(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 955 669**

(51) Int. CI.:

G10L 19/028 (2013.01)
G10L 19/02 (2013.01)
G10L 19/032 (2013.01)
G10L 25/18 (2013.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **25.06.2009 E 17175883 (2)**

(97) Fecha y número de publicación de la concesión europea: **14.06.2023 EP 3246918**

(54) Título: **Decodificador de audio, procedimiento para decodificar una señal de audio y programa de ordenador**

(30) Prioridad:

11.07.2008 US 79872 P
08.10.2008 US 103820 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
05.12.2023

(73) Titular/es:

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

(72) Inventor/es:

RETTELBACH, NIKOLAUS;
GRILL, BERNHARD;
FUCHS, GUILLAUME;
GEYERSBERGER, STEFAN;
MULTRUS, MARKUS;
POPP, HARALD;
HERRE, JUERGEN;
WABNIK, STEFAN;
SCHULLER, GERALD y
HIRSCHFELD, JENS

(74) Agente/Representante:

PONTI & PARTNERS, S.L.P.

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

DESCRIPCIÓN

Decodificador de audio, procedimiento para decodificar una señal de audio y programa de ordenador

ANTECEDENTES DE LA INVENCIÓN

[0001] Las realizaciones según la invención se refieren a un decodificador para proporcionar una representación decodificada de una señal de audio basándose en un flujo de audio codificado, un procedimiento para decodificar una señal de audio y un programa de ordenador para decodificar una señal de audio.

[0002] En términos generales, las realizaciones según la invención se refieren a un relleno de ruido.

[0003] Los conceptos de codificación de audio a menudo codifican una señal de audio en el dominio de la frecuencia. Por ejemplo, el concepto denominado de “codificación de audio avanzada” (AAC) codifica el contenido de diferentes segmentos espectrales (o segmentos de frecuencia) teniendo en cuenta un modelo psicoacústico. Para este propósito, se codifica la información de intensidad para diferentes segmentos espectrales. Sin embargo, la resolución utilizada para codificar intensidades en diferentes segmentos espectrales se adapta de acuerdo con las relevancias psicoacústicas de los diferentes segmentos espectrales. Así, algunos segmentos espectrales, que se considera que son de baja relevancia psicoacústica, se codifican con una resolución de intensidad muy baja, de modo que algunos de los segmentos espectrales que se consideran de baja relevancia psicoacústica, o incluso un número dominante de los mismos, se cuantifican a cero. Cuantificar la intensidad de un segmento espectral a cero lleva aparejada la ventaja de que el valor cuantificado a cero puede codificarse de una manera que ahorra muchos bits, lo cual ayuda a mantener la velocidad binaria lo más pequeña posible. No obstante, los segmentos espectrales cuantificados a cero a veces resultan en artefactos audibles, aunque el modelo psicoacústico indique que los segmentos espectrales son de baja relevancia psicoacústica.

[0004] Por lo tanto, existe el deseo de tratar los segmentos espectrales cuantificados a cero, tanto en un codificador de audio como en un decodificador de audio.

[0005] Se conocen diferentes estrategias para tratar los segmentos espectrales codificados a cero en los sistemas de codificación de audio en el dominio de la transformada y también en los codificadores de voz.

[0006] Por ejemplo, la norma MPEG-4 “AAC” (codificación de audio avanzada) utiliza el concepto de sustitución de ruido perceptivo (PNS). La sustitución de ruido perceptivo rellena bandas de factores de escala completas sólo con ruido. Los detalles respecto al MPEG-4 AAC pueden encontrarse, por ejemplo, en la Norma Internacional ISO /IEC 14496-3 (Tecnología de la Información – Codificación de Objetos Audiovisuales – Parte 3: Audio). Además, el codificador de voz AMR-WB+ reemplaza los vectores de cuantificación vectorial (vectores VQ) cuantificados a cero con un vector de ruido aleatorio, donde cada valor espectral complejo tiene una amplitud constante, pero una fase aleatoria. La amplitud se controla mediante un valor de ruido transmitido con el tren de bits. Los detalles respecto al codificador de voz AMR-WB+ pueden encontrarse, por ejemplo, en la especificación técnica titulada “Third Generation Partnership Project; Technical Specification Group Services and System Aspects: Audio Codec Processing Functions; Extended Adaptive Multi-Rate-Wide Band (AMR-WB+) Codec; Transcoding Functions (Release Six)”, que también se conoce como “3GPP TS 26.290 V6.3.0 (2005-06) – Technical Specification”.

[0007] Además, el documento EP1395980B1 describe un concepto de codificación de audio. La publicación describe un medio por el cual bandas de frecuencias seleccionadas de información de una señal de audio original, que son audibles, pero que son perceptivamente menos relevantes, no necesitan ser codificadas, sino que pueden reemplazarse por un parámetro de relleno de ruido. En cambio, aquellas bandas de señal que tienen contenido, que son más relevantes perceptivamente, se codifican totalmente. De esta manera se ahorran bits de codificación sin dejar huecos en el espectro de frecuencia de la señal recibida. El parámetro de relleno de ruido es una medida del valor de la señal RMS dentro de la banda en cuestión y se utiliza en el extremo de recepción por un algoritmo de decodificación para indicar la cantidad de ruido que hay que inyectar en la banda de frecuencias en cuestión.

[0008] Estrategias adicionales proporcionan una inserción de ruido no guiada en el decodificador, teniendo en cuenta la tonalidad del espectro transmitido.

[0009] Sin embargo, los conceptos convencionales típicamente llevan aparejado el problema de que o bien comprenden una escasa resolución en cuanto a la granularidad del relleno de ruido, lo cual típicamente degrada la impresión de audición, o requieren una cantidad comparativamente grande de información colateral de relleno de ruido, lo cual requiere velocidad binaria extra.

[0010] Las estrategias conocidas del documento de Neuendorf y col. “Detailed Technical Description of Reference Model 0 of the CfP on Unified Speech and Audio Coding (USAC)”m 86, Convención MPEG; Motion Picture Expert Group o la norma ISO/IEC JTC1/SC29/WG11, nº M15867, 9 de octubre de 2008, XP030044464, y del documento de Neuendorf y col, “A Novel Scheme for Low Bitrate Unified Speech and Audio Coding – MPEG RM0”, presentado en la 126 Convención, mayo de 2009, XP040508995, trata del problema mencionado anteriormente y permite el relleno de ruido de líneas espectrales individuales basándose en un único valor que representa el ruido de cuantificación o un error medio de cuantificación.

[0011] En vista de lo anterior, existe la necesidad de un concepto mejorado de relleno de ruido, que proporcione un mejor compromiso entre la impresión de audición alcanzable y la velocidad binaria requerida.

RESUMEN DE LA INVENCIÓN

[0012] Una realización según la invención crea un decodificador para proporcionar una representación decodificada de una señal de audio como se expone en la reivindicación 1.

[0013] El decodificador se basa en el hallazgo de que puede aplicarse un único valor de intensidad de ruido multibanda para un relleno de ruido con buenos resultados si se asocia información de ganancia de banda de frecuencia separada a las distintas bandas de frecuencia. Por consiguiente, se realiza un cambio de escala individual del ruido introducido en las diferentes bandas de frecuencia basándose en la información de ganancia de banda de frecuencia, de manera que, por ejemplo, el único valor de intensidad de ruido multibanda común proporciona, cuando se toma en combinación con información de ganancia de banda de frecuencia separada, información suficiente para introducir ruido de una manera adaptada a la psicoacústica humana. Así, el concepto descrito en este documento permite aplicar un relleno de ruido en el dominio cuantificado (pero no redimensionado). El ruido añadido en el decodificador puede cambiarse de escala con la relevancia psicoacústica de la banda sin requerir información colateral adicional (más allá de la información colateral, que es, de todos modos, necesaria para cambiar de escala el contenido de audio sin ruido de las bandas de frecuencia de acuerdo con la relevancia psicoacústica de las bandas de frecuencia)..

[0014] Otra realización según la invención crea un procedimiento para proporcionar una representación decodificada de una señal de audio basándose en un flujo de audio codificado, como se expone en la reivindicación 2.

