes se convierten en coeficientes de predicción lineal y se utilizan como coeficientes de filtro de conformación de ruido en el codificador. La unidad TNS del lado del codificador entonces realiza una operación de filtro en aquellos coeficientes espectrales, para los cuales se emplea TNS, para obtener coeficientes espectrales procesados

de la señal de audio. La información adicional que indica información TNS, por ejemplo, los coeficientes de reflexión (coeficientes PARCOR) se transmite al decodificador.

5 La Figura 4 ilustra un decodificador de acuerdo con el estado de la técnica que difiere del decodificador ilustrado en la Figura 2 en la medida en que el decodificador de la Figura 4 comprende además una unidad TNS del lado del decodificador 75. La unidad TNS del lado del decodificador recibe espectros modificados a escala cuantificados de forma inversa de la señal de audio y también recibe información TNS, por ejemplo, información que indica los coeficientes de reflexión (coeficientes PARCOR). La unidad TNS del lado del decodificador 75 procesa los espectros cuantificados de forma inversa de la señal de audio para obtener un espectro cuantificado de forma inversa  
10 procesado de la señal de audio.

La Figura 5 muestra un diagrama de bloques esquemático de un decodificador de señales de audio 100 de acuerdo con al menos una realización de la presente invención. El decodificador de señales de audio se configura para recibir una representación de señal de audio codificada. Típicamente, la presentación de señal de audio codificada está acompañada por información adicional. La representación de señal de audio codificada junto con la información  
15 adicional se puede proporcionar en la forma de un flujo de datos que se ha producido por, por ejemplo, un codificador de audio perceptual. El decodificador de señales de audio 100 se configura además para proporcionar una representación de señal de audio decodificada que puede ser idéntica a la señal etiquetada "representación en el dominio del tiempo sustancialmente compensada" en la Figura 5 o derivada a partir de la misma utilizando  
20 procesamiento posterior.

El decodificador de señales de audio 100 comprende una etapa de pre-procesamiento de decodificador 110 que se configura para obtener una pluralidad de señales de banda de frecuencia a partir de la representación de señal de audio codificada. Por ejemplo, la etapa de pre-procesamiento de decodificador 110 puede comprender un  
25 desempaquetador de flujo de bits en caso de que la representación de señal de audio codificada y la información adicional estén contenidas en un flujo de bits. Algunas de las normas de codificación de audio pueden utilizar resoluciones que varían en el tiempo y también resoluciones diferentes para la pluralidad de señales de banda de frecuencia, dependiendo del intervalo de frecuencia en el cual la presentación de señal de audio codificada lleve actualmente información relevante (alta resolución) o información irrelevante (baja resolución o ningún dato en absoluto). Esto significa que una banda de frecuencia en la cual la representación de señal de audio codificada tiene  
30 actualmente una gran cantidad de información relevante se codifica típicamente utilizando una resolución altamente precisa (es decir, que utiliza un número relativamente alto de bits) durante ese intervalo de tiempo, a diferencia de una señal de banda de frecuencia que lleva temporalmente ninguna o solamente muy poca información. Incluso puede suceder que para algunas de las señales de banda de frecuencia que el flujo de bits no contenga de forma temporal datos o bits, en absoluto, debido a que estas señales de banda de frecuencia no contienen ninguna información relevante durante el intervalo de tiempo correspondiente. El flujo de bits proporcionado a la etapa de pre-procesamiento de decodificador 110 contiene típicamente información (por ejemplo, como parte de la información adicional) que indica cuáles señales de banda de frecuencia de la pluralidad de señales de banda de frecuencia contienen datos para el intervalo de tiempo o "trama" actualmente considerada y la resolución de bits  
35 correspondiente.  
40

El decodificador de señales de audio 100 comprende además un estimador de recorte 120 configurado para analizar la información adicional con respecto a una ganancia de las señales de banda de frecuencia de la representación de señal de audio codificada para determinar un factor de cambio de nivel actual para la representación de señal de audio codificada. Algunas normas de codificación de audio perceptual utilizan factores de escala individuales para las diferentes señales de banda de frecuencia de la pluralidad de señales de banda de frecuencia. Los factores de escala individuales indican para cada señal de banda de frecuencia el intervalo de amplitud actual, con respecto a las otras señales de banda de frecuencia. Para algunas realizaciones de la presente invención un análisis de estos factores de escala permite una evaluación aproximada de una amplitud máxima que puede tener lugar en una  
45 representación en el dominio del tiempo correspondiente después de que la pluralidad de señales de banda de frecuencia se ha convertido de un dominio de la frecuencia a un dominio del tiempo. Esta información entonces se puede utilizar para determinar si, sin ningún procesamiento apropiado como se propone por la presente invención, sería probable que tuviera lugar el recorte dentro de la representación del dominio del tiempo para el intervalo de tiempo o "trama" considerada. El estimador de recorte 120 se configura para determinar un factor de cambio de nivel que cambia todas las señales de banda de frecuencia de la pluralidad de señales de banda de frecuencia en una cantidad idéntica con respecto al nivel (con respecto a una amplitud de señal o potencia de señal, por ejemplo). El factor de cambio de nivel se puede determinar para cada intervalo de tiempo (trama) de una forma individual, es decir, el factor de cambio de nivel es variable en el tiempo. Típicamente, el estimador de recorte 120 intentará  
50 ajustar los niveles de la pluralidad de señales de banda de frecuencia mediante el factor de cambio que es común para todas las señales de banda de frecuencia de una forma que es muy poco probable que tenga lugar recorte dentro de la representación del dominio del tiempo, pero al mismo tiempo manteniendo un intervalo dinámico razonable para las señales de banda de frecuencia. Como un ejemplo, se considera una trama de la representación de la señal de audio codificada en la cual un número de los factores de escala es relativamente alto. El estimador de recorte 120 ahora puede considerar el peor de los casos, es decir, picos de señal posibles dentro de la pluralidad de  
55  
60



señales de banda de frecuencia que se superponen o suman de una forma constructiva, dando como resultado una amplitud grande dentro de la representación en el dominio del tiempo. El factor de cambio de nivel ahora se puede determinar como un número que provoca que este pico hipotético dentro de la representación en el dominio del tiempo esté dentro de un intervalo dinámico deseado, posiblemente con la consideración adicional de un margen. Al menos de acuerdo con algunas realizaciones el estimador de recorte 120 no necesita la propia representación de señal de audio codificada para evaluar una probabilidad de recorte dentro de la representación en el dominio del tiempo para el intervalo de tiempo o trama considerada. La razón es que al menos algunas normas de codificación de audio perceptual eligen los factores de escala para las señales de banda de frecuencia para la pluralidad de señales de banda de frecuencia de acuerdo con la amplitud más grande que se tiene que codificar dentro de una cierta señal de banda de frecuencia y el intervalo de tiempo considerado. En otras palabras, el volumen más alto que se puede representar por la resolución de bits elegida para la señal de banda de frecuencia disponible, es muy probable que se tenga lugar al menos una vez durante el intervalo de tiempo o trama considerada, dadas las propiedades del esquema de codificación. Al usar esta suposición, el estimador de recorte 120 puede centrarse en evaluar la información adicional con respecto a la ganancia o ganancias de las señales de banda de frecuencia (por ejemplo, dicho factor de escala y los parámetros posiblemente adicionales) para determinar el factor de cambio de nivel actual para la representación de señal de audio codificada y el intervalo de tiempo considerado (trama).

El decodificador de señales de audio 100 comprende además un cambiador de nivel 130 configurado para cambiar los niveles de las señales de banda de frecuencia de acuerdo con el factor de cambio de nivel para obtener las señales de banda de frecuencia cambiadas de nivel.

El decodificador de señales de audio 100 comprende además un convertidor del dominio de la frecuencia al dominio del tiempo 140 configurado para convertir las señales de banda de frecuencia cambiadas de nivel en una representación en el dominio del tiempo. El convertidor del dominio de la frecuencia al dominio del tiempo 140 puede ser un banco de filtros inverso, una transformación discreta de coseno modificada inversa (MDCT inversa), un filtro de espejo en cuadratura inverso (QMF inverso), por nombrar algunos. Para algunas normas de codificación de audio el convertidor del dominio de la frecuencia al dominio del tiempo 140 se puede configurar para soportar formación de ventanas de tramas consecutivas, en el que dos tramas se superponen durante, por ejemplo, el 50 % de su duración.

La representación en el dominio del tiempo proporcionada por el convertidor del dominio de la frecuencia al dominio del tiempo 140 se proporciona a un compensador de cambio de nivel 150 que se configura para actuar sobre la representación en el dominio del tiempo para compensar al menos parcialmente un cambio de nivel aplicado a las señales de banda de frecuencia cambiadas de nivel por el cambiador de nivel 130, y para obtener una representación en el dominio del tiempo sustancialmente compensada. El compensador de cambio de nivel 150 recibe además el factor de cambio de nivel del estimador de recorte 140 o una señal derivada del factor de cambio de nivel. El cambiador de nivel 130 y el compensador de cambio de nivel 150 proporcionan un ajuste de ganancia de las señales de banda de frecuencia cambiadas de nivel y un ajuste de ganancia de compensación de la presentación en el dominio del tiempo, respectivamente, en el que dicho ajuste de ganancia evita el convertidor del dominio de la frecuencia al dominio del tiempo 140. De esta forma, las señales de banda de frecuencia cambiadas de nivel y la representación en el dominio del tiempo se pueden ajustar a un intervalo dinámico proporcionado por el convertidor del dominio de la frecuencia al dominio del tiempo 140 que se puede limitar debido a una longitud de palabra fija y/o una implementación aritmética de punto fijo del convertidor 140. En particular, el intervalo dinámico relevante de las señales de banda de frecuencia cambiadas de nivel y la representación en el dominio del tiempo correspondiente pueden estar en niveles de potencia de señal o valores de amplitud relativamente altos durante tramas relativamente sonoras. Por el contrario, el intervalo dinámico relevante de la señal de banda de frecuencia cambiada de nivel y en consecuencia también de la representación en el dominio del tiempo correspondiente pueden estar en valores de potencia de señal o valores de amplitud relativamente pequeños durante tramas relativamente suaves. En el caso de tramas sonoras, la información contenida en los bits inferiores de una representación binaria de las señales de banda de frecuencia cambiadas de nivel puede considerarse típicamente como despreciable en comparación con la información que está contenida dentro de los bits superiores. Típicamente, el factor de cambio de nivel es común para todas las señales de banda de frecuencia lo que hace posible compensar el cambio de nivel aplicado a las señales de banda de frecuencia cambiadas de nivel incluso corriente abajo del convertidor del dominio de la frecuencia al dominio del tiempo 140. A diferencia del factor de cambio de nivel propuesto que se determina por el propio decodificador de señal de audio 100, el llamado parámetro de ganancia global está contenido dentro del flujo de bits que se produjo por un codificador de señales de audio remoto y se proporcionó al decodificador de señales de audio 100 como una entrada. Además, la ganancia global se aplica a la pluralidad de señales de banda de frecuencia entre la etapa de pre-procesamiento de decodificador 110 y el convertidor del dominio de la frecuencia al dominio del tiempo 140. Típicamente, la ganancia global se aplica a la pluralidad de señales de banda de frecuencia en sustancialmente el mismo lugar dentro de la cadena de procesamiento de señales como los factores de escala para las diferentes señales de banda de frecuencia. Esto significa que para una trama relativamente sonora las señales de banda de frecuencia proporcionadas al convertidor del dominio de la frecuencia al dominio del tiempo 140 son ya relativamente sonoras, y por lo tanto pueden provocar recorte en la representación en el dominio del tiempo correspondiente, debido a que la pluralidad de señales de

banda de frecuencia no proporcionaron suficiente espacio en caso de que las señales de banda de frecuencia diferentes se sumaran de una forma constructiva, conduciendo por lo tanto a una amplitud de señal relativamente alta dentro de la representación en el dominio del tiempo.

5 El enfoque propuesto, es decir, por ejemplo implementado por el decodificador de señales de audio 100 ilustrado esquemáticamente en la Figura 5, permite la limitación de la señal sin perder precisión de datos o utilizar longitudes de palabra superiores para los bancos de filtros de decodificador (por ejemplo, el convertidor del dominio de la frecuencia al dominio del tiempo 140).

10 Para superar el problema de la longitud de palabra restringida de los bancos de filtros, la normalización de sonoridad como fuente de recorte potencial se puede mover al procesamiento en el dominio del tiempo. Esto permite que el banco de filtros 140 se implemente con la longitud de palabra original o la longitud de palabra reducida en comparación con una implementación donde la normalización de sonoridad se realiza dentro del procesamiento en el dominio de la frecuencia. Para realizar una mezcla suave de valores de ganancia, se puede realizar un ajuste de forma de transición como se explicará más adelante en el contexto de la Figura 9.

15 Además, las muestras de audio dentro del flujo de bits se cuantifican por lo general a menor precisión que la señal de audio reconstruida. Esto permite algún espacio en el banco de filtros 140. El decodificador 100 deriva alguna estimación del otro parámetro de flujo de bits  $p$  (tal como el factor de ganancia global) y, para el caso donde es probable el recorte de la señal de salida, se aplica un cambio de nivel ( $g_2$ ) para evitar el recorte en el banco de filtros 20 140. Este cambio de nivel se señala al dominio del tiempo para compensación apropiada mediante el compensador de cambio de nivel 150. Si no se estima el recorte, la señal de audio permanece sin cambios y por lo tanto el método no tiene ninguna pérdida de precisión.

El estimador de recorte se puede configurar además para determinar una probabilidad de recorte basándose en la 25 información adicional y/o para determinar el factor de cambio de nivel actual basándose en la probabilidad de recorte. Aunque la probabilidad de recorte sólo indica una tendencia, en lugar de un hecho permanente, puede proporcionar información útil relacionada con el factor de cambio de nivel que se puede aplicar de forma razonable a la pluralidad de señales de banda de frecuencia para una trama dada de la representación de señal de audio codificada. La determinación de la probabilidad de recorte puede ser relativamente sencilla en términos de 30 complejidad o esfuerzo computacional y en comparación con la conversión del dominio de la frecuencia al dominio del tiempo realizada por el convertidor del dominio de la frecuencia al dominio del tiempo 140.

La información adicional puede comprender al menos uno de un factor de ganancia global para la pluralidad de 35 señales de banda de frecuencia y una pluralidad de factores de escala. Cada factor de escala puede corresponder a una o más señales de banda de frecuencia de la pluralidad de señales de banda de frecuencia. El factor de ganancia global y/o la pluralidad de factores de escala proporcionan ya información útil relacionada con un nivel de sonoridad de la trama actual que se va a convertir al dominio del tiempo mediante el convertidor 140.

De acuerdo con al menos algunas realizaciones, la etapa de pre-procesamiento de decodificador 110 se puede 40 configurar para obtener la pluralidad de señales de banda de frecuencia en la forma de una pluralidad de tramas sucesivas. El estimador de recorte 120 se puede configurar para determinar el factor de cambio de nivel actual para una trama actual. En otras palabras, el decodificador de señales de audio 100 se puede configurar para determinar de forma dinámica los factores de cambio de nivel variables para diferentes tramas de la representación de la señal de audio codificada, por ejemplo dependiendo de un grado variable de la sonoridad dentro de las tramas sucesivas.

45 La representación de señal de audio decodificada se puede determinar basándose en la representación en el dominio del tiempo sustancialmente compensada. Por ejemplo, el decodificador de señales de audio 100 puede comprender además un limitador en el dominio del tiempo corriente abajo del compensador de cambio de nivel 150. De acuerdo con algunas realizaciones, el compensador de cambio de nivel 150 puede ser una parte de un limitador 50 en el dominio del tiempo de este tipo.

De acuerdo con realizaciones adicionales, la información adicional con respecto a la ganancia de las señales de 55 banda de frecuencia puede comprender una pluralidad de factores de ganancia relacionados con la banda de frecuencia.

La etapa de pre-procesamiento de decodificador 110 puede comprender un cuantificador inverso configurado para 60 pre-cuantificar cada señal de banda de frecuencia utilizando un indicador de cuantificación específico de la banda de frecuencia de una pluralidad de indicadores de cuantificación específicos de la banda de frecuencia. En particular, pueden haberse cuantificado diferentes señales de banda de frecuencia utilizando diferentes resoluciones de cuantificación (o resoluciones de bits) mediante un codificador de señales de audio que ha creado la presentación de señal de audio codificada y la información adicional correspondiente. Los diferentes indicadores de cuantificación específicos de la banda de frecuencia pueden por lo tanto proporcionar una información acerca de una resolución de amplitud para las diversas señales de banda de frecuencia, dependiendo de una resolución de amplitud requerida para esa señal de banda de frecuencia particular determinada anteriormente por el codificador de señales de audio.