[0015] Una realización adicional según la invención crea un programa de ordenador, como se expone en la reivindicación 3.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0016]

La **Fig. 1** muestra un diagrama esquemático de bloques de un codificador que puede utilizarse en relación con realizaciones de la invención;
la **Fig. 2** muestra un diagrama esquemático de bloques de un codificador que puede utilizarse en relación con realizaciones de la invención;
las **Figs. 3a** y **3b** muestran un diagrama esquemático de bloques de una codificación de audio avanzada extendida (AAC) que puede utilizarse en relación con realizaciones de la invención;
las **Figs. 4a** y **4b** muestran listados de programas de pseudocódigo de algoritmos ejecutados para la codificación de una señal de audio;
la **Fig. 5** muestra un diagrama esquemático de bloques de un decodificador según una realización de la invención;
la **Fig. 6** muestra un diagrama esquemático de bloques de un decodificador en el que puede implementarse la invención;
las **Figs. 7a** y **7b** muestran un diagrama esquemático de bloques de un decodificador de AAC (codificación de audio avanzada) extendida en el que puede implementarse la invención;
la **Fig. 8a** muestra una representación matemática de una cuantificación inversa, que puede llevarse a cabo en el decodificador de AAC extendida de la Fig. 7;
la **Fig. 8b** muestra un listado de programas de pseudocódigo de un algoritmo para cuantificación inversa, que puede llevarse a cabo por el decodificador de AAC extendida de la Fig. 7;
la **Fig. 8c** muestra una representación de diagrama de flujo de la cuantificación inversa;
la **Fig. 9** muestra un diagrama esquemático de bloques de un rellenador de ruido y un medio de reajuste de escala,

que pueden utilizarse en el decodificador de AAC extendida de la Fig. 7;
la **Fig. 10a** muestra una representación de código de pseudocódigo de programa de un algoritmo, que puede ser ejecutado por el rellenador de ruido mostrado en la Fig. 7 o por el rellenador de ruido mostrado en la Fig. 9;
la **Fig. 10b** muestra una leyenda de elementos del pseudocódigo de programa de la Fig. 10a;
la **Fig. 11** muestra un diagrama de flujo de un procedimiento, que puede implementarse en el rellenador de ruido de la Fig. 7 o en el rellenador de ruido de la Fig. 9;
la **Fig. 12** muestra una ilustración gráfica del procedimiento de la Fig. 11;
las **Figs. 13a** y **13b** muestran representaciones de pseudocódigo de programa de algoritmos, que pueden llevarse a cabo por el rellenador de ruido de la Fig. 7 o por el rellenador de ruido de la Fig. 9;
las **Figs. 14a** a **14d** muestran representaciones de elementos de tren de bits de un flujo de audio que pueden utilizarse en relación con realizaciones de la invención; y
la **Fig. 15** muestra una representación gráfica de un tren de bits que puede utilizarse en relación con realizaciones de la invención.

DESCRIPCIÓN DETALLADA DE LAS REALIZACIONES

1. Codificador

1.1 Codificador según la Fig. 1

[0017] La Fig. 1 muestra un diagrama esquemático de bloques de un codificador para proporcionar un flujo de audio basándose en la representación en el dominio de la transformada de una señal de audio de entrada, que puede utilizarse en relación con realizaciones de la invención.

[0018] El codificador 100 de la Fig. 1 comprende un calculador de error de cuantificación 110 y un proveedor de flujo de audio 120. El calculador de error de cuantificación 110 está configurado para recibir una información 112 respecto a una primera banda de frecuencias, para la cual se dispone de una información de ganancia de primera banda de frecuencias, y una información 114 acerca de una segunda banda de frecuencias, para la cual se dispone de una información de ganancia de segunda banda de frecuencias. El calculador de error de cuantificación está configurado para determinar un error de cuantificación multibanda sobre una pluralidad de bandas de frecuencias de la señal de audio de entrada, para las cuales se dispone de información de ganancia de banda separada. Por ejemplo, el calculador de error de cuantificación 110 está configurado para determinar el error de cuantificación multibanda por la primera banda de frecuencias y la segunda banda de frecuencias utilizando la información 112, 114. Por consiguiente, el calculador de error de cuantificación 110 está configurado para proporcionar la información 116 que describe el error de cuantificación multibanda al proveedor de flujo de audio 120. El proveedor de flujo de audio 120 está configurado para recibir también una información 122 que describe la primera banda de frecuencias y una información 124 que describe la segunda banda de frecuencias. Además, el proveedor de flujo de audio 120 está configurado para proporcionar un flujo de audio 126, de modo que el flujo de audio 126 comprende una representación de la información 116 y también una representación del contenido de audio de la primera banda de frecuencias y de la segunda banda de frecuencias.

[0019] Por consiguiente, el codificador 100 proporciona un flujo de audio 126, que comprende un contenido de información, que permite una decodificación eficiente del contenido de audio de la banda de frecuencias utilizando un relleno de ruido. En particular, el flujo de audio 126 proporcionado por el codificador lleva aparejado un buen compromiso entre velocidad binaria y flexibilidad de decodificación de relleno de ruido.

1.2 Codificador según la Fig. 2

1.2.1 Descripción general del codificador

[0020] En lo que viene a continuación, se describirá un codificador de audio mejorado, que puede utilizarse en relación con realizaciones de la invención, que está basado en el codificador de audio descrito en la Norma Internacional ISO/IEC 14496-3: 2005(E), Tecnología de la Información – Codificación de Objetos Audiovisuales – Parte 3: Audio, Subparte 4: Codificación General de Audio (GA) – AAC, Twin VQ, BSAC.

[0021] El codificador de audio 200 según la Fig. 2 está basado específicamente en el codificador de audio descrito en la norma ISO/IEC 14496-3: 2005(E), Parte 3: Audio, Subparte 4, Sección 4.1. Sin embargo, el codificador de audio 200 no necesita implementar la funcionalidad exacta del codificador de audio de la norma ISO/IEC 14494-3: 2005(E).

[0022] El codificador de audio 200 puede, por ejemplo, estar configurado para recibir una señal de tiempo de entrada 210 y para proporcionar, basándose en ello, un flujo de audio codificado 212. Un recorrido de procesamiento de señal puede comprender un reductor de muestreo opcional 220, un controlador de ganancia de AAC opcional 222, un banco de filtros de conmutación de bloques 224, un procesamiento de señal opcional 226, un codificador de AAC extendida 228 y un formateador de cabida útil de tren de bits 230. Sin embargo, el codificador 200 comprende típicamente un modelo psicoacústico 240.

[0023] En un caso muy simple, el codificador 200 sólo comprende la conmutación de bloques/banco de filtros 224, el codificador de AAC extendida 228, el formateador de cabida útil de tren de bits 230 y el modelo psicoacústico 240, mientras que los otros componentes (en particular, los componentes 220, 222, 226) deberían considerarse meramente opcionales.

[0024] En un caso simple, la conmutación de bloques/banco de filtros 224 recibe la señal de tiempo de entrada 210 (opcionalmente con un muestreo reducido por el reductor de muestreo 220, y opcionalmente con ganancia cambiada de escala por el controlador de ganancia de AAC 222), y proporciona, basándose en ello, una representación en el dominio de la frecuencia 224a. La representación en el dominio de la frecuencia 224a puede, por ejemplo, comprender una información que describe intensidades (por ejemplo, amplitudes o energías) de segmentos espectrales de la señal de tiempo de entrada 210. Por ejemplo, la conmutación de bloques/banco de filtros 224 puede estar configurada para realizar una transformada discreta del coseno modificada (MDCT) para obtener los valores en el dominio de la frecuencia de la señal de tiempo de entrada 210. La representación en el dominio de la frecuencia 224a puede ser dividida lógicamente en diferentes bandas de frecuencias, que también se designan como “bandas de factores de escala”. Por ejemplo, se supone que la conmutación de bloques/banco de filtros 224 proporciona valores espectrales (también designados como valores de segmentos de frecuencia) para un gran número de segmentos de frecuencia diferentes. El número de segmentos de frecuencia se determina, entre otras cosas, por la longitud de una ventana introducida en el banco de filtros 224, y también dependiendo de la velocidad de muestreo (y la velocidad binaria). Sin embargo, las bandas de frecuencias o las bandas de factores de escala definen subconjuntos de los valores espectrales proporcionados por la conmutación de bloques/bando de filtros. Los detalles respecto a la definición de las bandas de factores de escala resultan conocidos por los expertos en la materia, y también se describen en la norma ISO/IEC 14496-3: 2005(E), Parte 3, Subparte 4.

[0025] El codificador de AAC extendida 228 recibe los valores espectrales 224a proporcionados por la conmutación de bloques/banco de filtros 224 basándose en la señal de tiempo de entrada 210 (o una versión preprocesada de la misma) como información de entrada 228a. Como puede apreciarse a partir de la Fig. 2, la información de entrada 228a del codificador de AAC extendida 228 puede obtenerse de los valores espectrales 224a utilizando una o más de las etapas de procesamiento del procesamiento espectral opcional 226. Para los detalles respecto a las etapas de preprocesamiento opcional del procesamiento espectral 226, se hace referencia a la norma ISO/IEC 14496-3: 2005(E), y a más normas a las que se hace referencia en la misma.

[0026] El codificador de AAC extendida 228 está configurado para recibir la información de entrada 228a en forma de valores espectrales para una pluralidad de segmentos espectrales y para proporcionar, basándose en ello, una representación cuantificada y codificada sin ruido 228b del espectro. Para este propósito, el codificador de AAC extendida 228 puede, por ejemplo, utilizar información obtenida de la señal de audio de entrada 210 (o una versión preprocesada de la misma) utilizando el modelo psicoaústico 240. En términos generales, el codificador de AAC extendida 228 puede utilizar una información proporcionada por el modelo psicoacústico 240 para decidir qué exactitud debería aplicarse para la codificación de diferentes bandas de frecuencias (o banda de factores de escala) de la información de entrada espectral 228a. Así, el codificador de AAC extendida 228 puede adaptar en general su exactitud de cuantificación para diferentes bandas de frecuencias a las características específicas de la señal de tiempo de entrada 210, y también el número de bits disponibles. Así, el codificador de AAC extendida puede, por ejemplo, ajustar sus exactitudes de cuantificación, de modo que la información que representa el espectro cuantificado y codificado sin ruido comprende una velocidad binaria (o velocidad binaria media) apropiada.

[0027] El formateador de cabida útil de tren de bits 230 está configurado para incluir la información 228b que representa los espectros cuantificados y codificados sin ruido en el flujo de audio codificado 212 según una sintaxis predeterminada.

[0028] Para detalles adicionales respecto a la funcionalidad de los componentes del codificador descritos en este documento, se hace referencia a la norma ISO/IEC 14496-3: 2005(E) (incluyendo el anexo 4.B de la misma), y también a la norma ISO/IEC 13818-7: 2003.

[0029] Además, se hace referencia a la norma ISO/IEC 13818-7: 2005, Subcláusulas C1 a C9.

[0030] Asimismo, se hace referencia específica respecto a la terminología a la norma ISO/IEC 14496-3: 2005(E), Parte 3: Audio, Subparte 1: Principal.

[0031] Además, se hace referencia específica a la norma ISO/IEC 14496-3: 2005(E), Parte 3: Audio, Subparte 4: Codificación General de Audio (GA) – AAC, Twin VQ, BSAC.

1.2.2 Detalles del codificador

[0032] En lo que viene a continuación, se describirán detalles respecto al codificador haciendo referencia a las Figs. 3a, 3b, 4a y 4b.

[0033] Las Figs. 3a y 3b muestran un diagrama esquemático de bloques de un codificador de AAC extendida que puede utilizarse en relación con realizaciones de la invención. El decodificador de AAC extendida se designa con 228 y puede ocupar el lugar del codificador de AAC extendida 228 de la Fig. 2. El codificador de AAC extendida 228 está configurado para recibir, como información de entrada 228a, un vector de magnitudes de líneas espectrales, en el que el vector de líneas espectrales a veces está designado con mdct_line (0..1023). El codificador de AAC extendida 228 también recibe una información de umbral de códec 228c, que describe una energía de error admisible máximo en un nivel de MDCT. La información de umbral de códec 228c se proporciona típicamente de manera individual para diferentes bandas de factores de escala y se genera utilizando el modelo psicoacústico 240. La información de umbral de códec 228 a veces se designa con xmin (sb), en la que el parámetro sb indica la dependencia de la banda de factores de escala. El codificador de AAC extendida 228 también recibe una información de número de bits 228d, que describe un número de bits disponibles para codificar el espectro representado por el vector 228a de magnitudes de valores espectrales. Por ejemplo, la información de número de bits 228d puede comprender una información de bits medios (designada con mean_bits) y una información de bits adicionales (designada con more_bits. El codificador de AAC extendida 228 también está configurado para recibir una información de banda de factores de escala 228e, que describe, por ejemplo, un número y anchura de las bandas de factores de escala.

[0034] El codificador de AAC extendido comprende un cuantificador de valores espectrales 310, que está configurado para proporcionar un vector 312 de valores cuantificados de líneas espectrales, que también se designa con x_quant (0..1023). El cuantificador de valores espectrales 310, que incluye un cambio de escala, también está configurado para proporcionar una información de factor de escala 314, que puede representar un factor de escala por cada banda de factores de escala y también una información de factor de escala común. Además, el cuantificador de valores espectrales 310 puede estar configurado para proporcionar una información de utilización de bits 316, que puede describir un número de bits utilizados para cuantificar el vector 228a de magnitudes de valores espectrales. Efectivamente, el cuantificador de valores espectrales 310 está configurado para cuantificar diferentes valores espectrales del vector 228a con diferentes exactitudes dependiendo de la relevancia psicoacústica de los diferentes valores espectrales. Para este propósito, el cuantificador de valores espectrales 210 cambia de escala de los valores espectrales del vector 228a utilizando diferentes factores de escala dependientes de la banda de factores de escala y cuantifica los valores espectrales cambiados de escala resultantes. Típicamente, los valores espectrales asociados con bandas de factores de escala importantes psicoacústicamente serán cambiados de escala con grandes factores de escala, de modo que los valores espectrales cambiados de escala de bandas de factores de escala importantes psicoacústicamente abarquen un gran intervalo de valores. En cambio, los valores espectrales de bandas de factores de escala menos importantes psicoacústicamente son cambiados de escala con factores de escala más pequeños, de modo que los valores espectrales cambiados de escala de las bandas de factores de escala menos importantes psicoacústicamente abarcan sólo un intervalo más de valores más pequeño. Los valores espectrales cambiados de escala se cuantifican entonces, por ejemplo, a un valor entero. En esta cuantificación, muchos de los valores espectrales cambiados de escala de las bandas de factores de escala menos importantes psicoacústicamente se cuantifican a cero, porque los valores espectrales de las bandas de factores de escala menos importantes psicoacústicamente se cambian de escala sólo con un factor de escala pequeño.

[0035] Como resultado, puede decirse que los valores espectrales de las bandas de factores de escala más relevantes psicoacústicamente se cuantifican con elevada exactitud (porque las líneas espectrales cambiadas de escala de dichas bandas de factores de escala más relevantes abarcan un gran intervalo de valores y, por lo tanto, muchas etapas de cuantificación), mientras que los valores espectrales de las bandas de factores de escala menos importantes psicoacústicamente se cuantifican con exactitud de cuantificación más baja (porque los valores espectrales cambiados de escala de las bandas de factores de escala menos importantes abarcan un intervalo de valores más pequeño y, por lo tanto, se cuantifican a etapas de cuantificación menos diferentes).

[0036] El cuantificador de valores espectrales 310 está configurado típicamente para determinar factores de

cambio de escala apropiados utilizando el umbral de códec 228c y la información de número de bits 228d. Típicamente, el cuantificador de valores espectrales 310 también está configurado para determinar los factores de escala apropiados por sí mismo. Los detalles respecto a una posible implementación del cuantificador de valores espectrales 310 se describen en la norma ISO/IEC 14496-3: 2001, Capítulo 4.B.10. Además, la implementación del cuantificador de valores espectrales resulta bien conocida para un experto en la materia de la codificación MPEG4.