La pluralidad de indicadores de cuantificación específicos de la banda de frecuencia puede ser parte de la información adicional proporcionada a la etapa de pre-procesamiento de decodificador 110 y pueden proporcionar información adicional que se va a utilizar mediante el estimador de recorte 120 para determinar el factor de cambio de nivel.

5 El estimador de recorte 120 se puede configurar además para analizar la información adicional con respecto a si la información adicional sugiere un recorte potencial dentro de la representación en el dominio del tiempo. Un hallazgo de este tipo entonces se interpretaría como un bit menos significativo (LSB) que no contiene información relevante. En este caso el cambio de nivel aplicado por el cambiador de nivel 130 puede cambiar la información hacia el bit menos significativo de tal forma que al liberar un bit más significativo (MSB) se gana algún espacio en el bit más significativo, que puede ser necesario para la resolución en el dominio del tiempo en caso de que dos o más de las señales de banda de frecuencia se sumen de una forma constructiva. Este concepto también se puede extender a los  $n$  bits menos significativos y a los  $n$  bits más significativos.

15 El estimador de recorte 120 se puede configurar para considerar un ruido de cuantificación. Por ejemplo, en decodificación AAC, tanto la "ganancia global" como las "bandas de factor de escala" se utilizan para normalizar el audio/sub-banda. Como consecuencia, la información relevante de cada valor (espectral) se cambia a los MSB, mientras que los LSB se desprecian en la cuantificación. Después de la re-cuantificación en el decodificador, típicamente solamente los LSB contienen ruido. Si la "ganancia global" y los valores de "banda de factor de escala" (p) sugieren un recorte potencial después del banco de filtros de reconstrucción 140, se puede suponer de forma razonable que los LSB no contienen información. Con el método propuesto, el decodificador 100 cambia la información también en estos bits para ganar algún espacio con los MSB. Esto provoca sustancialmente que no haya pérdida de información.

25 El aparato propuesto (decodificador o codificador de señales de audio) y los métodos permiten la prevención de recortes para decodificadores/codificadores de audio sin emplear un banco de filtro de alta resolución para el espacio requerido. Esto es típicamente mucho menos costoso en términos de requisitos de memoria y complejidad computacional que realizar/implementar un banco de filtros con resolución más alta.

30 La Figura 6 muestra un diagrama de bloques esquemático de un decodificador de señales de audio 100 de acuerdo con realizaciones adicionales de la presente invención. El decodificador de señales de audio 100 comprende un cuantificador inverso 210 ( $Q^{-1}$ ) que se configura para recibir la representación de señal de audio codificada y típicamente también la información adicional o una parte de la información adicional. En algunas realizaciones, el cuantificador inverso 210 puede comprender un desempaquetador de flujo de bits configurado para desempaquetar un flujo de bits que contiene la representación de señal de audio codificada y la información adicional, por ejemplo en la forma de paquetes de datos, en el que cada paquete de datos puede corresponder a un cierto número de tramas de la representación de señal de audio codificada. Como se explicó anteriormente, dentro de la representación de señal de audio codificada y dentro de cada trama, cada banda de frecuencia puede tener su propia resolución de cuantificación individual. De esta forma, las bandas de frecuencia que requieren de forma temporal una cuantificación relativamente precisa, para representar de forma correcta las porciones de la señal de audio dentro de las bandas de frecuencia, pueden tener una resolución de cuantificación precisa de este tipo. Por otra parte, las bandas de frecuencia que contienen, durante una trama dada, nada o solamente una pequeña cantidad de información se puede cuantificar utilizando una cuantificación mucho más basta, ahorrando por lo tanto bits de datos. El cuantificador inverso 210 se puede configurar para proporcionar las diferentes bandas de frecuencia, que se han cuantificado utilizando resoluciones de cuantificación variables en el tiempo e individuales, a una resolución de cuantificación común. La resolución de cuantificación común puede ser, por ejemplo, la resolución proporcionada por una representación aritmética de punto fijo que se utiliza por el decodificador de señales de audio 100 de forma interna para los cálculos y el procesamiento. Por ejemplo, el decodificador de señales de audio 100 puede utilizar de forma interna una representación de punto fijo de 16 bits o de 24 bits. La información adicional proporcionada al cuantificador inverso 210 puede contener información relacionada con diferentes resoluciones de cuantificación para la pluralidad de señales de banda de frecuencia para cada nueva trama. El cuantificador inverso 210 se puede considerar como un caso especial de la etapa de pre-procesamiento de decodificador 110 representada en la Figura 5.

55 El estimador de recorte 120 mostrado en la Figura 6 es similar al estimador de recorte 120 en la Figura 5.

El decodificador de señales de audio 100 comprende además el cambiador de nivel 230 que se conecta a una salida del cuantificador inverso 210. El cambiador de nivel 230 recibe además información adicional o una parte de la información adicional, así como el factor de cambio de nivel que se determina mediante el estimador de recorte 120 de una forma dinámica, es decir, para cada intervalo de tiempo o trama, el factor de cambio de nivel puede suponer un valor diferente. El factor de cambio de nivel se aplica de forma consistente a la pluralidad de señales de banda de frecuencia utilizando una pluralidad de multiplicadores o elementos de modificación de escala 231, 232 y 233. Puede ocurrir que algunas de las señales de banda de frecuencia sean relativamente fuertes cuando dejan el cuantificador inverso 210, posiblemente utilizando ya sus MSB respectivos. Cuando estas señales de banda de frecuencia fuertes

se suman dentro del convertidor del dominio de la frecuencia al dominio del tiempo 140, se puede observar un desbordamiento dentro de la representación en el dominio del tiempo emitida por el convertidor del dominio de la frecuencia del dominio del tiempo 140. El factor de cambio de nivel determinado por el estimador de recorte 120 y aplicado por los elementos de modificación de escala 231, 232, 233 hace posible reducir de forma selectiva (es decir, teniendo en cuenta la información adicional actual) los niveles de las señales de banda de frecuencia de tal forma que es menos probable que tenga lugar un desbordamiento de la representación en el dominio del tiempo. El cambiador de nivel 230 comprende además una segunda pluralidad de multiplicadores o elementos de modificación de escala 236, 237, 238 configurados para aplicar los factores de escala específicos de la banda de frecuencia a las bandas de frecuencia correspondientes. La información adicional puede comprender  $M$  factores de escala. El cambiador de nivel 230 proporciona la pluralidad de señales de banda de frecuencia cambiadas de nivel al convertidor del dominio de la frecuencia al dominio del tiempo 140 que se configura para convertir las señales de banda de frecuencia cambiadas de nivel a la representación en el dominio del tiempo.

El decodificador de señales de audio 100 de la Figura 6 comprende además el compensador de cambio de nivel 150 que comprende en la realización representada un multiplicador o elemento de modificación de escala adicional 250 y un calculador recíproco 252. El calculador recíproco 252 recibe el factor de cambio de nivel y determina la recíproca ( $1/x$ ) del factor de cambio de nivel. La recíproca del factor de cambio de nivel se reenvía al elemento de modificación de escala adicional 250 donde se multiplica por la representación en el dominio del tiempo para producir la representación en el dominio del tiempo sustancialmente compensada. Como alternativa a los multiplicadores o elementos de modificación de escala 231, 232, 233 y 252 también puede ser posible utilizar elementos de suma/resta para aplicar el factor de cambio de nivel a la pluralidad de señales de banda de frecuencia y a la representación en el dominio del tiempo.

De forma opcional, el decodificador de señales de audio 100 en la Figura 6 comprende además un elemento de procesamiento posterior 260 conectado a una salida del compensador de cambio de nivel 150. Por ejemplo, el elemento de procesamiento posterior 260 puede comprender un limitador en el dominio del tiempo que tiene una característica fija para reducir o eliminar cualquier recorte que pueda aún estar presente dentro de la representación en el dominio del tiempo sustancialmente compensada, a pesar de la provisión del cambiador de nivel 230 y el compensador de cambio de nivel 150. Una salida del elemento de procesamiento posterior opcional 260 proporciona la representación de señal de audio decodificada. En caso de que no esté presente el elemento de procesamiento posterior opcional 260, la representación de señal de audio decodificada puede estar disponible en la salida del compensador de cambio de nivel 150.

La Figura 7 muestra un diagrama de bloques esquemático de un decodificador de señales de audio 100 de acuerdo con realizaciones adicionales posibles de la presente invención. Un decodificador de flujo de bits/cuantificador inverso 310 se configura para procesar un flujo de bits entrante y para derivar la siguiente información a partir del mismo: la pluralidad de señales de banda de frecuencia  $X_1(f)$ , los parámetros de flujo de bits  $\mathbf{p}$ , y una ganancia global  $g_1$ . Los parámetros de flujo de bits  $\mathbf{p}$  pueden comprender los factores de escala para las bandas de frecuencia y/o la ganancia global  $g_1$ .

Los parámetros de flujo de bits  $\mathbf{p}$  se proporcionan al estimador de recorte 320 que deriva el factor de escala  $1/g_2$  a partir de los parámetros de flujo de bits  $\mathbf{p}$ . El factor de escala  $1/g_2$  se alimenta al cambiador de nivel 330 que en la realización representada también implementa un control de intervalo dinámico (DRC). El cambiador de nivel 330 puede recibir además los parámetros de flujo de bits  $\mathbf{p}$  o una porción de los mismos para aplicar los factores de escala a la pluralidad de señales de banda de frecuencia. El cambiador de nivel 330 emite la pluralidad de señales de banda de frecuencia cambiadas de nivel  $X_2(f)$  al banco de filtros inversos 340 que proporciona la conversión del dominio de la frecuencia al dominio del tiempo. En una salida del banco de filtros inversos 340, se proporciona la representación en el dominio del tiempo  $X_3(t)$  que se va a suministrar al compensador de cambio de nivel 350. El compensador de cambio de nivel 350 es un multiplicador o elemento de modificación de escala, como en la realización representada en la Figura 6. El compensador de cambio de nivel 350 es una parte de un procesamiento en el dominio del tiempo posterior 360 para procesamiento de alta precisión, por ejemplo, que soporta una longitud de palabra más larga que el banco de filtros inversos 340. Por ejemplo, el banco de filtros inversos puede tener una longitud de palabra de 16 bits y el procesamiento de alta precisión realizado por el procesamiento en dominio del tiempo posterior se puede realizar utilizando 20 bits. Como otro ejemplo, la longitud de palabra del banco de filtros inversos 340 puede ser de 24 bits y la longitud de palabra del procesamiento de alta precisión puede ser de 30 bits. En cualquier caso, el número de bits no deberá considerarse como que limita el alcance de la presente patente/solicitud de patente a menos que se indique explícitamente. El procesamiento en el dominio del tiempo posterior 360 emite la representación de señal de audio decodificada  $X_4(t)$ .

El cambio de ganancia aplicado  $g_2$  se alimenta hacia la implementación del limitador 360 para compensación. El limitador 362 se puede implementar con alta precisión.

Si el estimador de recorte 320 no estima ningún recorte, las muestras de audio permanecen sustancialmente sin cambios, es decir como si no se hubiera realizado el cambio de nivel y la compensación de cambio de nivel.

El estimador de recorte proporciona la recíproca  $g_2$  del factor de cambio de nivel  $1/g_2$  a un combinador 328 donde se combina con la ganancia global  $g_1$  para producir una ganancia combinada  $g_3$ .

5 El decodificador de señales de audio 100 comprende además un ajuste de forma de transición 370 que se configura para proporcionar transiciones suaves cuando la ganancia combinada  $g_3$  cambia de forma abrupta de una trama precedente a una trama actual (o de la trama actual a una trama posterior). El ajustador de forma de transición 370 se puede configurar para encadenar el factor de cambio de nivel actual y un factor de cambio de nivel posterior para obtener un factor de cambio de nivel encadenado  $g_4$  para uso por el compensador cambio de nivel 350. Para permitir la transición suave de los factores de ganancia que cambian, se tiene que realizar un ajuste de forma de transición. Esta herramienta crea un vector de factores de ganancia  $g_4(t)$  (un factor para cada muestra de la señal de audio correspondiente). Para simular el mismo comportamiento del ajuste de ganancia que produciría el procesamiento de la señal en el dominio de la frecuencia, se tienen que utilizar las mismas ventanas de transición  $W$  del banco de filtros 340. Una trama cubre una pluralidad de muestras. El factor de ganancia combinado  $g_3$  es típicamente constante para la duración de una trama. La ventana de transición  $W$  es típicamente una trama larga y proporciona valores de ventana diferentes para cada muestra dentro de la trama (por ejemplo, el primer medio periodo de un coseno). Los detalles relacionados con una implementación posible del ajuste de forma de transición se proporcionan en la Figura 9 y en la descripción correspondiente más adelante.

20 La Figura 8 ilustra esquemáticamente el efecto de un cambio de nivel aplicado a la pluralidad de señales de banda de frecuencia. Una señal de audio (por ejemplo, cada una de la pluralidad de señales de banda de frecuencia) se puede representar utilizando una resolución de 16 bits, como se simboliza por el rectángulo 402. El rectángulo 404 ilustra esquemáticamente cómo se emplean los bits de la resolución de 16 bits para representar la muestra cuantificada dentro de una de las señales de banda de frecuencia proporcionadas por la etapa de pre-procesamiento de decodificador 110. Se puede ver que la muestra cuantificada puede utilizar un cierto número de bits empezando del bit más significativo (MSB) hasta un último bit utilizado para la muestra cuantificada. Los bits restantes hasta el bit menos significativo (LSB) contienen solamente ruido de cuantificación. Esto se puede explicar por el hecho de que para la trama actual la señal de banda de frecuencia correspondiente se representó dentro del flujo de bits solamente mediante un número reducido de bits (< 16 bits). Incluso si se utilizó la resolución de bits completa de 16 bits dentro del flujo de bits para la trama actual y para la banda de frecuencia correspondiente, el bit menos significativo contiene típicamente una cantidad significativa de ruido de cuantificación.

Un rectángulo 406 en la Figura 8 ilustra esquemáticamente el resultado del cambio de nivel de la señal de banda de frecuencia. Puesto que se puede esperar que el contenido del bit o bits menos significativos contenga una cantidad considerable de ruido de cuantificación, la muestra cuantificada se puede cambiar hacia el bit menos significativo, sustancialmente sin perder información relevante. Esto se puede lograr al cambiar simplemente los bits hacia abajo (“desplazamiento a la derecha”), o recalculando de forma real la representación binaria. En ambos casos, el factor de cambio de nivel se puede memorizar para compensación posterior del cambio de nivel aplicado (por ejemplo, por medio del compensador de cambio de nivel 150 o 350). El cambio de nivel da como resultado espacio adicional en el bit o bits más significativos.

La Figura 9 ilustra esquemáticamente una implementación posible del ajuste de forma de transición 370 mostrado en la Figura 7. El ajustador de forma de transición 370 puede comprender una memoria 371 para un factor de cambio de nivel anterior, un primer sistema de formación de ventanas 372 configurado para generar una primera pluralidad de muestras en ventanas al aplicar una forma de ventana al factor de cambio de nivel actual, un segundo sistema de formación de ventanas 376 configurado para generar una segunda pluralidad de muestras en ventanas al aplicar una forma de ventana anterior al factor de cambio de nivel anterior proporcionado por la memoria 371, y un combinador de muestras 379 configurado para combinar mutuamente muestras en ventanas correspondientes de la primera pluralidad de muestras en ventanas y de la segunda pluralidad de muestras en ventanas para obtener una pluralidad de muestras combinadas. El primer sistema de formación de ventanas 372 comprende un proveedor de forma de ventana 373 y un multiplicador 374. El segundo sistema de formación de ventanas 376 comprende un proveedor de forma de ventana anterior 377 y un multiplicador adicional 378. El multiplicador 374 y el multiplicador adicional 378 emiten vectores durante el tiempo. En el caso del primer sistema de formación de ventanas 372 cada elemento de vector corresponde a la multiplicación del factor de ganancia combinado actual  $g_3(t)$  (constante durante la trama actual) con la forma de ventana actual proporcionada por el proveedor de forma de ventana 373. En el caso del segundo sistema de formación de ventanas 376 cada elemento de vector corresponde a la multiplicación del factor de ganancia combinado anterior  $g_3(t-T)$  (constante durante la trama anterior) con la forma de ventana anterior proporcionada por el proveedor de forma de ventana anterior 377.

60 De acuerdo con la realización ilustrada esquemáticamente en la Figura 9, el factor de ganancia de la trama anterior se tiene que multiplicar por la ventana de la “segunda mitad” del banco de filtros 340, mientras que el factor de ganancia real se multiplica por la secuencia de ventana de la “primera mitad”. Estos dos vectores se pueden sumar para formar un vector de ganancia  $g_4(t)$  que se va a multiplicar por elemento con la señal de audio  $X_3(t)$  (véase la Figura 7).

Las formas de ventana se pueden guiar por la información adicional  $w$  del banco de filtros 340, si se requiere.

5 La forma de ventana y la forma de ventana anterior también se pueden utilizar mediante el convertidor del dominio de la frecuencia al dominio del tiempo 340 de tal forma que la misma forma de ventana y la forma de ventana anterior se utilizan para convertir las señales de banda de frecuencia cambiadas de nivel a la representación en el dominio del tiempo y para la formación de ventanas del factor de cambio de nivel actual y del factor de cambio de nivel anterior.

10 El factor de cambio de nivel actual puede ser válido para una trama actual de la pluralidad de señales de banda de frecuencia. El factor de cambio de nivel anterior puede ser válido para una trama anterior de la pluralidad de señales de banda de frecuencia. La trama actual y la trama anterior se pueden superponer, por ejemplo por 50 %.

15 El ajuste de forma de transición 370 se puede configurar para combinar el factor de nivel anterior con una segunda porción de la forma de ventana anterior que da como resultado una secuencia de factor de trama anterior. El ajuste de forma de transición 370 se puede configurar además para combinar el factor de cambio de nivel actual con una primera porción de la forma de ventana actual que da como resultado una secuencia de factor de trama actual. Se puede determinar una secuencia del factor de cambio de nivel encadenado basándose en la secuencia de factor de trama anterior y la secuencia de factor de trama actual.

20 El enfoque propuesto no se limita necesariamente a decodificadores, sino que también los codificadores pueden tener un ajuste de ganancia o limitador en combinación con un banco de filtros que puede beneficiarse del método propuesto.

25 La Figura 10 ilustra cómo se conectan la etapa de pre-procesamiento de decodificador 110 y el estimador de recorte 120. La etapa de pre-procesamiento del decodificador 110 corresponde a o comprende el determinador del libro de códigos 1110. El estimador de recorte 120 comprende una unidad de estimación 1120. El determinador de libro de códigos 1110 se adapta para determinar un libro de códigos a partir de una pluralidad de libros de códigos como un libro de códigos identificado, en el que la señal de audio se ha codificado empleando el libro de códigos identificado. La unidad de estimación 1120 se adapta para derivar un valor de nivel, por ejemplo un valor de energía, un valor de amplitud o un valor de sonoridad, asociado con el libro de códigos identificado como un valor de nivel derivado. Además, la unidad de estimación 1120 se adapta para estimar una estimación de nivel, por ejemplo una estimación de energía, una estimación de amplitud o una estimación de sonoridad, de la señal de audio utilizando el valor del nivel derivado. Por ejemplo, el determinador del libro de códigos 1110 puede determinar el libro de códigos, que se ha utilizado por un codificador para codificar la señal de audio, recibiendo información adicional transmitida junto con la señal de audio codificada. En particular, la información adicional puede comprender información que identifica el libro de códigos utilizado para codificar una sección considerada de la señal de audio. Tal información puede transmitirse, por ejemplo, desde el codificador al decodificador como un número, que identifica un libro de códigos Huffman utilizado para codificar la sección considerada de la señal de audio.

40 La Figura 11 ilustra una unidad de estimación de acuerdo con una realización. La unidad de estimación comprende un derivador de valor de nivel 1210 y una unidad de modificación de escala 1220. El derivador de valor de nivel se adapta para derivar un valor de nivel asociado con el libro de códigos identificado, es decir, el libro de códigos que se utilizó para codificar los datos espectrales por el codificador, buscando el valor de nivel en una memoria, solicitando el valor de nivel desde una base de datos local o solicitando el valor de nivel asociado con el libro de códigos identificado desde una computadora remota. En una realización, el valor de nivel, que se busca o solicita mediante el derivador de valor de nivel, puede ser un valor de nivel promedio que indica un nivel promedio de un valor espectral no modificado a escala codificado al utilizar el libro de códigos identificado.

50 Mediante esto, el valor de nivel derivado no se calcula a partir de los valores espectrales reales sino en su lugar, se utiliza un valor de nivel promedio que depende solamente del libro de códigos empleado. Como se ha explicado anteriormente, el codificador se adapta en general para seleccionar el libro de códigos a partir de una pluralidad de libros de códigos que se adaptan mejor para codificar los datos espectrales respectivos de una sección de la señal de audio. Puesto que los libros de códigos difieren, por ejemplo con respecto a sus valores máximos absolutos que se pueden codificar, el valor promedio que se codifica por un libro de códigos Huffman difiere del libro de códigos a libro de códigos y, por lo tanto, también el nivel de valor promedio de un coeficiente espectral codificado mediante un libro de códigos particular difiere del libro de códigos al libro de códigos.

60 Por lo tanto, de acuerdo con una realización, un valor de nivel promedio para codificar un coeficiente espectral de una señal de audio que emplea un libro de códigos Huffman particular se puede determinar para cada libro de códigos Huffman y, por ejemplo, se puede almacenar en una memoria, una base de datos o en una computadora remota. El derivador de valor de nivel entonces tiene que buscar o solicitar simplemente el valor de nivel asociado con el libro de códigos identificado que se ha empleado para codificar los datos espectrales, para obtener el valor de nivel derivado asociado con el libro de códigos identificado.

Sin embargo, se tiene que tener en cuenta que los libros de códigos Huffman a menudo se emplean para codificar valores espectrales no modificados a escala, como es el caso para la MPEG AAC. Entonces, sin embargo, se debe tener en cuenta la modificación a escala cuando se lleva a cabo una estimación de nivel. Por lo tanto, la unidad de estimación de la Figura 11 también comprende una unidad de modificación de escala 1220. La unidad de modificación de escala se adapta para derivar un factor de escala relacionado con la señal de audio codificada o a una porción de la señal de audio codificada como un factor de escala derivado. Por ejemplo, con respecto a un decodificador, la unidad de modificación de escala 1220 determinará un factor de escala para cada banda de factor de escala. Por ejemplo, la unidad de modificación de escala 1220 puede recibir información acerca del factor de escala de una banda de factor de escala recibiendo información adicional transmitida desde un codificador al decodificador. La unidad de modificación de escala 1220 se adapta además para determinar un valor de nivel modificado a escala basándose en el factor de escala y el valor de nivel derivado.

En una realización, donde el valor de nivel derivado es un valor de energía derivado, la unidad de modificación de escala se adapta para aplicar el factor de escala derivado al valor de energía derivado para obtener un valor de nivel modificado a escala multiplicando el valor de energía derivado por el cuadrado del factor de escala derivado.

En otra realización, donde el valor de nivel derivado es un valor de amplitud derivado, y la unidad de modificación de escala se adapta para aplicar el factor de escala derivado al valor de amplitud derivado para obtener un valor de nivel modificado a escala multiplicando el valor de amplitud derivado por el factor de escala derivado.

En una realización adicional, en la que el valor de nivel derivado es un valor de sonoridad derivado, y la unidad de modificación de escala 1220 se adapta para aplicar el factor de escala derivado al valor de sonoridad derivado para obtener un valor de nivel modificado a escala multiplicando el valor de sonoridad derivado por el cubo del factor de escala derivado. Existen formas alternativas para calcular la sonoridad tal como mediante un exponente 3/2. En general, los factores de escala se tienen que transformar al dominio de sonoridad, cuando el valor de nivel derivado es un valor de sonoridad.

Estas realizaciones tienen en cuenta, que se determina un valor de energía basándose en el cuadrado de los coeficientes espectrales de una señal de audio, que se determina un valor de amplitud basándose en los valores absolutos de los coeficientes espectrales de una señal de audio, y que se determina un valor de sonoridad basándose en los coeficientes espectrales de una señal de audio que se ha transformado al dominio de sonoridad.

La unidad de estimación se adapta para estimar una estimación de nivel de la señal de audio utilizando el valor de nivel modificado a escala. En la realización de la Figura 11, la unidad de estimación se adapta para emitir el valor de nivel modificado a escala como la estimación de nivel. En este caso, no se lleva a cabo pos-procesamiento del valor de nivel modificado a escala. Sin embargo, como se ilustra en la realización de la Figura 12, la unidad de estimación también se puede adaptar para llevar a cabo un pos-procesamiento. Por lo tanto, la unidad de estimación de la Figura 12 comprende un pos-procesador 1230 para pos-procesar uno o más valores de nivel modificados a escala para estimar una estimación de nivel. Por ejemplo, la estimación de nivel de la unidad de estimación se puede determinar mediante el pos-procesador 1230 determinando un valor promedio de una pluralidad de valores de nivel modificados a escala. Este valor promediado se puede emitir por la unidad de estimación como la estimación de nivel.

En comparación con las realizaciones presentadas, un enfoque del estado de la técnica para estimar, por ejemplo, la energía de una banda de factor de escala sería realizar la decodificación Huffman y la cuantificación inversa para todos los valores espectrales y calcular la energía sumando el cuadrado de todos los valores espectrales cuantificados de forma inversa.

En las realizaciones propuestas, sin embargo, este proceso computacionalmente complejo del estado de la técnica se reemplaza por una estimación del nivel promedio que solamente depende del factor de escala y los usos del libro de códigos y no de los valores cuantificados reales.

Las realizaciones de la presente invención emplean el hecho de que un libro de códigos Huffman se diseña para proporcionar codificación óptima después de una estadística especializada. Esto significa que el libro de códigos se ha diseñado de acuerdo con la probabilidad de los datos, por ejemplo, AAC-ELD (AAC-ELD = codificación de audio avanzada - bajo retardo mejorado): líneas espectrales. Este proceso se puede invertir para obtener la probabilidad de los datos de acuerdo con el libro de códigos. La probabilidad de cada entrada de datos dentro de un libro de códigos (índice) está dada por la longitud de palabra de código. Por ejemplo,

$$p(\text{índice}) = 2^{-\text{longitud}(\text{palabra de código})}$$

es decir

$$p(\text{índice}) = 2^{-\text{longitud}(\text{palabra de código})}$$

en el que  $p$ (índice) es la probabilidad de una entrada de datos (un índice) dentro de un libro de códigos.

5 Basándose en esto, el nivel esperado se puede pre-calcular y almacenar de la siguiente manera: cada índice representa una secuencia de valores enteros ( $x$ ), por ejemplo, líneas espectrales, donde la longitud de la secuencia depende de la dimensión del libro de códigos, por ejemplo, 2 o 4 para AAC-ELD.

10 La Figura 13a y 13b ilustran un método para generar un valor de nivel, por ejemplo un valor de energía, un valor de amplitud o un valor de sonoridad, asociado con un libro de códigos de acuerdo con una realización. El método comprende:

15 Determinar una secuencia de valores numéricos asociados con una palabra de código del libro de códigos para cada palabra de código del libro de códigos (etapa 1310). Como se ha explicado anteriormente, un libro de códigos codifica una secuencia de valores numéricos, por ejemplo, 2 o 4 valores numéricos por una palabra de código del libro de códigos. El libro de códigos comprende una pluralidad de libros de códigos para codificar una pluralidad de secuencias de valores numéricos. La secuencia de valores numéricos, que se determina, es la secuencia de valores numéricos que se codifica por la palabra de código considerada del libro de códigos. La etapa 1310 se lleva a cabo para cada palabra de código del libro de códigos. Por ejemplo, si el libro de códigos comprende 81 palabras de código, se determinan 81 secuencias de valores numéricos en la etapa 1310.

20 En la etapa 1320, se determina una secuencia cuantificada de forma inversa de valores numéricos para cada palabra de código del libro de códigos aplicando un cuantificador inverso a los valores numéricos de la secuencia de valores numéricos de una palabra de código para cada palabra de código del libro de códigos. Como se ha explicado anteriormente, un codificador puede emplear en general cuantificación cuando se codifican los valores espectrales de la señal de audio, por ejemplo cuantificación no uniforme. Como consecuencia, esta cuantificación se tiene que invertir en el lado del decodificador.

25 Después, en la etapa 1330, se determina una secuencia de valores de nivel para cada palabra de código del libro de códigos.

30 Si se va a generar un valor de energía como el valor de nivel del libro de códigos, entonces se determina una secuencia de valores de energía para cada palabra de código, y se calcula el cuadrado de cada valor de la secuencia cuantificada de forma inversa de los valores numéricos para cada palabra de código del libro de códigos.

35 Si, sin embargo, se va a generar un valor de amplitud como el valor del nivel de libros de códigos, entonces se determina una secuencia de valores de amplitud para cada palabra de código, y se calcula el valor absoluto de cada valor de la secuencia cuantificada de forma inversa de valores numéricos para cada palabra de código del libro de códigos.

40 Si, aunque, se va a generar un valor de sonoridad como el valor de nivel del libro de códigos, entonces se determina una secuencia de valores de sonoridad para cada palabra de código, y se calcula el cubo de cada valor de la secuencia cuantificada de forma inversa de valores numéricos para cada palabra de código del libro de códigos. Existen formas alternativas para calcular la sonoridad tal como mediante un exponente  $3/2$ . En general, los valores de la secuencia cuantificada de forma inversa de valores numéricos se tienen que transformar al dominio de sonoridad, cuando se va a generar un valor de sonoridad como el valor del nivel del libro de códigos.

Posteriormente, en la etapa 1340, se calcula un valor de suma de nivel para cada palabra de código del libro de códigos sumando los valores de la secuencia de valores de nivel para cada palabra de código del libro de códigos.

50 Entonces, en la etapa 1350, se determina un valor de suma de nivel de probabilidad ponderada para cada palabra de código del libro de códigos multiplicando el valor de suma de nivel de una palabra de código por un valor de probabilidad asociado con la palabra de código para cada palabra de código del libro de códigos. Mediante esto, se tiene en cuenta que algunas de las secuencias de valores numéricos, por ejemplo, secuencias de coeficientes espectrales, no aparecerán tan a menudo como otras secuencias de coeficientes espectrales. El valor de probabilidad asociado con la palabra de código tiene esto en cuenta. Un valor de probabilidad de este tipo se puede derivar a partir de la longitud de palabra de código, ya que las palabras de códigos que es más probable que aparezcan se codifican utilizando palabras de códigos que tienen una longitud más corta, mientras que otras palabras de código que es menos probable que aparezcan se codificarán utilizando palabras de código que tienen una longitud más larga, cuando se emplea codificación Huffman.

60 En la etapa 1360, se determinará un valor de suma de nivel de probabilidad ponderada promediado para cada palabra de código del libro de códigos dividiendo el valor de suma de nivel de probabilidad ponderada de una palabra de código por un valor de dimensión asociado con el libro de códigos para cada palabra de código del libro de códigos. Un valor de dimensión indica el número de valores espectrales que se codifican por una palabra de



código del libro de códigos. Mediante esto, se determina un valor de suma de nivel de probabilidad ponderada, promediado que representa un valor de nivel (probabilidad ponderada) para un coeficiente espectral que se codifica por la palabra de código.

- 5 Entonces, en la etapa 1370, el valor de nivel del libro de códigos se calcula sumando los valores de suma del nivel de probabilidad ponderada promediados de todas las palabras de código.

Se tiene que señalar, que una generación de este tipo de un valor de nivel se tiene que realizar solo una vez para un libro de códigos. Si se determina el valor de nivel de un libro de códigos, este valor se puede buscar y utilizar simplemente, por ejemplo por un aparato para estimación de nivel de acuerdo con las realizaciones descritas anteriormente.

10 A continuación, se presenta un método para generar un valor de energía asociado con un libro de códigos de acuerdo con una realización. Para estimar el valor esperado de la energía de los datos codificados con el libro de códigos dado, se tienen que realizar las siguientes etapas solamente una vez para cada índice del libro de códigos:

A) aplicar el cuantificador inverso a los valores enteros de la secuencia (por ejemplo AAC-ELD:  $X^{(4/3)}$ ).

20 B) calcular la energía elevando al cuadrado cada valor de la secuencia de A)

C) crear la suma de la secuencia B)

D) multiplicar C) con la probabilidad dada del índice

25 E) dividir por la dimensión del libro de códigos para obtener la energía esperada por línea espectral.

Finalmente, se tienen que sumar todos los valores calculados por E) para obtener la energía esperada del libro de códigos completo.