[0037] El codificador de AAC extendida 228 también comprende un calculador de error de cuantificación multibanda 330, que está configurado para recibir, por ejemplo, el vector 228a de magnitudes de valores espectrales, el vector 312 de valores cuantificados de líneas espectrales y la información de factor de escala 314. El calculador de error de cuantificación multibanda 330 está configurado, por ejemplo, para determinar una desviación entre una versión cambiada de escala no cuantificada de los valores espectrales del vector 228a (por ejemplo, cambada de escala utilizando una operación de cambio de escala no lineal y un factor de escala) y una versión cambiada de escala y cuantificada (por ejemplo, cambiada de escala utilizando una operación de cambio de escala no lineal y un factor de escala, y cuantificada utilizando una operación de redondeo a “entero”) de los valores espectrales. Además, el calculador de error de cuantificación multibanda 330 puede estar configurado para calcular un error medio de cuantificación sobre una pluralidad de bandas de factores de escala. Cabe destacar que el calculador de error de cuantificación multibanda 330 calcula preferentemente el error de cuantificación multibanda en un dominio cuantificado (más exactamente en un dominio cambiado de escala psicoacústicamente), de modo que se pone de relieve el peso de un error de cuantificación en bandas de factores de escala relevantes psicoacústicamente cuando se compara con un error de cuantificación en bandas de factores de escala menos relevantes psicoacústicamente. Los detalles respecto a la operación del calculador de error de cuantificación multibanda se describirán posteriormente haciendo referencia a las Figs. 4a y 4b.

[0038] El codificador de AAC extendida 328 también comprende un adaptador de factor de escala 340, que está configurado para recibir el vector 312 de valores cuantificados, la información de factor de escala 314 y también la información de error de cuantificación multibanda 332, proporcionada por el calculador de error de cuantificación multibanda 340. El adaptador de factor de escala 340 está configurado para identificar bandas de factores de escala, que se “cuantifican a cero”, es decir, bandas de factores de escala para las cuales todos los valores espectrales (o líneas espectrales) se cuantifican a cero. Para tales bandas de factores de escala cuantificadas enteramente a cero, el adaptador de factor de escala 340 adapta el factor de escala respectivo. Por ejemplo, el adaptador de factor de escala 340 puede establecer el factor de escala de una banda de factores de escala cuantificada enteramente a cero en un valor que representa una relación entre una energía residual (antes de la cuantificación) de la banda de factores de escala respectiva y una energía del error de cuantificación multibanda 334. Por consiguiente, el adaptador de factor de escala 340 proporciona factores de escala adaptados 342. Cabe destacar que tanto los factores de escala proporcionados por el cuantificador de valores espectrales 310 como los factores de escala adaptados proporcionados por el adaptador de factor de escala se designan con “factor de escala (sb)”, “scf[band]”, “sf[g][sfb]”, “scf[g][sfb]” en la bibliografía y también dentro de esta solicitud. Los detalles respecto al funcionamiento del adaptador de factor de escala 340 se describirán posteriormente haciendo referencia a las Figs. 4a y 4b.

[0039] El codificador de AAC extendida 228 también comprende una codificación sin ruido 350, que se explica, por ejemplo, en la norma ISO/IEC 14496-3: 2001, Capítulo 4.B.11. En resumen, la codificación sin ruido 350 recibe el vector de valores cuantificados de líneas espectrales (también designado como “valores cuantificados de los espectros”) 312, la representación entera 342 de los factores de escala (o bien proporcionada por el cuantificador de valores espectrales 310, o bien adaptada por el adaptador de factor de escala 340), y también un parámetro de relleno de ruido 332 (por ejemplo, en forma de una información de nivel de ruido) proporcionado por el calculador de error de cuantificación multibanda 330.

[0040] La codificación sin ruido 350 comprende una codificación de coeficientes espectrales 350a para codificar los valores cuantificados 312 de las líneas espectrales, y para proporcionar valores cuantificados y codificados 352 de las líneas espectrales. Los detalles respecto a la codificación de coeficientes espectrales se describen, por ejemplo, en las secciones 4.B.11.2, 4.B.11.3, 4.B.11.4 y 4.B.11.6 de la norma ISO/IEC 14496-3: 2001. La codificación sin ruido 350 también comprende una codificación de factor de escala 350b para codificar la representación entera 342 del factor de escala para obtener una información de factor de escala codificado 354. La codificación sin ruido 350 también comprende una codificación de parámetros de relleno de ruido 350c para codificar el uno o más parámetros de relleno de ruido 332, para obtener uno o más parámetros de relleno de ruido codificados 356. En consecuencia, el codificador de AAC extendida proporciona una información que describe los espectros cuantificados como codificados sin ruido, en la que esta información comprende valores cuantificados y codificados de las líneas espectrales, información de factor de escala codificado e información de parámetros de relleno de ruido codificados.

[0041] En lo que viene a continuación, se describirá la funcionalidad del calculador de error de cuantificación

multibanda 330 y del adaptador de factor de escala 340, que son componentes clave del codificador de AAC extendida inventivo 228, haciendo referencia a las Figs. 4a y 4b. Para este propósito, la Fig. 4a muestra un listado de programa de un algoritmo llevado a cabo por el calculador de error de cuantificación multibanda 330 y el adaptador de factor de escala 340.

[0042] Una primera parte del algoritmo, representada por las líneas 1 a 12 del pseudocódigo de la Fig. 4a, comprende un cálculo de un error medio de cuantificación, que se lleva a cabo por el calculador de error de cuantificación multibanda 330. El cálculo del error medio de cuantificación se lleva a cabo, por ejemplo, por todas las bandas de factores de escala, excepto aquellas que se cuantifican a cero. Si una banda de factores de escala está cuantificada enteramente a cero (es decir, todas las líneas espectrales de la banda de factores de escala se cuantifican a cero), dicha banda de factores de escala se omite para el cálculo del error medio de cuantificación. Si, sin embargo, una banda de factores de escala no se cuantifica enteramente a cero (es decir, comprende al menos una línea espectral, que no se cuantifica a cero), todas las líneas espectrales de dicha banda de factores de escala se consideran para el cálculo del error medio de cuantificación. El error medio de cuantificación se calcula en un dominio cuantificado (o, más exactamente, en un dominio cambiado de escala). El cálculo de una contribución al error promedio puede apreciarse en la línea 7 del pseudocódigo de la Fig. 4a. En particular, la línea 7 muestra la contribución de una sola línea espectral al error promedio, en la que el promediado se lleva a cabo por todas las líneas espectrales (en la que nLines indica el número de líneas consideradas totales).

[0043] Como puede apreciarse en la línea 7 del pseudocódigo, la contribución de una línea espectral al error promedio es el valor absoluto (operador “fabs”) de una diferencia entre un valor de magnitud de línea espectral no cuantificada cambiada de escala y un valor de magnitud de línea espectral cuantificada cambiada de escala. En el valor de magnitud de línea espectral no cuantificada cambiada de escala, la “línea” de valor de magnitud (que puede ser igual a mdct_line) se cambia de escala de manera no lineal utilizando una función de potencia (pow(line, 0.75) = $line^{0,75}$) y utilizando un factor de escala (por ejemplo, un factor de escala 314 proporcionado por el cuantificador de valores espectrales 310). En el cálculo del valor de magnitud de línea espectral cuantificada cambiada de escala, la “línea” de valor de magnitud de línea espectral puede cambiarse de escala de manera no lineal utilizando las funciones de potencia mencionadas anteriormente y cambiarse de escala utilizando el factor de escala mencionado anteriormente. El resultado de este cambio de escala no lineal y lineal puede cuantificarse utilizando un operador entero (“INT”). Utilizando el cálculo tal como se indica en la línea 7 del pseudocódigo, se considera el diferente impacto de la cuantificación sobre las bandas de frecuencias más importantes psicoacústicamente y las menos importantes psicoacústicamente.

[0044] Después del cálculo del error de cuantificación (promedio) multibanda (avgError), el error de cuantificación promedio puede cuantificarse opcionalmente, tal como se muestra en las líneas 13 y 14 del pseudocódigo. Cabe destacar que la cuantificación del error de cuantificación multibanda tal como aquí se muestra está adaptado específicamente al intervalo esperado de valores y las características estadísticas del error de cuantificación, de modo que el error de cuantificación puede representarse de manera eficiente en cuanto a bits. Sin embargo, pueden aplicarse otras cuantificaciones del error de cuantificación multibanda.

[0045] Una tercera parte del algoritmo, que se representa en las líneas 15 a 25, puede ser ejecutada por el adaptador de factor de escala 340. La tercera parte del algoritmo sirve para establecer factores de escala de bandas de frecuencias de factor de escala, que han sido cuantificadas enteramente a cero, en un valor bien definido, lo cual permite un relleno de ruido sencillo, lo cual lleva aparejada una buena impresión de audición. La tercera parte del algoritmo comprende opcionalmente una cuantificación inversa del nivel de ruido (por ejemplo, representada por el error de cuantificación multibanda 332). La tercera parte del algoritmo también comprende un cálculo de un valor de factor de escala de reemplazo para las bandas de factores de escala cuantificadas a cero (mientras que los factores de escala de las bandas de factores de escala no cuantificadas a cero se dejarán sin cambios). Por ejemplo, el valor de factor de escala de reemplazo para una cierta banda de factores de escala (“band”) se calcula utilizando la ecuación mostrada en la línea 20 del algoritmo de la Fig. 4a. En esta ecuación, “(INT)” representa un operador entero, “2.f” representa el número 2 en una representación en coma flotante, “log” designa un operador logarítmico, “energy” designa una energía de la banda de factores de escala en consideración (antes de la cuantificación), “(float)” designa un operador en coma flotante, “sfbWidth” designa una anchura de la cierta banda de factores de escala en términos de líneas espectrales (o segmentos espectrales), y “noiseVal” designa un valor de ruido que describe el error de cuantificación multibanda. En consecuencia, el factor de escala de reemplazo describe una relación entre una energía promedio por segmento de energía (energy/sfbWidth) de las ciertas bandas de factores de escala en consideración, y una energía ($noiseVal^2$) del error de cuantificación multibanda.

1.2.3 Conclusión del codificador

[0046] Esta memoria descriptiva describe un codificador que tiene un nuevo tipo de cálculo de nivel de ruido. El nivel de ruido se calcula en el dominio cuantificado basándose en el error de cuantificación promedio.

[0047] Calcular el error de cuantificación en el dominio cuantificado lleva aparejadas ventajas significativas, por ejemplo, porque se considera la relevancia psicoacústica de diferentes bandas de frecuencias (bandas de factores de escala). El error de cuantificación por línea (es decir, por línea espectral, o segmento espectral) en el dominio cuantificado está típicamente en el intervalo [-0,5; 0,5] (nivel de cuantificación 1) con un error absoluto promedio de 0,25 (para valores de entrada de distribución normal que normalmente son mayores que 1). Utilizando un codificador, que proporciona información acerca de un error de cuantificación multibanda, las ventajas del relleno de ruido en el dominio cuantificado pueden aprovecharse en un codificador, tal como se describirá posteriormente.

[0048] El cálculo del nivel de ruido y la detección de sustitución de ruido en el codificador puede comprender las siguientes etapas:

• Detectar y marcar las bandas espectrales que pueden reproducirse perceptivamente equivalentes en el decodificador por sustitución de ruido. Por ejemplo, para este propósito puede comprobarse una tonalidad o una medición de planicidad espectral;
• Calcular y cuantificar el error medio de cuantificación (que puede calcularse por todas las bandas de factores de escala no cuantificadas a cero); y
• Calcular el factor de escala (scf) para una banda cuantificada a cero de modo que el ruido introducido (por el decodificador) coincida con la energía original.

[0049] Una cuantificación de nivel de ruido apropiada puede ayudar a producir el número de bits requeridos para transportar la información que describe el error de cuantificación multibanda. Por ejemplo, el nivel de ruido puede cuantificarse en 8 niveles de cuantificación en el dominio logarítmico, teniendo en cuenta la percepción humana de la sonoridad. Por ejemplo, puede utilizarse el algoritmo mostrado en la Fig. 4b, en el que “(INT)” designa un operador entero, en el que “LD” designa una operación logarítmica para una base 2, y en el que “meanLineError” designa un error de cuantificación por línea de frecuencia. “min(.,.)” designa un operador de valor mínimo, y “max(.,.)” designa un operador de valor máximo.

2. Decodificador

2.1. Decodificador según la Fig. 5

[0050] La Fig. 5 muestra un diagrama esquemático de bloques de un decodificador según una realización de la invención. El decodificador 500 está configurado para recibir una información de audio codificada, por ejemplo, en forma de un flujo de audio codificado 510, y para proporcionar, basándose en ello, una representación decodificada de la señal de audio, por ejemplo, basándose en componentes espectrales 522 de una primera banda de frecuencias y componentes espectrales 524 de una segunda banda de frecuencias. El decodificador 500 comprende un rellenador de ruido 520, que está configurado para recibir una representación 522 de componentes espectrales de una primera banda de frecuencias, a la cual se asocia información de ganancia de primera banda de frecuencias, y una representación 524 de componentes espectrales de una segunda banda de frecuencias, a la cual se asocia información de ganancia de segunda banda de frecuencias. Además, el rellenador de ruido 520 está configurado para recibir una representación 526 de un valor de intensidad de ruido multibanda. Además, el rellenador de ruido está configurado para introducir ruido en componentes espectrales (por ejemplo, en valores de líneas espectrales o valores de segmentos espectrales) de una pluralidad de bandas de frecuencias a las cuales se asocia información de ganancia de banda de frecuencias separada (por ejemplo, en forma de factores de escala) basándose en el valor de intensidad de ruido multibanda común 526. Por ejemplo, el rellenador de ruido 520 puede estar configurado para introducir ruido en las componentes espectrales 522 de la primera banda de frecuencias para obtener las componentes espectrales afectadas por ruido 512 de la primera banda de frecuencias, y también para introducir ruido en las componentes espectrales 524 de la segunda banda de frecuencias para obtener las componentes espectrales afectadas por ruido 514 de la segunda banda de frecuencias.

[0051] Así, comenzando desde una línea de comienzo de relleno de ruido, el decodificador está configurado para reemplazar cada línea espectral cuantificada a cero por un valor de reemplazo, que es un valor de ruido indicado, cuya magnitud se determina mediante el valor de intensidad de ruido multibanda, con un signo aleatorio, para realizar el relleno de ruido en un dominio cuantificado.

[0052] El decodificador está configurado para cambiar de escala el valor de reemplazo con un factor de escala transmitido para una banda de factores de escala real, para realizar un cambio de escala individual del ruido introducido en distintas bandas de frecuencias basándose en la información de ganancia de banda de frecuencia.

[0053] Aplicando el ruido descrito por un solo valor de intensidad de ruido multibanda 526 a componentes espectrales de diferentes bandas de frecuencias a las que se asocia información de ganancia de diferentes bandas de frecuencias, puede introducirse ruido en las diferentes bandas de frecuencias de manera muy afinada, teniendo en cuenta la diferente relevancia psicoacústica de una banda de frecuencias diferente, que se expresa por la información de ganancia de banda de frecuencias. Así, el decodificador 500 es capaz de llevar a cabo un relleno de ruido afinado basándose en una información colateral de relleno de ruido muy pequeña (eficiente en cuanto a bits).

[0054] En otras palabras, el decodificador se basa en el hallazgo de que puede aplicarse un único valor de intensidad de ruido multibanda para un relleno de ruido con buenos resultados si se asocia información de ganancia de banda de frecuencia separada a las distintas bandas de frecuencia. Por consiguiente, es posible un cambio de escala individual del ruido introducido en las distintas bandas de frecuencia basándose en la información de ganancia de banda de frecuencia.

2.2 Decodificador según la Fig. 6

2.2.1 Descripción general del decodificador

[0055] La Fig. 6 muestra un diagrama esquemático de bloques de un decodificador 600 en el que puede implementarse la invención.