30 Después de que la salida de estas etapas se almacena en una tabla, los valores de energía estimados se pueden buscar simplemente basándose en el índice del libro de códigos, es decir, dependiendo de en qué libro de códigos se utilice. Los valores espectrales reales no se tienen que someter a decodificación Huffman para esta estimación.

35 Para estimar la energía total de los datos espectrales de una trama de audio completa, se tiene que tener en cuenta el factor de escala. El factor de escala se puede extraer del flujo de bits sin una cantidad significativa de complejidad. El factor de escala se puede modificar antes de que se aplique a la energía esperada, por ejemplo se puede calcular el cuadrado del factor de escala utilizado. Entonces la energía esperada se multiplica por el cuadrado del factor de escala utilizado.

40 De acuerdo con las realizaciones descritas anteriormente, el valor espectral para cada banda de factor de escala se puede estimar sin decodificar los valores espectrales sometidos a codificación Huffman. Las estimaciones del nivel se pueden utilizar para identificar los flujos con un nivel bajo, por ejemplo con baja potencia, que son los que típicamente no dan como resultado recortes. Por lo tanto, la decodificación completa de estos tales se puede omitir.

45 De acuerdo con una realización, un aparato para estimación de nivel comprende además una memoria o una base de datos que tiene almacenada en la misma una pluralidad de valores de memoria del nivel del libro de códigos que indican un valor de nivel que está asociado con un libro de códigos, en el que cada uno de la pluralidad de libros de códigos tiene un valor de memoria del nivel del libro de códigos asociado con él almacenado en la memoria o la base de datos. Además, el derivador de valor de nivel se configura para derivar el valor de nivel asociado con el libro de códigos identificado derivando un valor de memoria de nivel del libro de códigos asociado con el libro de códigos identificado desde la memoria o desde la base de datos.

50 El nivel estimado de acuerdo con las realizaciones descritas anteriormente puede variar si una etapa de procesamiento adicional como predicción, tal como filtrado de predicción, se aplica en el códec, por ejemplo, para filtrado AAC-ELD TNS (conformación de ruido temporal). En este punto, los coeficientes de la predicción se transmiten dentro del flujo de bits, por ejemplo, para TNS como coeficientes PARCOR.

60 La Figura 14 ilustra una realización en la que la unidad de estimación comprende además un ajustador de filtro de predicción 1240. El ajustador de filtro de predicción se adapta para derivar uno o más coeficientes de filtro de predicción relacionados con la señal de audio codificada o con una porción de la señal de audio codificada como coeficientes de filtro de predicción derivados. Además, el ajustador de filtro de predicción se adapta para obtener un valor de nivel ajustado de filtro de predicción basándose en los coeficientes de filtro predicción y el valor de nivel derivado. Además, la unidad de estimación se adapta para estimar una estimación de nivel de la señal de audio utilizando el valor de nivel ajustado de filtro de predicción.

En una realización, los coeficientes PARCOR para TNS se utilizan como los coeficientes de filtro de predicción. La ganancia de predicción del proceso de filtrado se puede determinar a partir de aquellos coeficientes de una forma muy eficiente. Con respecto a TNS, la ganancia de predicción se puede calcular de acuerdo con la fórmula:

$$5 \quad \text{ganancia} = 1/\text{prod}(1-\text{parcor}_i^2).$$

Por ejemplo, si se tienen que tener en cuenta los 3 coeficientes PARCOR, por ejemplo,  $\text{parcor}_1$ ,  $\text{parcor}_2$  y  $\text{parcor}_3$ , la ganancia se calcula de acuerdo con la fórmula:

$$\text{ganancia} = \frac{1}{(1 - \text{parcor}_1^2)(1 - \text{parcor}_2^2)(1 - \text{parcor}_3^2)}$$

10 Para n coeficientes PARCOR  $\text{parcor}_1$ ,  $\text{parcor}_2$ , ...  $\text{parcor}_n$ , se aplica la siguiente fórmula:

$$\text{ganancia} = \frac{1}{(1 - \text{parcor}_1^2)(1 - \text{parcor}_2^2) \dots (1 - \text{parcor}_n^2)}$$

15 Esto significa que la amplificación de la señal de audio a través del filtrado se puede estimar sin aplicar la propia operación de filtrado.

La Figura 15 muestra un diagrama de bloques esquemático de un codificador 1500 que implementa el ajuste de ganancia propuesto que “evita” el banco de filtros. El codificador de señales de audio 1500 se configura para proporcionar una representación de señal de audio codificada basándose en una representación en el dominio del tiempo de una señal de audio de entrada. La representación en el dominio del tiempo puede ser, por ejemplo, una señal de entrada de audio modulada por codificación de impulsos.

El codificador de señales de audio comprende un estimador de recorte 1520 configurado para analizar la representación en el dominio del tiempo de la señal de audio de entrada para determinar un factor de cambio de nivel actual para la representación de señal de entrada. El codificador de señales de audio comprende además un cambiador de nivel 1530 configurado para cambiar un nivel de la representación en el dominio del tiempo de la señal de audio de entrada de acuerdo con el factor de cambio de nivel para obtener una representación en el dominio del tiempo cambiada de nivel. Un convertidor del dominio del tiempo al dominio de la frecuencia 1540 (por ejemplo, un banco de filtros, tal como un banco de filtros de espejo en cuadratura, una transformada discreta de coseno modificada, etc.) se configura para convertir la representación en el dominio del tiempo cambiada de nivel en una pluralidad de señales de banda de frecuencia. El codificador de señales de audio 1500 también comprende un compensador de cambio de nivel 1550 configurado para actuar sobre la pluralidad de señales de banda de frecuencia para compensar al menos parcialmente un cambio de nivel aplicado a la representación en el dominio del tiempo cambiada de nivel por el cambiador de nivel 1530 y para obtener una pluralidad de señales de banda de frecuencia sustancialmente compensadas.

El codificador de señales de audio 1500 puede comprender además un componente de asignación de bits/ruido, cuantificador y de codificación 1510 y un modelo psicoacústico 1508. El modelo psicoacústico 1508 determina los umbrales de enmascaramiento de tiempo-frecuencia variables en (y/o resoluciones de cuantificación de trama individual y de banda de frecuencia individual, y factores de escala) basándose en la señal de audio de entrada PCM, que se va a utilizar mediante la asignación de bits/ruido, cuantificador y la codificación 1610. Los detalles relacionados con una posible implementación del modelo psicoacústico y otros aspectos de la codificación de audio perceptual se pueden encontrar, por ejemplo, en las normas internacionales ISO/IEC 11172-3 e ISO/IEC 13818-3. La asignación de bits/ruido, cuantificador, y la codificación 1510 se configuran para cuantificar la pluralidad de señales de banda de frecuencia de acuerdo con sus resoluciones de cuantificación de trama individual y de banda de frecuencia individual, y para proporcionar estos datos a un formateador de flujo de bits 1505 que emite un flujo de bits codificado que se va a proporcionar a uno o más decodificadores de señales de audio. La asignación de bits/ruido, cuantificador y la codificación 1510 se pueden configurar para determinar la información adicional además de la pluralidad de señales de frecuencia cuantificadas. Esta información adicional también se puede proporcionar al formateador de flujo de bits 1505 para inclusión en el flujo de bits.

La Figura 16 muestra un diagrama de flujo esquemático de un método para decodificar una representación de señal de audio codificada para obtener una representación de señal de audio decodificada. El método comprende una etapa 1602 de pre-procesar la representación de señal de audio codificada para obtener una pluralidad de señales de banda de frecuencia. En particular, el pre-procesamiento puede comprender desempaquetar un flujo de bits en datos que corresponden a tramas sucesivas, y re-cuantificar (cuantificación inversa) los datos relacionados con la

banda de frecuencia de acuerdo con las resoluciones de cuantificación específicas de la banda de frecuencia para obtener una pluralidad de señales de banda de frecuencia.

5 En una etapa 1604 del método para decodificar, se analiza la información adicional con respecto a una ganancia de las señales de banda de frecuencia para determinar un factor de cambio de nivel actual para la representación de señal de audio codificada. La ganancia relacionada con las señales de banda de frecuencia puede ser individual para cada señal de banda de frecuencia (por ejemplo, los factores de escala conocidos en algunos esquemas de codificación de audio perceptual o parámetros similares) o común para todas las señales de banda de frecuencia (por ejemplo, la ganancia global conocida en algunos esquemas de codificación de audio perceptual). El análisis de la información adicional permite reunir información acerca de una sonoridad de la señal de audio codificada durante la trama disponible. La sonoridad, a su vez, puede indicar una tendencia de la representación de señal de audio decodificada a entrar en recorte. El factor de cambio de nivel se determina típicamente como un valor que evita tal recorte mientras conserva un intervalo dinámico relevante y/o contenido de información relevante de (todas) las señales de banda de frecuencia.

15 El método para decodificar comprende además una etapa 1606 para cambiar los niveles de la señal de banda de frecuencia de acuerdo con el factor de cambio de nivel. En caso de que las señales de banda de frecuencia se cambien de nivel a un nivel inferior, el cambio de nivel crea algún espacio adicional en el bit o bits más significativos de una representación binaria de las señales de banda de frecuencia. Este espacio adicional puede ser necesario cuando se convierte la pluralidad de señales de banda de frecuencia del dominio de la frecuencia al dominio del tiempo para obtener una representación en el dominio del tiempo, que se realiza en una etapa posterior 1608. En particular, el espacio adicional reduce el riesgo de que la representación en el dominio del tiempo se recorte si algunas de las señales de banda de frecuencia están cerca de un límite superior relacionado con su amplitud y/o potencia. Como consecuencia, la conversión del dominio de la frecuencia al dominio del tiempo se puede realizar utilizando una longitud de palabra relativamente corta.

20 El método para decodificar también comprende una etapa 1609 para actuar sobre la representación en el dominio del tiempo para compensar al menos parcialmente un cambio de nivel aplicado a las señales de banda de frecuencia cambiadas de nivel. Posteriormente, se obtiene una representación en el dominio del tiempo sustancialmente compensada.

30 En consecuencia, un método para decodificar una representación de señal de audio codificada a una representación de señal de audio decodificada comprende:

- 35
- pre-procesar la representación de señal de audio codificada para obtener una pluralidad de señales de banda de frecuencia;
  - analizar la información adicional con respecto a una ganancia de las señales de banda de frecuencia para determinar un factor de cambio de nivel actual para la representación de señal de audio codificada;
  - cambiar los niveles de las señales de banda de frecuencia de acuerdo con el factor de cambio de nivel para obtener las señales de banda de frecuencia cambiadas de nivel,
  - realizar una conversión del dominio de la frecuencia al dominio del tiempo de las señales de banda de frecuencia a una representación en el dominio del tiempo; y
  - actuar sobre la representación en el dominio del tiempo para compensar al menos parcialmente un cambio de nivel aplicado a las señales de banda de frecuencia cambiadas de nivel y para obtener una representación en el dominio del tiempo sustancialmente compensada.
- 40
- 45

50 De acuerdo con aspectos adicionales, analizar la información adicional puede comprender: determinar una probabilidad de recorte basándose en la información adicional y para determinar el factor de cambio de nivel actual basándose en la probabilidad de recorte.

55 De acuerdo con aspectos adicionales, la información adicional puede comprender al menos uno de un factor de ganancia global para la pluralidad de señales de banda de frecuencia y una pluralidad de factores de escala, correspondiendo cada factor de escala a una señal de banda de frecuencia de la pluralidad de señales de banda de frecuencia.

60 De acuerdo con aspectos adicionales, el pre-procesamiento de la representación de señal de audio codificada puede comprender obtener la pluralidad de señales de banda de frecuencia en la forma de una pluralidad de tramas sucesivas, y analizar la información adicional puede comprender determinar el factor de cambio de nivel actual para una trama actual.

De acuerdo con aspectos adicionales, la representación de señal de audio decodificada se puede determinar basándose en la representación en el dominio del tiempo sustancialmente compensada.

De acuerdo con aspectos adicionales, el método puede comprender además: aplicar una característica de limitador

en el dominio del tiempo posterior para actuar sobre la representación en el dominio del tiempo para compensar al menos parcialmente el cambio de nivel.

5 De acuerdo con aspectos adicionales, la información adicional con respecto a la ganancia de las señales de banda de frecuencia puede comprender una pluralidad de factores de ganancia relacionados con la banda de frecuencia.

10 De acuerdo con aspectos adicionales, pre-procesar la señal de audio codificada puede comprender re-cuantificar cada señal de banda de frecuencia utilizando un indicador de cuantificación específico de la banda de frecuencia de una pluralidad de indicadores de cuantificación específicos de la banda de frecuencia.

15 De acuerdo con aspectos adicionales, el método puede comprender además realizar un ajuste en forma de transición, comprendiendo el ajuste de forma de transición: encadenar el factor de cambio de nivel actual y un factor de cambio de nivel posterior para obtener un factor de cambio de nivel encadenado para uso durante la acción de la compensación al menos parcialmente del cambio de nivel.

De acuerdo con aspectos adicionales, el ajuste de forma de transición puede comprender además:

- almacenar temporalmente un factor de cambio de nivel anterior,
- 20 - generar una primera pluralidad de muestras en ventanas aplicando una forma de ventana al factor de cambio de nivel actual,
- generar una segunda pluralidad de muestras en ventanas aplicando una forma de ventana anterior al factor de cambio de nivel anterior proporcionado por la acción de almacenar temporalmente el factor de cambio de nivel anterior, y
- 25 - combinar mutuamente las muestras en ventanas correspondientes de la primera pluralidad de muestras en ventanas y de la segunda pluralidad de muestras en ventanas para obtener una pluralidad de muestras combinadas.

30 De acuerdo con aspectos adicionales, la forma de ventana y la forma de ventana anterior también se puede utilizar mediante la conversión del dominio de la frecuencia al dominio del tiempo por lo que se utiliza la misma forma de ventana y la forma de ventana anterior para convertir las señales de banda de frecuencia cambiadas de nivel en la representación en el dominio del tiempo y para aplicar formación de ventanas del factor de cambio de nivel actual y el factor de cambio de nivel anterior.

35 De acuerdo con aspectos adicionales, el factor de cambio de nivel actual puede ser válido para una trama actual de la pluralidad de las señales de banda de frecuencia, en la que el factor de cambio de nivel anterior puede ser válido para una trama anterior de la pluralidad de señales de banda de frecuencia, y en la que la trama actual y la trama anterior pueden superponerse. El ajuste de forma de transición se puede configurar

- para combinar el factor de cambio de nivel anterior con una segunda porción de la forma de ventana anterior que da como resultado una secuencia de factor de trama anterior,
- 40 - para combinar el factor de cambio de nivel actual con una primera porción de la forma de ventana actual que da como resultado una secuencia de factor de trama actual, y
- para determinar una secuencia del factor de cambio de nivel encadenado basándose en la secuencia de factor de trama anterior y la secuencia de factor de trama actual.

45 De acuerdo con aspectos adicionales, analizar la información adicional se puede realizar con respecto a si la información adicional sugiere un recorte potencial dentro de la representación en el dominio del tiempo lo que significa que un bit menos significativo no contiene información relevante, y en el que en este caso el cambio de nivel cambia la información hacia el bit menos significativo por lo que al liberar un bit más significativo se gana algún espacio en el bit más significativo.

50 De acuerdo con aspectos adicionales, se puede proporcionar un programa informático para implementar el método para decodificar o el método para codificar, cuando el programa informático se ejecuta en una computadora o procesador de señales.

55 Aunque algunos aspectos se han descrito en el contexto de un aparato, es claro que estos aspectos también representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a una etapa de método o una característica de una etapa de método. De forma análoga, los aspectos descritos en el contexto de una etapa de método también representan una descripción de un bloque o elemento o característica correspondiente de un aparato correspondiente.

60 La señal descompuesta inventiva se puede almacenar en un medio de almacenamiento digital o se puede transmitir en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión alámbrico tal como internet.

5 Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención se pueden implementar en hardware o software. La implementación se puede realizar utilizando un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM, o una memoria FLASH, que tienen señales de control legibles de forma electrónica almacenadas en los mismos, que cooperan (o son capaces de cooperar) con un sistema informático programable de manera que se realiza el método respectivo.

10 Algunas realizaciones de acuerdo con la invención comprenden un portador de datos no transitorio que tiene señales de control legibles de forma electrónica, que son capaces de cooperar con un sistema informático programable, de tal forma que se realiza uno de los métodos descritos en el presente documento.

15 En general, las realizaciones de la presente invención se pueden implementar como un producto de programa informático con un código de programa, siendo el código de programa operativo para realizar uno de los métodos cuando el producto de programa informático se ejecuta en una computadora. El código de programa puede almacenarse, por ejemplo, en un portador legible por máquina.

Otras realizaciones comprenden el programa informático para realizar uno de los métodos descritos en el presente documento, almacenado en un portador legible por máquina.

20 En otras palabras, una realización del método inventivo es, por lo tanto, un programa informático que tiene un código de programa para realizar uno de los métodos descritos en el presente documento, cuando el programa informático se ejecuta en una computadora.

25 Una realización adicional del método inventivo es, por lo tanto, un portador de datos (o un medio de almacenamiento digital, o un medio legible por computadora) que comprende, grabado en el mismo, el programa informático para realizar uno de los métodos descritos en el presente documento.

30 Una realización adicional del método inventivo es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa informático para realizar uno de los métodos descritos en el presente documento. El flujo de datos o la secuencia de señales pueden configurarse, por ejemplo, para que se transfieran mediante una conexión de comunicación de datos, por ejemplo mediante internet.

35 Una realización adicional comprende un medio de procesamiento, por ejemplo una computadora, o un dispositivo lógico programable, configurado para o adaptado para realizar uno de los métodos descritos en el presente documento.

Una realización adicional comprende una computadora que tiene instalado en la misma el programa informático para realizar uno de los métodos descritos en el presente documento.

40 En algunas realizaciones, un dispositivo lógico programable (por ejemplo un campo de matriz de puertas programables) se puede utilizar para realizar algunas o todas las funcionalidades de los métodos descritos en el presente documento. En algunas realizaciones, un campo de matriz de puertas programables puede cooperar con un microprocesador para realizar uno de los métodos descritos en el presente documento. En general, los métodos se realizan de manera preferente por cualquier aparato de hardware.

45 Las realizaciones descritas anteriormente son simplemente ilustrativas para los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en el presente documento serán evidentes para otros expertos en la técnica. Se propone, por lo tanto, que se limite solamente por el alcance de las reivindicaciones de la patente próximas y no por los detalles específicos presentados a manera de descripción y explicación de las realizaciones en el presente documento.

50

## REIVINDICACIONES

1. Un decodificador de señales de audio (100) configurado para proporcionar una representación de señal de audio decodificada basándose en una representación de señal de audio codificada, comprendiendo el decodificador de  
5 señales de audio:
- una etapa de pre-procesamiento de decodificador (110) configurada para obtener una pluralidad de señales de banda de frecuencia a partir de la representación de la señal de audio codificada;
- 10 un estimador de recorte (120) configurado para analizar la información adicional con respecto a una ganancia de las señales de banda de frecuencia de la representación de la señal de audio codificada en cuanto a si la información adicional sugiere un recorte potencial para determinar un factor de cambio de nivel actual para la representación de señal de audio codificada, en el que cuando la información adicional sugiere el recorte potencial, el factor de cambio de nivel actual provoca que la información de la pluralidad de señales de banda de frecuencia se cambie hacia un bit  
15 menos significativo de tal forma que se gana un espacio en al menos un bit más significativo;
- un cambiador de nivel (130) configurado para cambiar los niveles de las señales de banda de frecuencia de acuerdo con el factor de cambio de nivel actual para obtener las señales de banda de frecuencia cambiadas de nivel;
- 20 un convertidor del dominio de la frecuencia al dominio del tiempo (140) configurado para convertir las señales de banda de frecuencia cambiadas de nivel a una representación en el dominio del tiempo; y
- un compensador de cambio de nivel (150) configurado para actuar sobre la representación en el dominio del tiempo para compensar al menos parcialmente un cambio de nivel aplicado a las señales de banda de frecuencia cambiadas de nivel por el cambiador de nivel (130) y para obtener una representación en el dominio del tiempo  
25 sustancialmente compensada.
2. Decodificador de señales de audio (100) de acuerdo con la reivindicación 1, en el que el estimador de recorte (120) se configura además para determinar una probabilidad de recorte basándose en al menos una de la información adicional y la representación de señal de audio codificada, y para determinar el factor de cambio de nivel actual basándose en la probabilidad de recorte.
3. Decodificador de señales de audio (100) de acuerdo con la reivindicación 1 o 2, en el que la información adicional comprende al menos uno de un factor de ganancia global para la pluralidad de señales de banda de frecuencia y una pluralidad de factores de escala, correspondiendo cada factor de escala a una señal de banda de frecuencia o un grupo de señales de banda de frecuencia dentro de la pluralidad de señales de banda de frecuencia.
- 35 4. Decodificador de señales de audio (100) de acuerdo con una cualquiera de las reivindicaciones precedentes, en el que la etapa de pre-procesamiento de decodificador (110) se configura para obtener la pluralidad de señales de banda de frecuencia en la forma de una pluralidad de tramas sucesivas, y en el que el estimador de recorte (120) se configura para determinar el factor de cambio de nivel actual para una trama actual.
- 40 5. Decodificador de señales de audio (100) de acuerdo con una cualquiera de las reivindicaciones precedentes, en el que la representación de señal de audio decodificada se determina basándose en la representación en el dominio del tiempo sustancialmente compensada.
- 45 6. Decodificador de señales de audio (100) de acuerdo con una cualquiera de las reivindicaciones precedentes, que comprende además un limitador en el dominio del tiempo corriente abajo del compensador de cambio de nivel (150).
- 50 7. Decodificador de señales de audio (100) de acuerdo con una cualquiera de las reivindicaciones precedentes, en el que la información adicional con respecto a la ganancia de las señales de banda de frecuencia comprende una pluralidad de factores de ganancia relacionados con la banda de frecuencia.
8. Decodificador de señales de audio (100) de acuerdo con una cualquiera de las reivindicaciones precedentes, en el que la etapa de pre-procesamiento de decodificador (110) comprende un cuantificador inverso configurado para re-cuantificar cada señal de banda de frecuencia utilizando un indicador de cuantificación específico de la banda de frecuencia de una pluralidad de indicadores de cuantificación específicos de la banda de frecuencia.
- 55 9. Decodificador de señales de audio (100) de acuerdo con una cualquiera de las reivindicaciones precedentes, que comprende además un ajustador de forma de transición configurado para encadenar el factor de cambio de nivel actual y un factor de cambio de nivel posterior para obtener un factor de cambio de nivel encadenado para uso por el compensador (150) de cambio de nivel.
- 60 10. Decodificador de señales de audio (100) de acuerdo con la reivindicación 9, en el que el ajustador de forma de

- transición comprende una memoria (371) para un factor de cambio de nivel anterior, un primer sistema de formación de ventanas (372) configurado para generar una primera pluralidad de muestras en ventanas aplicando una forma de ventana al factor de cambio de nivel actual, un segundo sistema de formación de ventanas (376) configurado para generar una segunda pluralidad de muestras en ventanas aplicando una forma de ventana anterior al factor de cambio de nivel anterior proporcionado por la memoria (371), y un combinador de muestras (379) configurado para combinar mutuamente las muestras en ventanas correspondientes de la primera pluralidad de muestras en ventanas y de la segunda pluralidad de muestras en ventanas para obtener una pluralidad de muestras combinadas.
- 5
11. Decodificador de señales de audio (100) de acuerdo con la reivindicación 10,
- 10 en el que el factor de cambio de nivel actual es válido para una trama actual de la pluralidad de señales de banda de frecuencia, en el que el factor de cambio de nivel anterior es válido para una trama anterior de la pluralidad de señales de banda de frecuencia, y en el que la trama actual y la trama anterior se superponen;
- en el que el ajuste de forma de transición se configura
- 15 para combinar el factor de cambio de nivel anterior con una segunda porción de la forma de ventana anterior que da como resultado una secuencia de factor de trama anterior,
- para combinar el factor de cambio de nivel actual con una porción de la forma de ventana actual que da como resultado una secuencia de factor de trama actual, y
- 20 para determinar una secuencia del factor de cambio de nivel encadenado basándose en la secuencia de factor de trama anterior y la secuencia de factor de trama actual.
12. Decodificador de señales de audio (100) de acuerdo con una cualquiera de las reivindicaciones precedentes, en el que el estimador de recorte (120) se configura para analizar al menos una de la representación de señal de audio codificada y la información adicional con respecto a si al menos una de la representación de señal de audio codificada y la información adicional sugieren un recorte potencial dentro de la representación en el dominio del tiempo lo que significa que un bit menos significativo no contiene información relevante, y en el que en este caso el cambio de nivel aplicado por el cambiador de nivel cambia la información hacia el bit menos significativo por lo que al liberar un bit más significativo se gana algún espacio en el bit más significativo.
- 25
13. Decodificador de señales de audio (100) de acuerdo con una cualquiera de las reivindicaciones precedentes, en el que el estimador de recorte (120) comprende:
- 30 un determinador del libro de códigos (1110) para determinar un libro de códigos a partir de una pluralidad de libros de códigos como un libro de códigos identificado, en el que la representación de señal de audio codificada se ha codificado empleando el libro de códigos identificado, y
- 35 una unidad de estimación (1120) configurada para derivar un valor de nivel asociado con el libro de códigos identificado como un valor de nivel derivado y, para estimar una estimación de nivel de la señal de audio utilizando el valor de nivel derivado.
- 40
14. Codificador de señales de audio configurado para proporcionar una representación de señal de audio codificada basándose en una representación en el dominio del tiempo de una señal de audio de entrada, comprendiendo el codificador de señales de audio:
- 45 un estimador de recorte configurado para analizar la representación en el dominio del tiempo de la señal de audio de entrada en cuanto a sí se sugiere el recorte potencial para determinar un factor de cambio de nivel actual para la representación de señal de entrada, en el que cuando se sugiere el recorte potencial, el factor de cambio de nivel actual provoca que la representación en el dominio del tiempo de la señal de audio de entrada se cambie hacia un bit menos significativo por lo que se gana un espacio en al menos un bit más significativo;
- 50 un cambiador de nivel configurado para cambiar un nivel de la representación en el dominio del tiempo de la señal de audio de entrada de acuerdo con el factor de cambio de nivel actual para obtener una representación en el dominio del tiempo cambiada de nivel;
- 55 un convertidor del dominio del tiempo al dominio de la frecuencia configurado para convertir la representación en el dominio del tiempo cambiada de nivel a una pluralidad de señales de banda de frecuencia; y
- un compensador de cambio de nivel configurado para actuar sobre la pluralidad de señales de banda de frecuencia para compensar al menos parcialmente un cambio de nivel aplicado a la representación en el dominio del tiempo cambiada de nivel mediante el cambiador de nivel y para obtener una pluralidad de señales de banda de frecuencia sustancialmente compensadas.
- 60
15. Método para decodificar una representación de señal de audio codificada y para proporcionar una representación de señal de audio decodificada correspondiente, comprendiendo el método:

pre-procesar la representación de señal de audio codificada para obtener una pluralidad de señales de banda de frecuencia;

5 analizar la información adicional con respecto a una ganancia de las señales de banda de frecuencia en cuanto a si la información adicional sugiere un recorte potencial para determinar un factor de cambio de nivel actual para la representación de señal de audio codificada, en el que cuando la información adicional sugiere el recorte potencial, el factor de cambio de nivel actual provoca que la información de la pluralidad de señales de banda de frecuencia se cambie hacia un bit menos significativo por lo que se gana un espacio en al menos un bit más significativo;

10 cambiar los niveles de las señales de banda de frecuencia de acuerdo con el factor de cambio de nivel para obtener las señales de banda de frecuencia cambiadas de nivel;

15 realizar una conversión del dominio de la frecuencia al dominio del tiempo de las señales de banda de frecuencia a una representación en el dominio del tiempo; y

20 actuar sobre la representación en el dominio del tiempo para compensar al menos parcialmente un cambio de nivel aplicado a las señales de banda de frecuencia cambiadas de nivel y para obtener una representación en el dominio del tiempo sustancialmente compensada.

25 16. Método de codificación de señal de audio para proporcionar una representación de señal de audio codificada basándose en una representación en el dominio del tiempo de una señal de audio de entrada, comprendiendo el método:

30 analizar la representación del dominio del tiempo de la señal de audio de entrada en cuanto a si se sugiere recorte potencial para determinar un factor de cambio de nivel actual para la representación de señal de entrada, en el que cuando se sugiere el recorte potencial, el factor de cambio de nivel actual provoca que la representación en el dominio del tiempo de la señal de audio de entrada se cambie hacia un bit menos significativo por lo que se gana un espacio en al menos un bit más significativo;

35 cambiar un nivel de la representación en el dominio del tiempo de la señal de audio de entrada de acuerdo con el factor de cambio de nivel actual para obtener una representación en el dominio de tiempo cambiada de nivel;

40 convertir la representación en el dominio del tiempo cambiada de nivel en una pluralidad de señales de banda de frecuencia; y

actuar sobre la pluralidad de señales de banda de frecuencia para compensar al menos parcialmente un cambio de nivel aplicado a la representación en el dominio del tiempo cambiada de nivel mediante el cambio y para obtener una pluralidad de señales de banda de frecuencia sustancialmente compensadas.

17. Programa informático adaptado para ordenar a una computadora para que realice el método de la reivindicación 15 o 16.