[0056] El decodificador 600 es similar al decodificador descrito en la norma ISO/IEC 14496.3: 2005 (E), de modo que se hace referencia a esta norma internacional. El decodificador 600 está configurado para recibir un flujo de audio codificado 610 y para proporcionar, basándose en ello, señales de tiempo de salida 612. El flujo de audio codificado puede comprender alguna o toda la información descrita en la norma ISO/IEC 14496.3: 2005 (E), y comprende adicionalmente información que describe un valor de intensidad de ruido multibanda. El decodificador 600 comprende además un deformateador de cabida útil de tren de bits 620, que está configurado para extraer del flujo de audio codificado 610 una pluralidad de parámetros de audio codificados, algunos de los cuales se explicarán en detalle en lo que viene a continuación. El decodificador 600 comprende además un decodificador de “codificación de audio avanzada” (AAC) extendida 630, cuya funcionalidad se describirá en detalle, haciendo referencia a las Figs. 7a, 7b, 8a a 8c, 9, 10a, 10b, 11, 12, 13a y 13b. El decodificador de AAC extendida 630 está configurado para recibir una información de entrada 630a, que comprende, por ejemplo, una información de líneas espectrales cuantificadas y codificadas, una información de factor de escala codificado y una información de parámetros de relleno de ruido codificados. Por ejemplo, la información de entrada 630a del codificador de AAC extendida 630 puede ser idéntica a la información de salida 228b proporcionada por el codificador de AAC extendida 220a descrita con referencia a la Fig. 2.

[0057] El decodificador de AAC extendida 630 puede estar configurado para proporcionar, basándose en la información de entrada 630a, una representación 630b de un espectro cambiado de escala y cuantificado inversamente, por ejemplo, en forma de valores de líneas espectrales cambiadas de escala, cuantificadas inversamente para una pluralidad de segmentos de frecuencia (por ejemplo, para 1024 segmentos de frecuencia).

[0058] Opcionalmente, el decodificador 600 puede comprender decodificadores de espectro adicionales, como, por ejemplo, un decodificador de espectro TwinVQ y/o un decodificador de espectro BSAC, que pueden utilizarse alternativamente al decodificador de espectro de AAC extendida 630 en algunos casos.

[0059] El decodificador 600 puede comprender opcionalmente un procesamiento de espectro 640, que está configurado para procesar la información de salida 630b del decodificador de AAC extendida 630 con el fin de obtener una información de entrada 640a de una conmutación de bloques/banco de filtros 640. El procesamiento espectral opcional 630 puede comprender una o más, o incluso todas las funcionalidades M/S, PNS, predicción, intensidad, predicción a largo plazo, acoplamiento conmutado dependientemente, TNS, acoplamiento conmutado dependientemente, funcionalidades que se describen en detalle en la norma ISO/IEC 14493.3: 2005 (E) y los documentos a los que se hace referencia en la misma. Si, sin embargo, se omite el procesamiento espectral 630, la información de salida 630b del decodificador de AAC extendida 630 puede servir directamente como información de entrada 640a de la conmutación de bloques/banco de filtros 640. Así, el decodificador de AAC extendida 630 puede proporcionar, como la información de salida 630b, espectros cambiados de escala y cuantificados inversamente. La

conmutación de bloques/banco de filtros 640 utiliza, como la información de entrada 640a, los espectros cuantificados inversamente (opcionalmente preprocesados) y proporciona, basándose en ello, una o más señales de audio reconstruidas en el dominio del tiempo como información de salida 640b. El banco de filtros/conmutación de bloques puede, por ejemplo, estar configurado para aplicar la inversa de la correspondencia de frecuencia que se llevó a cabo en el codificador (por ejemplo, en la conmutación de bloques/banco de filtros 224). Por ejemplo, puede utilizarse una transformada inversa del coseno discreta modificada (IMDCT) por el banco de filtros. Por ejemplo, la IMDCT puede estar configurada para soportar o bien un conjunto de 120, 128, 480, 512, 960 o 1024, o bien cuatro conjuntos de 32 o 256 coeficientes espectrales.

[0060] Para obtener detalles, se hace referencia, por ejemplo, a la Norma Internacional ISO/IEC 14496-3: 2005 (E). El decodificador 600 puede comprender además opcionalmente un control de ganancia de AAC 650, un decodificador de SBR 652 y un acoplamiento conmutado independientemente 654, para obtener la señal de tiempo de salida 612 a partir de la señal de salida 640b de la conmutación de bloques/banco de filtros 640.

[0061] Sin embargo, la señal de salida 640b de la conmutación de bloques/banco de filtros 640 también puede servir como la señal de tiempo de salida 612 en ausencia de la funcionalidad 650, 652, 654.

2.2.2 Detalles del decodificador de AAC extendida

[0062] En lo que viene a continuación, se describirán detalles respecto al decodificador de AAC extendida, haciendo referencia a las Figs. 7a y 7b. Las Figs. 7a y 7b muestran un diagrama esquemático de bloques del decodificador de AAC de la Fig. 6 en combinación con el deformateador de cabida útil de tren de bits 620 de la Fig. 6.

[0063] El deformateador de cabida útil de flujo de bits 620 recibe un flujo de audio decodificado 610, que puede comprender, por ejemplo, un flujo de datos de audio codificado que comprende un elemento de sintaxis con el título “ac_raw_data_block”, que es un bloque de datos sin procesar del codificador de audio. Sin embargo, el formateador de cabida útil de tren de bits 620 está configurado para proporcionar al decodificador de AAC extendida 630 un espectro cuantificado y codificado sin ruido o una representación, que comprende una información de líneas espectrales cuantificadas y codificadas aritméticamente 630aa (por ejemplo, designada como ac_spectral_data), una información de factor de escala 630ab (por ejemplo, designada como scale_factor_data) y una información de parámetros de relleno de ruido 630ac. La información de parámetros de relleno de ruido 630ac comprende, por ejemplo, un valor de compensación de ruido (designado como noise_offset) y un valor de nivel de ruido (designado como noise_level).

[0064] Con respecto al decodificador de AAC extendida, cabe destacar que el decodificador de AAC extendida 630 es muy similar al decodificador de AAC de la Norma Internacional ISO/IEC 19496–3: 2005 (E), de modo que se hace referencia a la descripción detallada en dicha Norma.

[0065] El decodificador de AAC extendida 630 comprende un decodificador de factor de escala 740 (también designado como herramienta de decodificación sin ruido de factor de escala) que está configurado para recibir la información de factor de escala 630ab y para proporcionar, basándose en ello, una representación entera decodificada 742 de los factores de escala (la cual también se designa como sf[g] [sfb] o scf[g] [sfb]). Con respecto al decodificador de factor de escala 740, se hace referencia a la norma ISO/IEC 14496–3: 2005, Capítulos 4.6.2 y 4.6.3. Cabe destacar que la representación entera decodificada 742 de los factores de escala refleja una exactitud de cuantificación con la cual se cuantifican diferentes bandas de frecuencias (también designadas como bandas de factores de escala) de una señal de audio. Los factores de escala más grandes indican que las bandas de factores de escala correspondientes se han cuantificado con una exactitud elevada, y los factores de escala más pequeños indican que las bandas de factores de escala correspondientes se han cuantificado con exactitud baja.

[0066] El decodificador de AAC extendida 630 también comprende un decodificador espectral 750 que está configurado para recibir la información de líneas espectrales cuantificadas y codificadas por entropía (por ejemplo, codificadas según Huffman o codificadas aritméticamente) 630aa y para proporcionar, basándose en ello, valores cuantificados 752 del uno o más espectros (por ejemplo, designados como x_ac_quant o x_quant). Con respecto al decodificador espectral, se hace referencia, por ejemplo, a la sección 4.6.3 de la Norma Internacional anteriormente mencionada. Sin embargo, se pueden aplicar naturalmente implementaciones alternativas del decodificador espectral. Por ejemplo, el decodificador Huffman de la norma ISO/IEC 14496–3: 2005 puede ser reemplazado por un decodificador aritmético si la información de líneas espectrales 630aa se codifica aritméticamente.

[0067] El decodificador de AAC extendida 630 comprende además un cuantificador inverso 760, que puede ser un cuantificador inverso no uniforme. Por ejemplo, el cuantificador inverso 760 puede proporcionar valores espectrales

cuantificados inversamente no cambiados de escala 762 (por ejemplo, designados con x_ac_invquant o x_invquant). Por ejemplo, el cuantificador inverso 760 puede comprender la funcionalidad descrita en la norma ISO/IEC 14496–3: 2005, Capítulo 4.6.2. Alternativamente, el cuantificador inverso 760 puede comprender la funcionalidad descrita con referencia a las Figs. 8a a 8c.