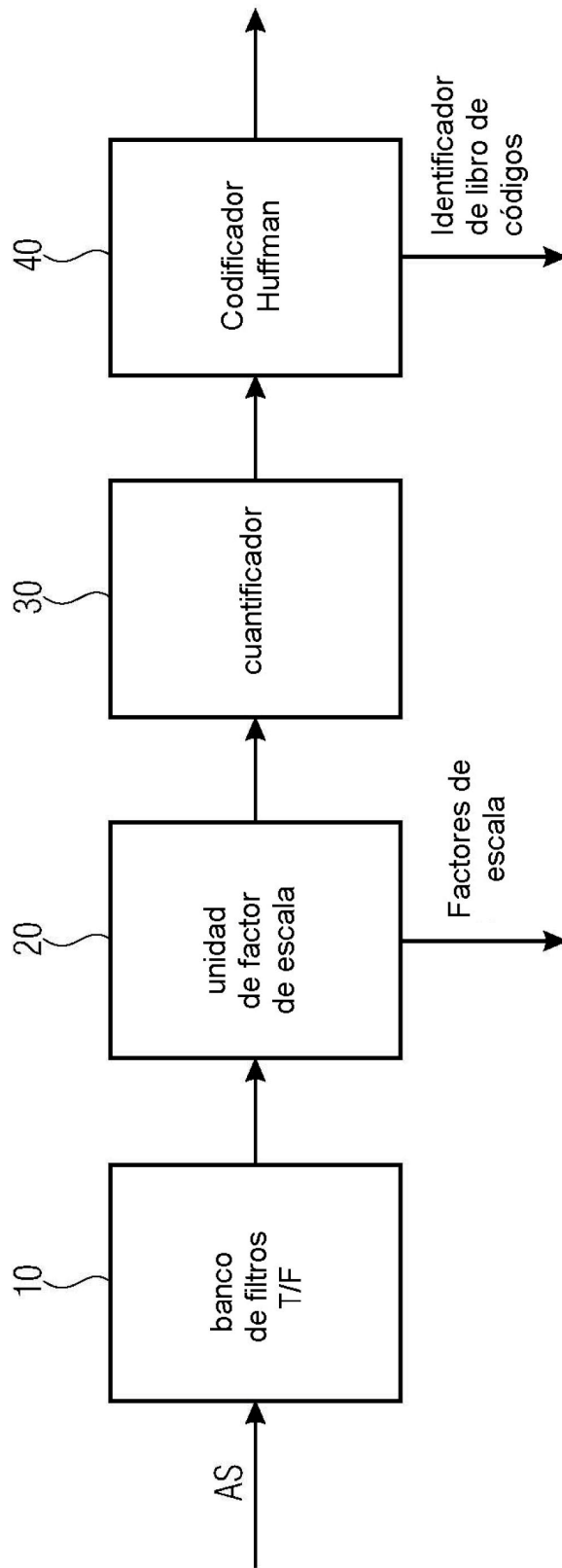


FIG 1

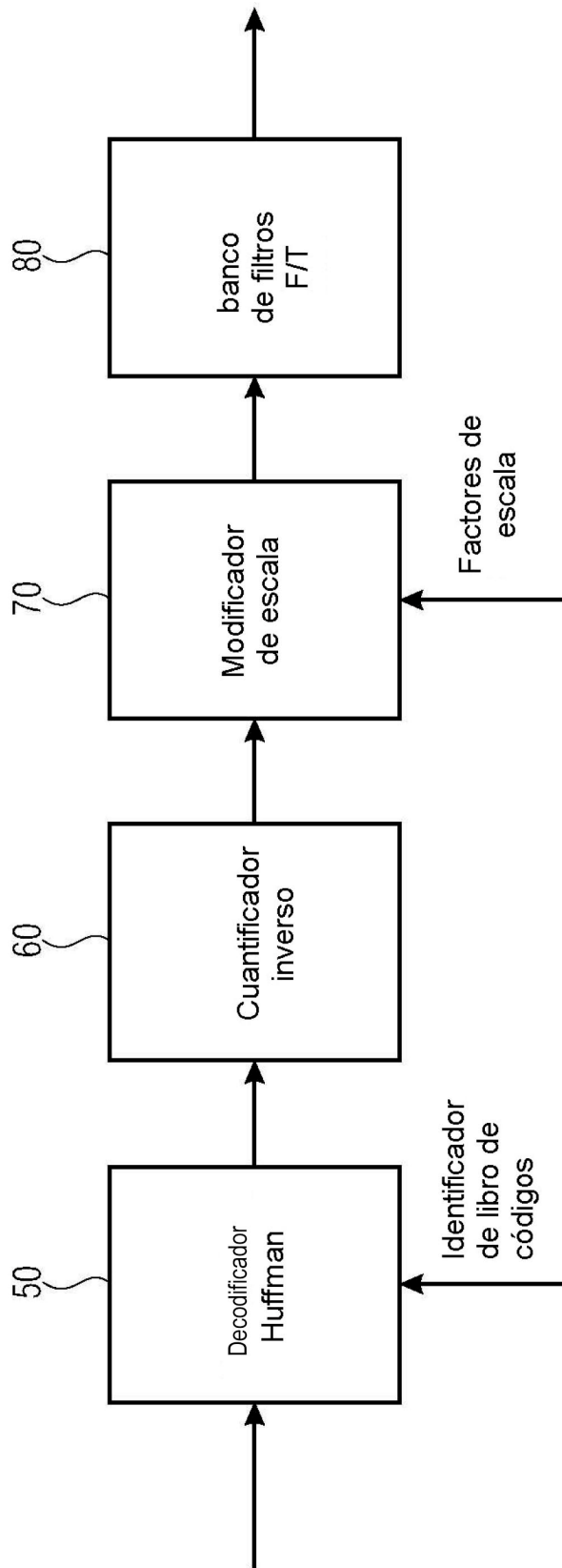


FIG 2

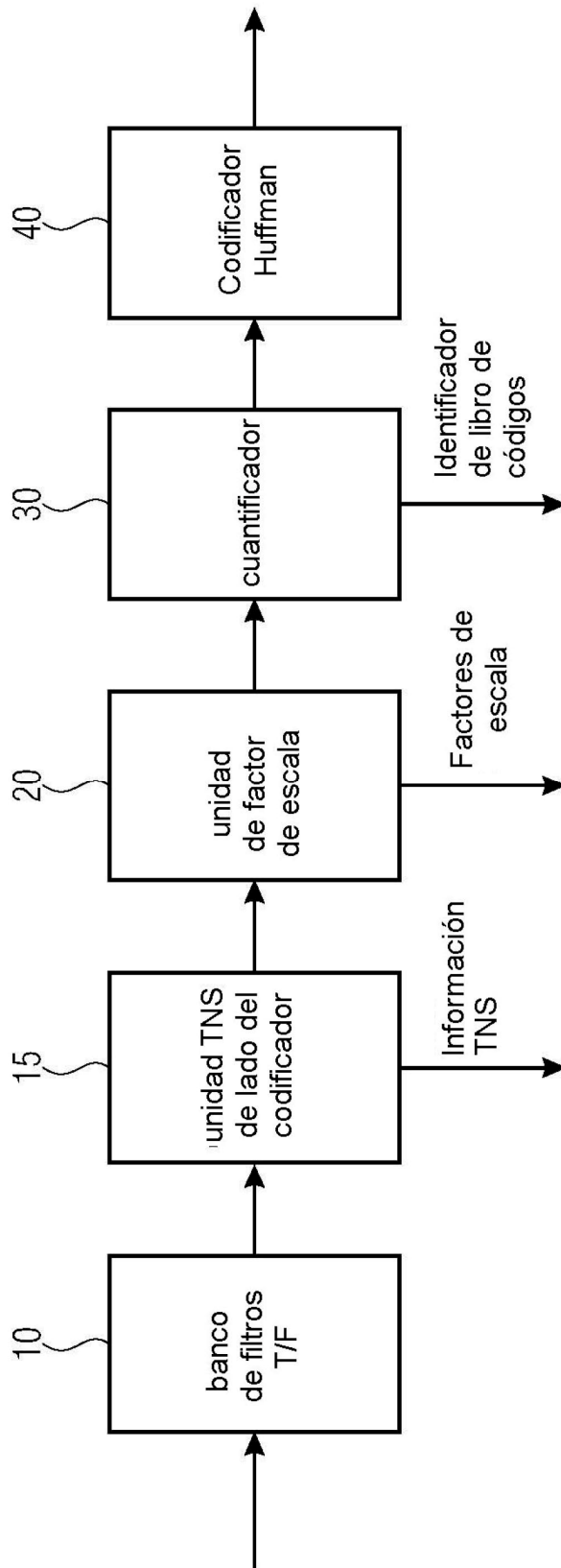


FIG 3

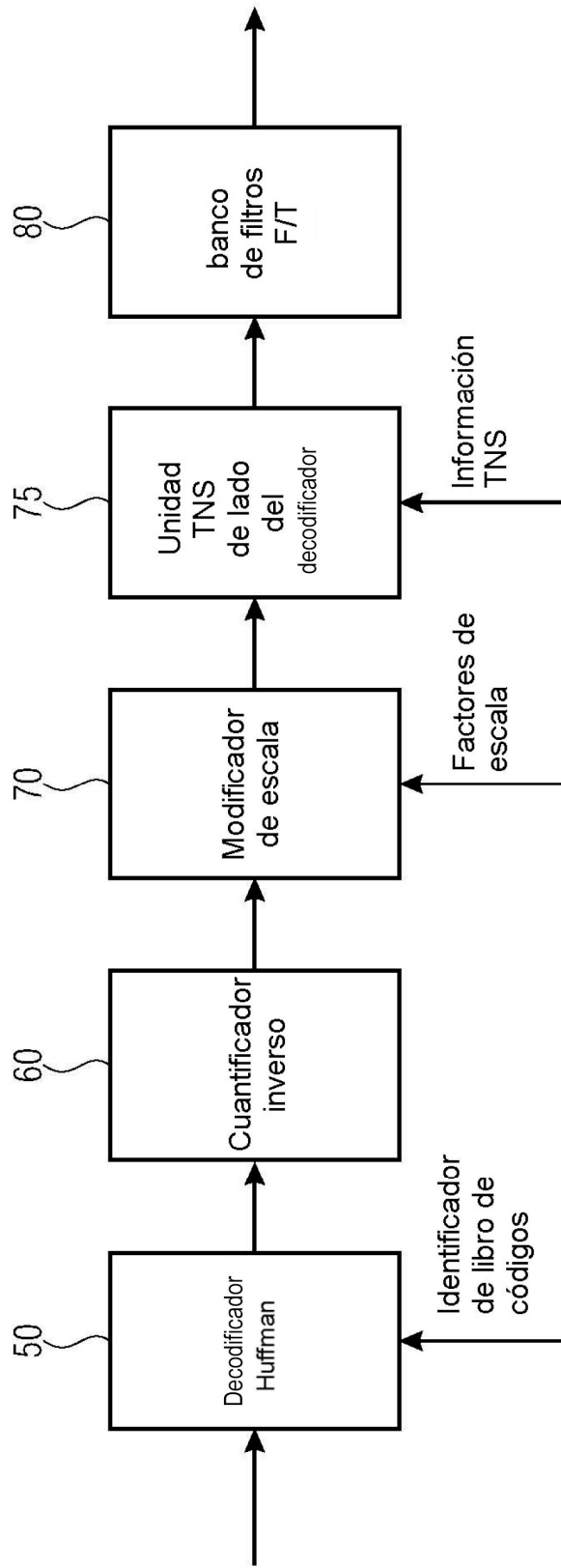


FIG 4

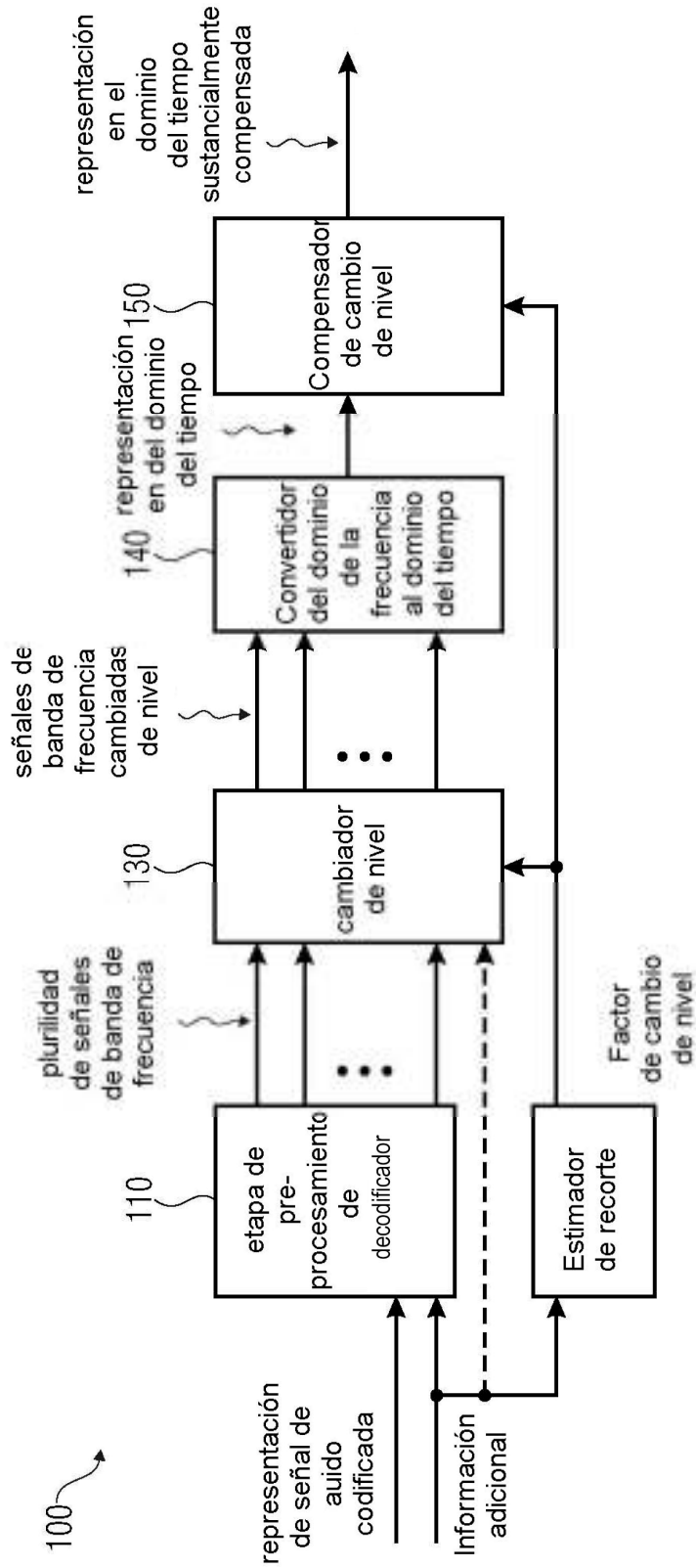


FIG 5

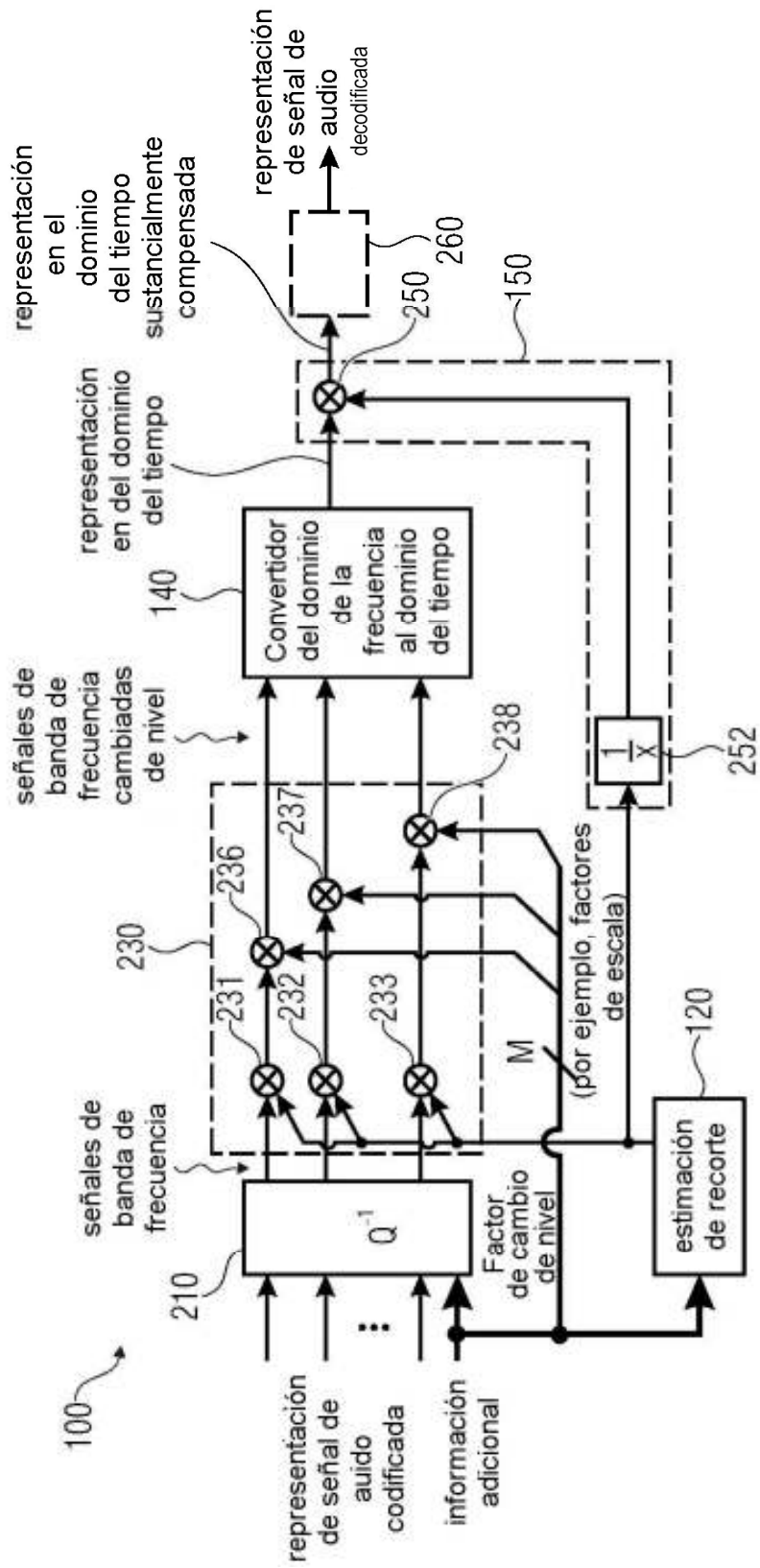


FIG 6

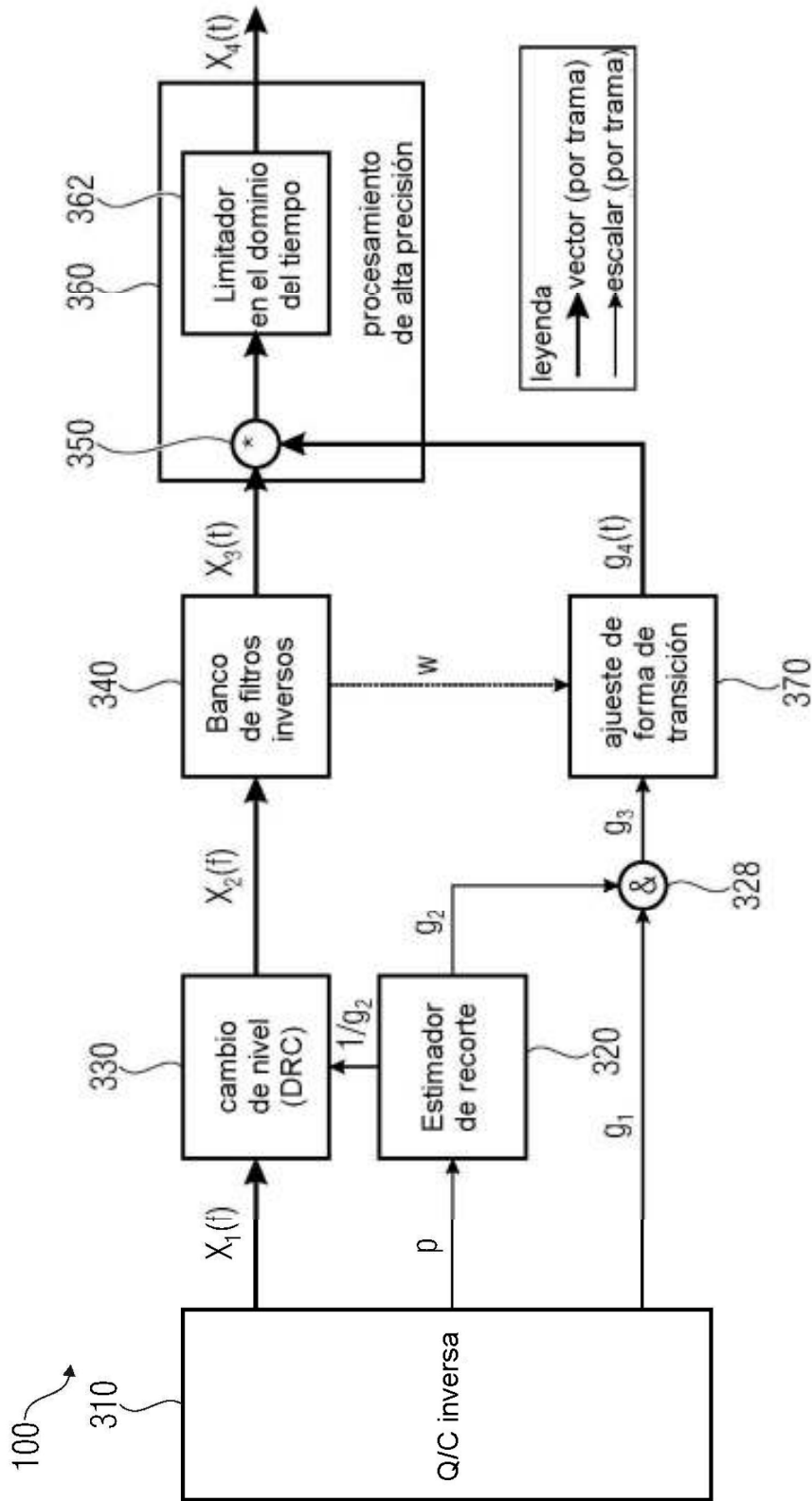


FIG 7

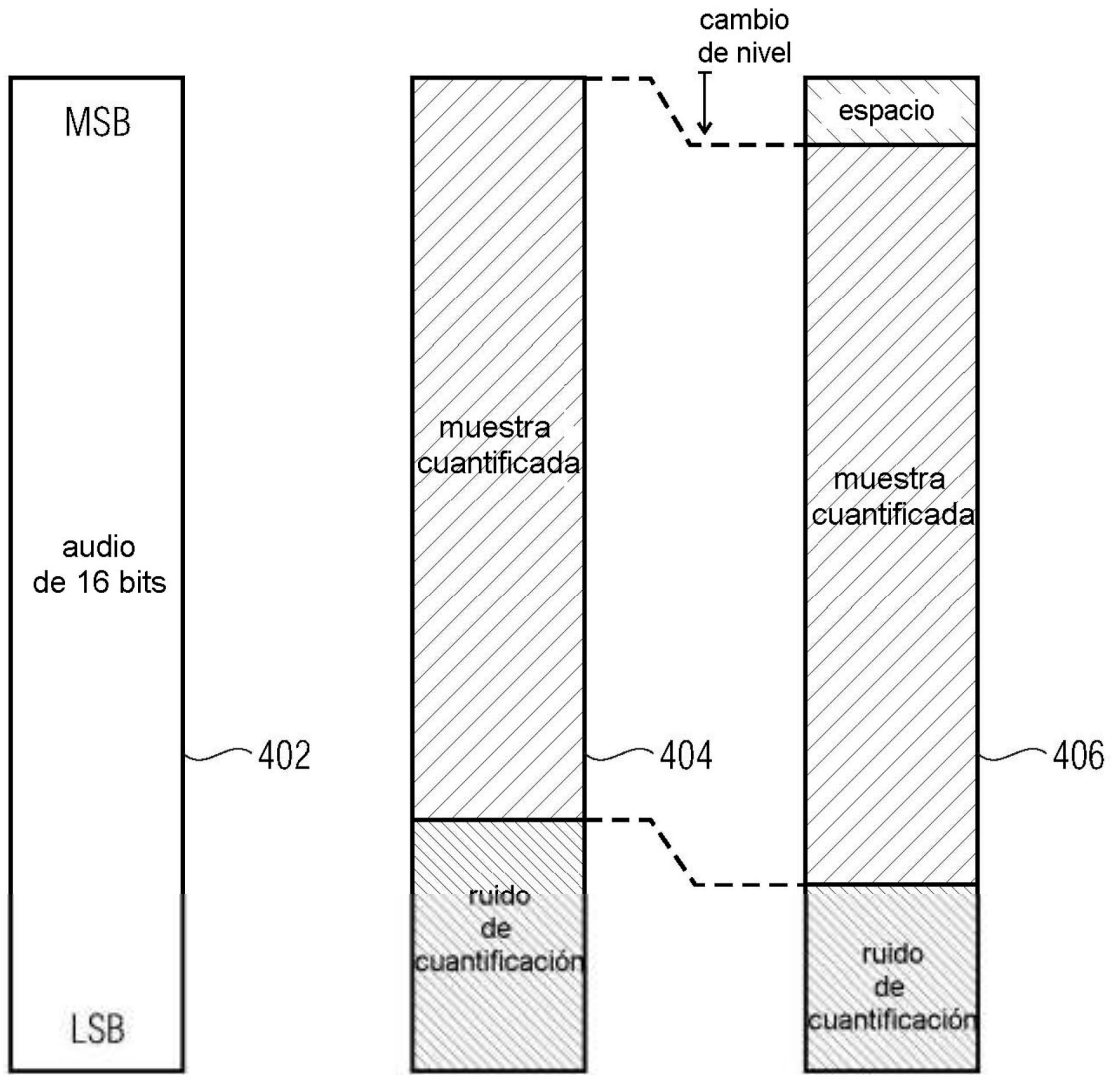


FIG 8



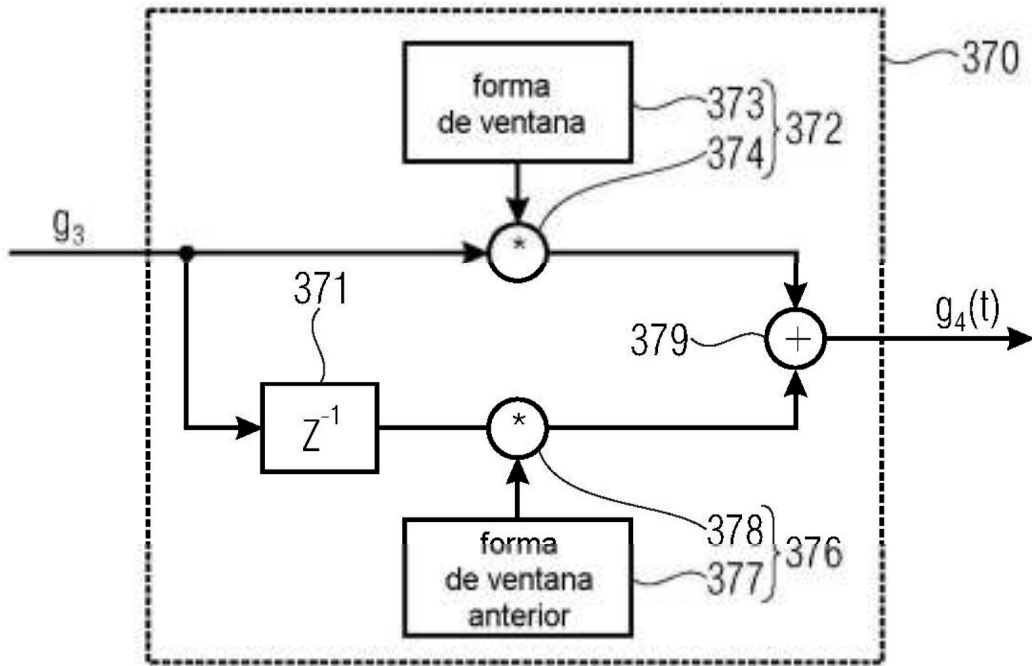


FIG 9

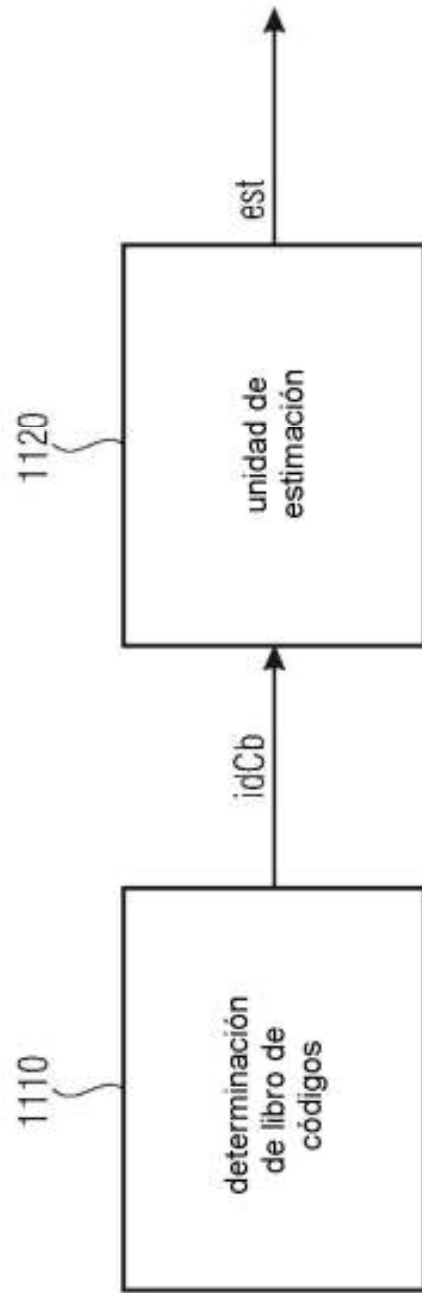


FIG 10

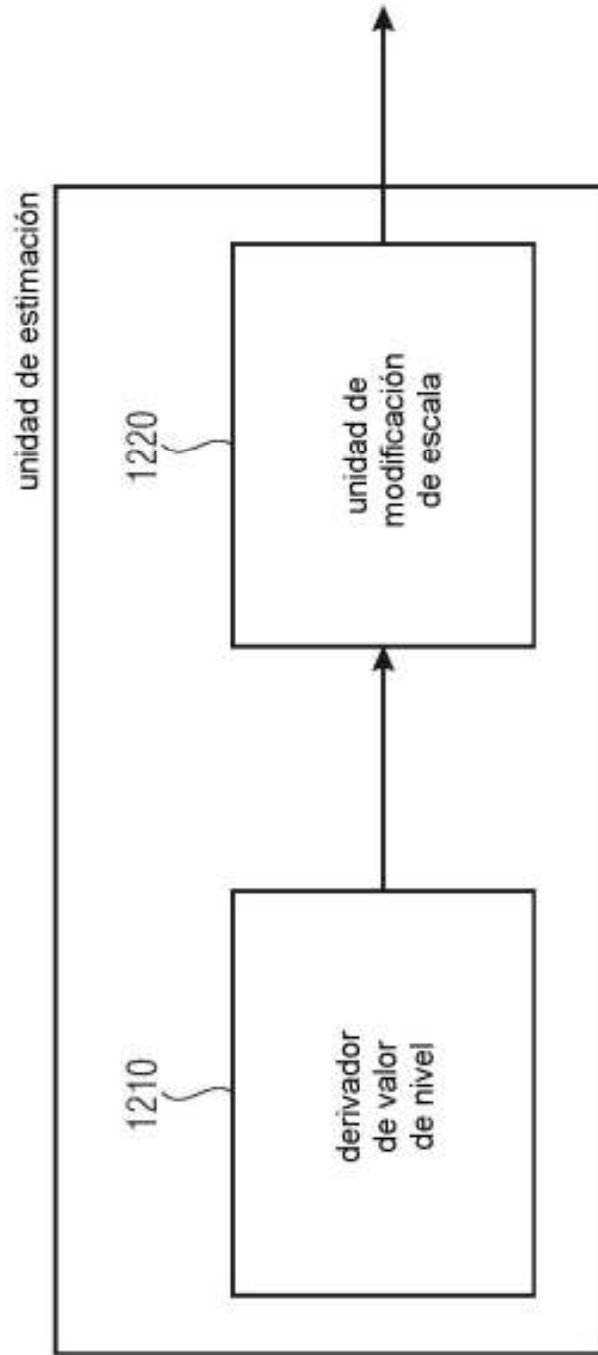


FIG 11