[0068] El decodificador de AAC extendida 630 también comprende un rellenador de ruido 770 (también designado como herramienta de relleno de ruido), el cual recibe la representación entera decodificada 742 de los factores de escala procedente del decodificador de factor de escala 740, los valores espectrales cuantificados inversamente sin cambiar de escala 762 procedentes del cuantificador inverso 760 y la información de parámetros de relleno de ruido 630ac procedente del deformateador de cabida útil de tren de bits 620. El rellenador de ruido está configurado para proporcionar, basándose en ello, la representación modificada (típicamente entera) 772 de los factores de escala, la cual también está designada en este documento con sf[g] [sfb] o scf[g] [sfb]. El rellenador de ruido 770 también está configurado para proporcionar valores espectrales cuantificados inversamente no cambiados de escala 774, también designados como x_ac_invquant o x_invquant basándose en su información de entrada. Los detalles con respecto a la funcionalidad del rellenador de ruido se describirán posteriormente, haciendo referencia a las Figs. 9, 10a, 10b, 11, 12, 13a y 13b.

[0069] El decodificador de AAC extendida 630 también comprende un medio de reajuste de escala 780, que está configurado para recibir la representación entera modificada de los factores de escala 772 y los valores espectrales cuantificados inversamente no cambiados de escala 774, y para proporcionar, basándose en ello, valores espectrales cuantificados inversamente cambiados de escala 782, que también pueden designarse como x_rescal, y los cuales pueden servir como la información de salida 630b del decodificador de AAC extendida 630. El medio de reajuste de escala 780 puede comprender, por ejemplo, la funcionalidad que se describe en la norma ISO/IEC 14496–3: 2005, Capítulo 4.6.2.3.3.

2.2.3. Cuantificador inverso

[0070] En lo que viene a continuación, se describirá la funcionalidad del cuantificador inverso 760, haciendo referencia a las Figs. 8a, 8b, y 8c. La Fig. 8a muestra una representación de una ecuación para obtener los valores espectrales cuantificados inversamente no cambiados de escala 762 a partir de los valores espectrales cuantificados 752. En las ecuaciones alternativas de la Fig. 8a, “sign(.)” designa un operador de signo, y “.” designa un operador de valor absoluto. La Fig. 8b muestra un pseudocódigo de programa que representa la funcionalidad del cuantificador inverso 760. Como puede apreciarse, se lleva a cabo la cuantificación inversa según la regla de correspondencia matemática mostrada en la Fig. 8a para todos los grupos de ventanas (designados por la variable en curso g), para todas las bandas de factores de escala (designadas por la variable en curso sfb), para todas las ventanas (designadas por el índice en curso win) y todas las líneas espectrales (o segmentos espectrales) (designadas por la variable en curso bin). La Fig. 8c muestra una representación de diagrama de flujo del algoritmo de la Fig. 8b. Para las bandas de factores de escala por debajo de una banda de factores de escala máxima predeterminada (designada con max_sfb), se obtienen los valores espectrales cuantificados inversamente no cambiados de escala como una función de los valores espectrales cuantificados no cambiados de escala. Se aplica una regla de cuantificación inversa no lineal.

2.2.4 Rellenador de ruido

2.2.4.1 Rellenador de ruido según las Figs. 9 a 12

[0071] La Fig. 9 muestra un diagrama de bloques esquemático de un rellenador de ruido 900. El rellenador de ruido 900 puede, por ejemplo, ocupar el lugar del rellenador de ruido 770 descrito con referencia a las Figs. 7A y 7B.

[0072] El rellenador de ruido 900 recibe la representación entera decodificada 742 de los factores de escala, que pueden considerarse como valores de ganancia de banda de frecuencias. El rellenador de ruido 900 también recibe los valores espectrales cuantificados inversamente no cambiados de escala 762. Además, el rellenador de ruido 900 recibe la información de parámetros de relleno de ruido 630ac, que comprende, por ejemplo, los parámetros de relleno de ruido noise_value y noise_offset. Además, el rellenador de ruido 900 proporciona la representación entera modificada 772 de los factores de escala y los valores espectrales cuantificados inversamente no cambiados de escala 774. El rellenador de ruido 900 comprende un detector de líneas espectrales cuantificadas a cero 910, que está configurado para determinar si una línea espectral (o segmento espectral) está cuantificada a cero (y posiblemente cumple requisitos de relleno de ruido adicionales). Para este propósito, el detector de líneas espectrales cuantificadas a cero 910 recibe directamente los espectros cuantificados inversamente no cambiados de escala 762 como información de entrada. El rellenador de ruido 900 comprende además un dispositivo de reemplazo selectivo de líneas espectrales 920, que está configurado para reemplazar selectivamente los valores espectrales de la información de

entrada 762 por los valores de reemplazo de líneas espectrales 922 dependiendo de la decisión del detector de líneas espectrales cuantificadas a cero 910. De ese modo, el detector de líneas espectrales cuantificadas a cero 910 indica que una cierta línea espectral de la información de entrada 762 debería ser reemplazada por un valor de reemplazo, entonces el dispositivo de reemplazo selectivo de líneas espectrales 920 reemplaza la cierta línea espectral con el valor de reemplazo de las líneas espectrales 922 para obtener la información de salida 774. Si no, el dispositivo de reemplazo selectivo de líneas espectrales 920 transmite el valor de la cierta línea espectral sin cambio para obtener la información de salida 774. El rellenador de ruido 900 también comprende un modificador selectivo de factores de escala 930, que está configurado para modificar selectivamente los factores de escala de la información de entrada 742. Por ejemplo, el modificador selectivo de factores de escala 930 está configurado para incrementar los factores de escala de las bandas de frecuencias de factores de escala, que han sido cuantificadas a cero por un valor predeterminado, que se designa como “noise_offset”. De ese modo en la información de salida 772, los factores de escala de las bandas de frecuencias cuantificadas a cero se incrementan en comparación con los valores de factores de escala correspondientes dentro de la información de entrada 742. En cambio, los valores de factores de escala correspondientes de las bandas de frecuencias de factores de escala, que no se cuantifican a cero, son idénticos en la información de entrada 742 y en la información de salida 772.

[0073] Para determinar si una banda de frecuencias de factores de escala está cuantificada a cero, el rellenador de ruido 900 también comprende un detector de bandas cuantificadas a cero 940, que está configurado para controlar el modificador selectivo de factores de escala 930 proporcionando una señal o un indicador de “permitir modificación de factores de escala” 942 basándose en la información de entrada 762. Por ejemplo, el detector de bandas cuantificadas a cero 940 puede proporcionar una señal o un indicador que indica la necesidad de un incremento de un factor de escala al modificador selectivo de factores de escala 930 si todos los segmentos de frecuencia (también designadas como segmentos espectrales) de una banda de factores de escala están cuantificados a cero.

[0074] Cabe destacar aquí que el modificador selectivo de factores de escala también puede adoptar la forma de un dispositivo de reemplazo selectivo de factores de escala, que está configurado para establecer los factores de escala de las bandas de factores de escala cuantificadas enteramente a cero en un valor predeterminado, independientemente de la información de entrada 742.

[0075] En lo que viene a continuación, se describirá un medio de reajuste de escala 950, que puede adoptar la función del medio de reajuste de escala 780. El medio de reajuste de escala 950 está configurado para recibir la representación entera modificada 772 de los factores de escala proporcionados por el rellenador de ruido y también para los valores espectrales cuantificados inversamente no cambiados de escala 774 proporcionados por el rellenador de ruido. El medio de reajuste de escala 950 comprende un computador de ganancia de factores de escala 960, que está configurado para recibir una representación entera del factor de escala por banda de factores de escala y para proporcionar un valor de ganancia por banda de factores de escala. Por ejemplo, el computador de ganancia de factores de escala 960 puede estar configurada para calcular un valor de ganancia 962 para una banda de frecuencias i–ésima basándose en una representación entera modificada 772 del factor de escala para la banda de factores de escala i-ésima. De ese modo, el computador de ganancia de los factores de escala 960 proporciona valores de ganancia individuales para las diferentes bandas de factores de escala. El medio de reajuste de escala 950 también comprende un multiplicador 970, que está configurado para recibir los valores de ganancia 962 y los valores espectrales cuantificados inversamente no cambiados de escala 774. Cabe destacar que cada uno de los valores espectrales cuantificados inversamente no cambiados de escala 774 está asociado con una banda de frecuencias de factores de escala (sfb). Por consiguiente, el multiplicador 970 está configurado para cambiar la escala de cada uno de los valores espectrales cuantificados inversamente no cambiados de escala 774 con un valor de ganancia correspondiente asociado con la misma banda de factores de escala. En otras palabras, todos los valores espectrales cuantificados inversamente no cambiados de escala 774 asociados con una banda de factores de escala dada se cambian de escala con el valor de ganancia asociado con la banda de factores de escala dada. Por consiguiente, los valores espectrales cuantificados inversamente no cambiados de escala asociados con bandas de factores de escala diferentes se cambian de escala típicamente con diferentes valores de ganancia asociados con diferentes bandas de factores de escala.