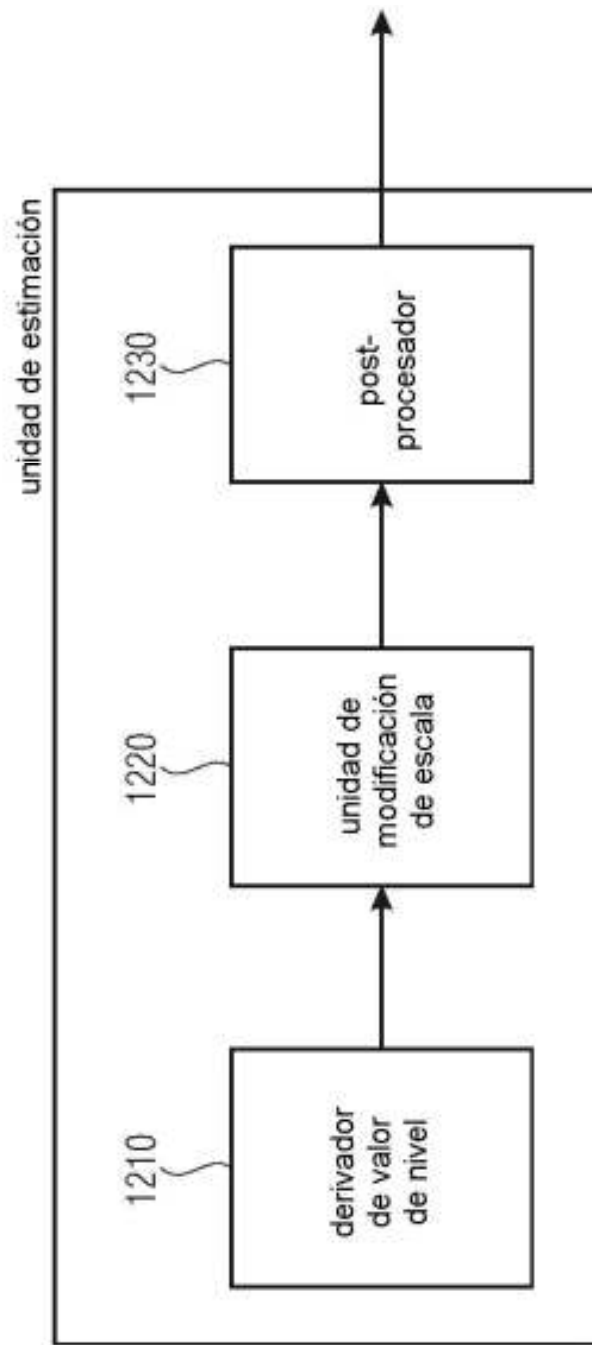


FIG 12

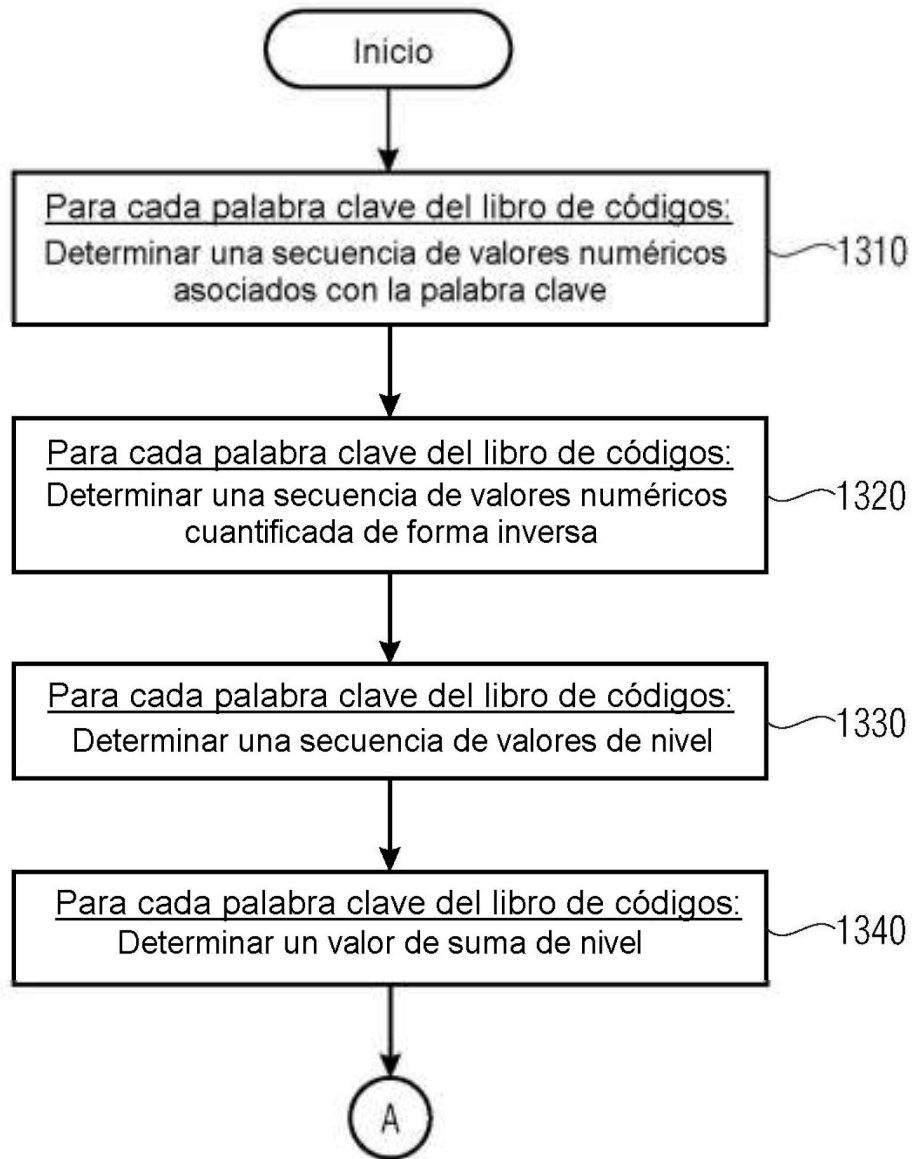


FIG 13A

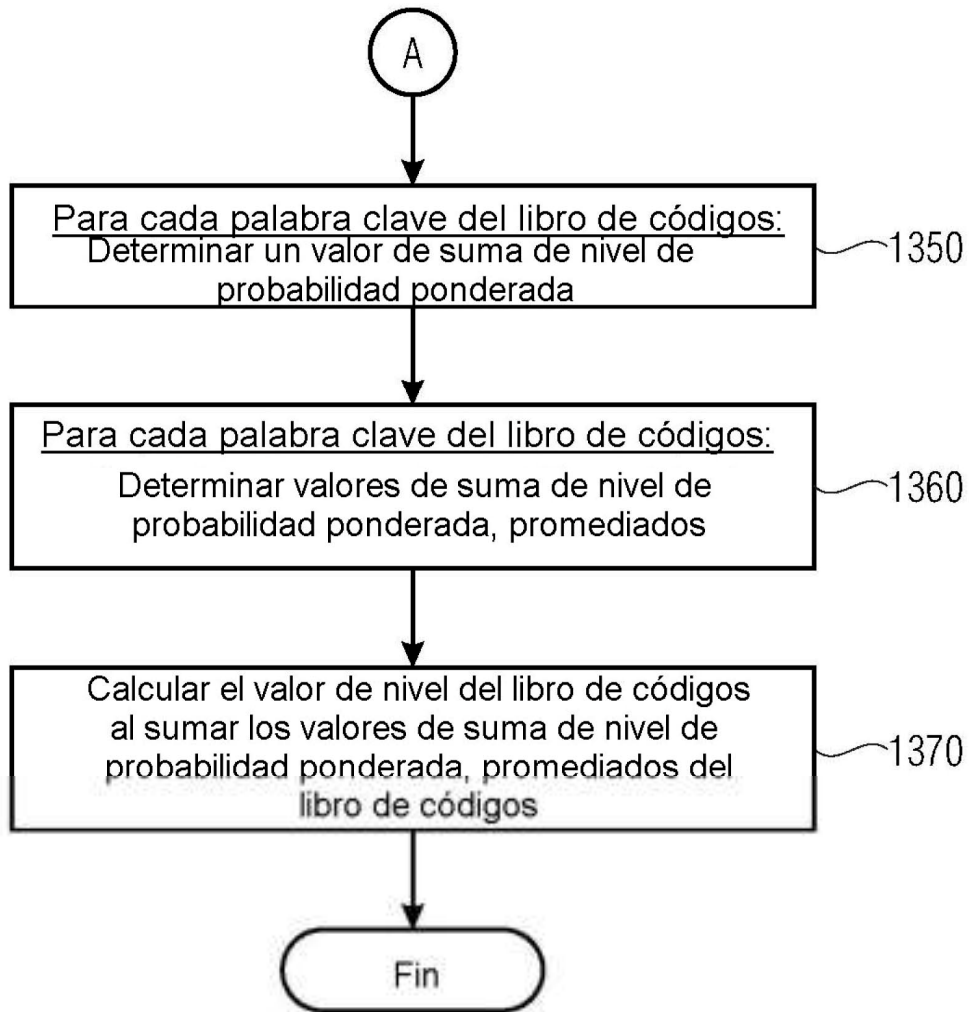


FIG 13B

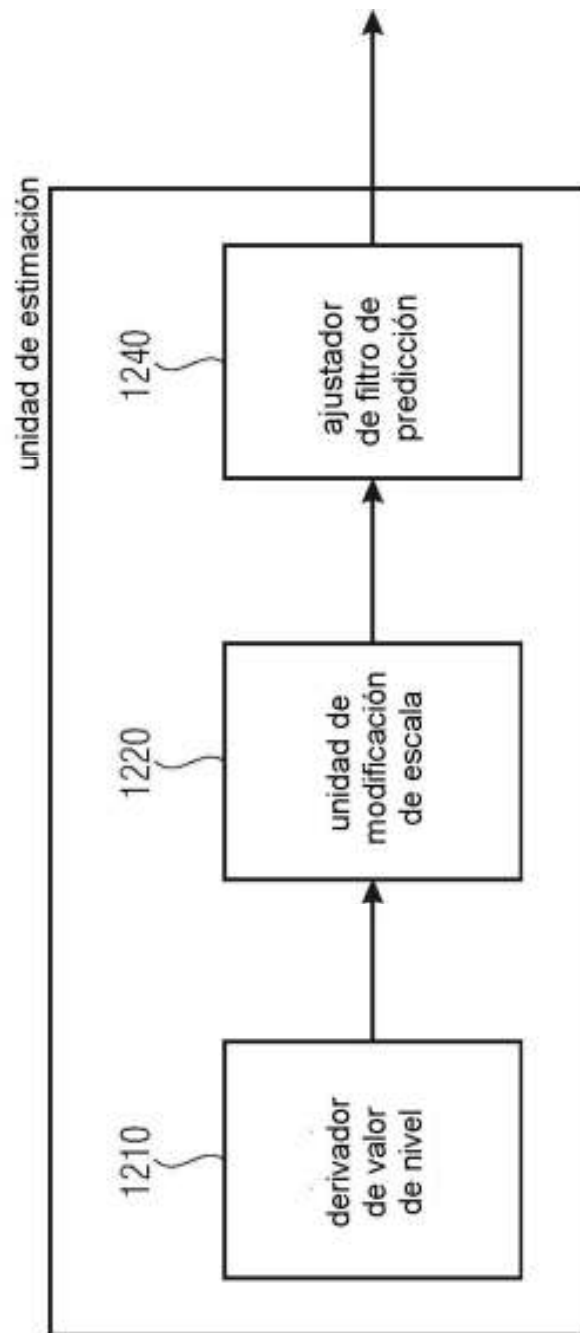


FIG 14

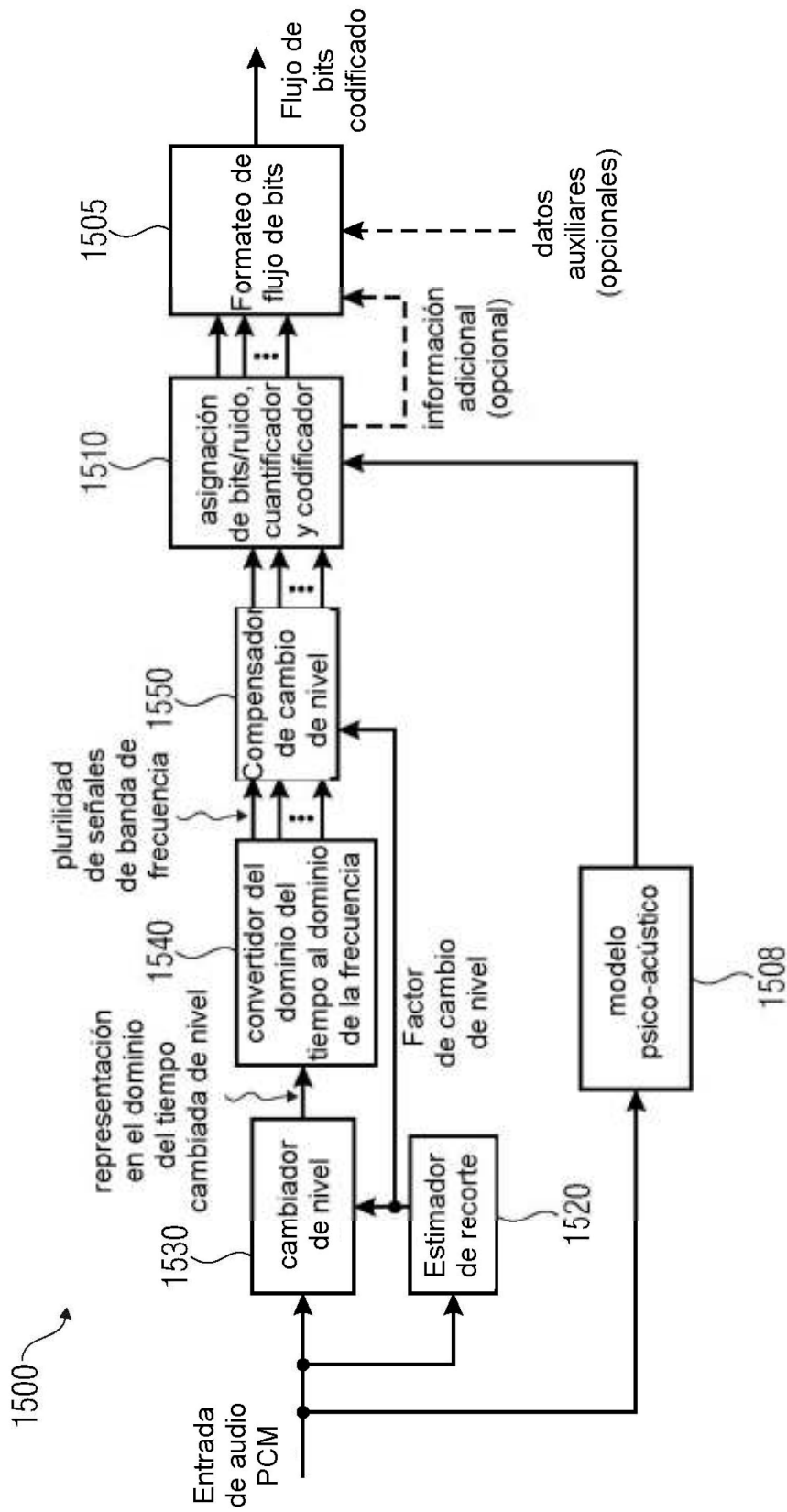


FIG 15



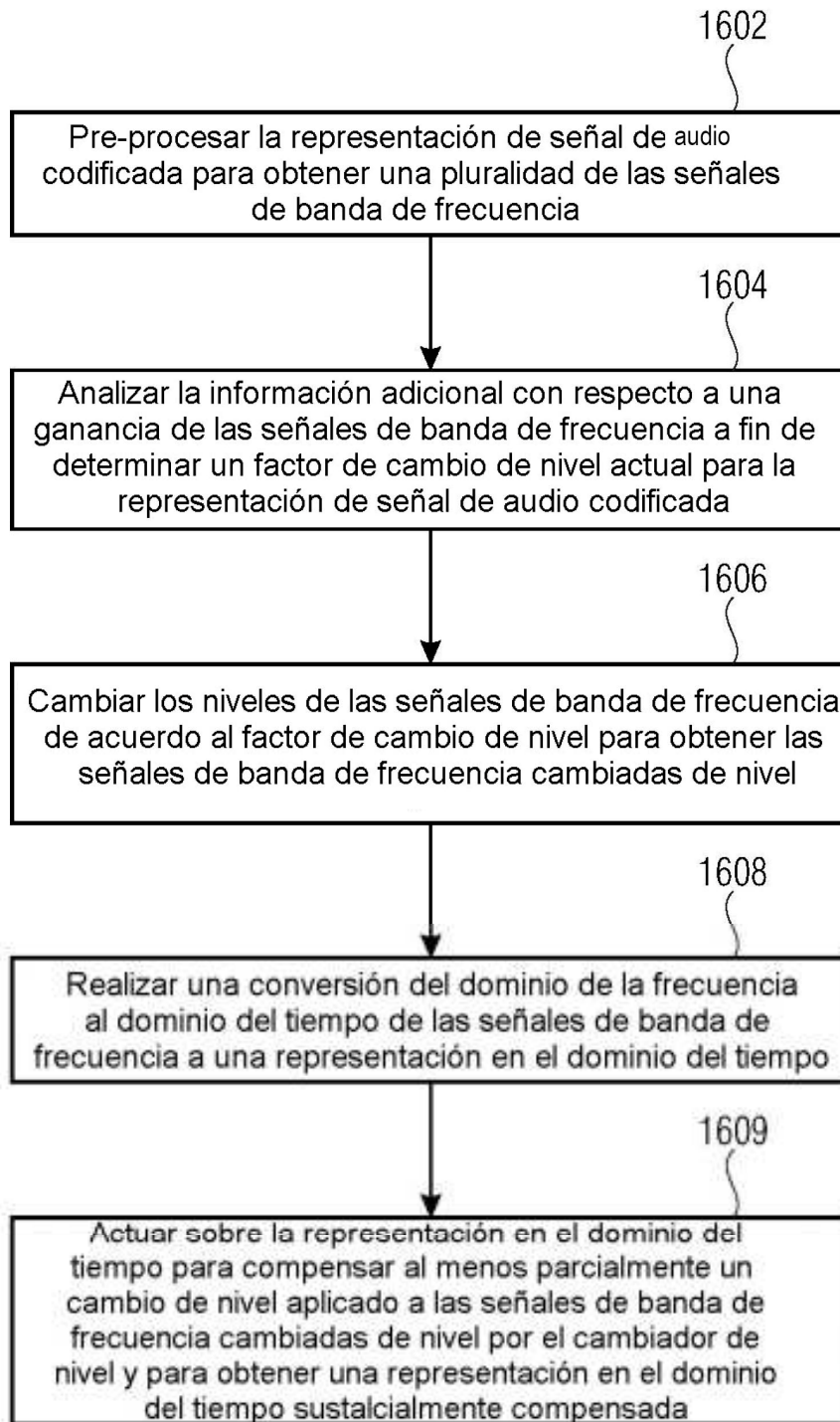


FIG 16