[0076] De este modo, los diferentes valores espectrales cuantificados inversamente no cambiados de escala se cambian de escala con diferentes valores de ganancia, dependiendo de a qué bandas de factores de escala están asociados.

Representación de pseudocódigo de programa

[0077] En lo que viene a continuación, se describirá la funcionalidad del rellenador de ruido 900 haciendo referencia a las Figs. 10A y 10B, las cuales muestran una representación de pseudocódigo de programa (Fig. 10A) y

una leyenda correspondiente (Fig. 10B). Los comentarios empiezan con "- -".

[0078] El algoritmo de relleno de ruido representado por el listado del pseudocódigo de programa de la Fig. 10 comprende una primera parte (líneas 1 a 8) para obtener un valor de ruido (noiseVal) a partir de una representación de nivel de ruido (noise_level). Además, se obtiene una compensación de ruido (noise_offset). La obtención del valor de ruido a partir del nivel de ruido comprende un cambio de escala no lineal, en el que el valor de ruido se calcula según:

$$noiseVal = 2^{((noise_level-14)/3)}.$$

[0079] Además, se lleva a cabo un desplazamiento de intervalo del valor de compensación de ruido de modo que el valor de compensación de ruido de intervalo desplazado puede adoptar valores positivos y negativos.

[0080] Una segunda parte del algoritmo (líneas 9 a 29) es responsable del reemplazo selectivo de los valores espectrales cuantificados inversamente no cambiados de escala con valores de reemplazo de líneas espectrales y de una modificación selectiva de los factores de escala. Tal como se puede apreciarse a partir del pseudocódigo de programa, el algoritmo puede ejecutarse para todos los grupos de ventanas disponibles (bucle "for" de las líneas 9 a 29). Además, se pueden procesar todas las bandas de factores de escala entre cero y una banda de factores de escala máxima (max_sfb) aun cuando el procesamiento puede ser diferente para diferentes bandas de factores de ajuste (bucle "for" entre las líneas 10 a 28). Un aspecto importante es el hecho de que se supone generalmente que una banda de factores de escala se cuantifica a cero a menos que se descubra que la banda de factores de escala no está cuantificada a cero (consúltese la línea 11). Sin embargo, la verificación de si una banda de factores de escala está cuantificada a cero o no se ejecuta sólo para las bandas de factores de escala, de lo cual una línea de frecuencia de comienzo (swb_offset[sfb]) está por encima de un índice de coeficiente espectral predeterminado (noiseFillingStartOffset). Una rutina condicional entre las líneas 13 y 24 se ejecuta sólo si un índice de los coeficientes espectrales más bajos de una banda de factores de escala sfb es mayor que la compensación de comienzo de relleno de ruido. En cambio, para cualquier banda de factores de escala para la cual un índice del coeficiente espectral más bajo (swb_offset[sfb]) es menor o igual que un valor predeterminado (noiseFillingStartOffset), se supone que las bandas no están cuantificadas a cero, independientemente de los valores reales de las líneas espectrales (véanse las líneas 24a, 24b y 24c).

[0081] Sin embargo, si el índice de los coeficientes espectrales más bajos de una cierta banda de factores de escala es mayor que el valor predeterminado (noiseFillingStartOffset), entonces se considera que la cierta banda de factores de escala está cuantificada a cero sólo si todas las líneas espectrales de la cierta banda de factores de escala están cuantificadas a cero (se restablece el indicador "band_quantized_to_zero" mediante el bucle "for" entre las líneas 15 y 22 si un solo segmento espectral de la banda de factores de escala no está cuantificado a cero.

[0082] En consecuencia, un factor de escala de una banda de factores de escala dada se modifica utilizando de la compensación de ruido si el indicador "band_quantized_to_zero", que se establece inicialmente por defecto (línea 11) no se borra durante la ejecución del código de programa entre las líneas 12 y 24. Tal como se menciona anteriormente, un restablecimiento del indicador sólo puede producirse para las bandas de factores de escala para las cuales un índice del coeficiente espectral más bajo está por encima del valor predeterminado (noiseFillingStartOffset). Además, el algoritmo de la Fig. 10A comprende un reemplazo de valores de líneas espectrales con valores de reemplazo de líneas espectrales si la línea espectral está cuantificada a cero (condición de la línea 16 y operación de reemplazo de la línea 17). Sin embargo, dicho reemplazo sólo se lleva a cabo para las bandas de factores de escala para las cuales un índice de los coeficientes espectrales más bajos está por encima del valor predeterminado (noiseFillingStartOffset). Para las bandas de frecuencias espectral más baja, se omite el reemplazo de los valores espectrales cuantificados a cero con los valores espectrales de reemplazo.

[0083] Además, cabe destacar que los valores de reemplazo podrían calcularse de una manera sencilla porque se añade un signo aleatorio o pseudoaleatorio al valor de ruido (noiseVal) calculado en la primera parte del algoritmo (consúltese la línea 17).

[0084] Cabe destacar que la Fig. 10B muestra una leyenda de los símbolos relevantes utilizados en el pseudocódigo de programa de la Fig. 10A para facilitar una mejor comprensión del pseudocódigo de programa.

[0085] Aspectos importantes de la funcionalidad del rellenador de ruido se ilustran en la Fig. 11. Tal como puede apreciarse, la funcionalidad del rellenador de ruido comprende opcionalmente calcular 1110 un valor de ruido basándose en el nivel de ruido. La funcionalidad del rellenador de ruido comprende también el reemplazo 1120 de valores de líneas espectrales de las líneas espectrales cuantificadas a cero con valores de reemplazo de líneas

espectrales dependiendo del valor de ruido para obtener valores de líneas espectrales reemplazados. Sin embargo, el reemplazo 1120 sólo se lleva a cabo para las bandas de factores de escala que tienen un coeficiente espectral más bajo por encima de un índice de coeficiente espectral predeterminado.

[0086] La funcionalidad del rellenador de ruido también comprende modificar 1130 un factor de escala de bandas dependiendo del valor de compensación de ruido, si, y sólo si la banda de factores de escala está cuantificada a cero. Sin embargo, la modificación 1130 se ejecuta de esa forma para las bandas de factores de escala que tienen un coeficiente espectral más bajo por encima del índice de coeficiente espectral predeterminado.

[0087] El rellenador de ruido también comprende una funcionalidad de dejar 1140 sin cambio a los factores de escala de bandas, independiente de si la banda de factores de escala está cuantificada a cero, para las bandas de factores de escala que tienen un coeficiente espectral más bajo por debajo del índice de coeficiente espectral predeterminado.

[0088] Además, el medio de reajuste de escala comprende una funcionalidad 1150 de aplicar los factores de escala de bandas no modificadas o modificadas (cualquiera que esté disponible) a los valores de líneas espectrales no reemplazados o reemplazados (cualquiera que esté disponible), para obtener espectros cambiados de escala y cuantificados inversamente.

[0089] La Fig. 12 muestra una representación esquemática del concepto descrito con referencia a las Figs. 10A, 10B y 11. En particular, las diferentes funcionalidades se representan dependiendo de un segmento de comienzo de bandas de factores de escala.

2.2.4.2 Rellenador de ruido según las Figs. 13A y 13B

[0090] Las Figs. 13A y 13B muestran listados de pseudocódigos de programa de algoritmos, que pueden llevarse a cabo en una implementación alternativa del rellenador de ruido 770. La Fig. 13A describe un algoritmo para obtener un valor de ruido (para uso dentro del rellenador de ruido) a partir de una información de nivel de ruido, que puede representarse por la información de parámetros de relleno de ruido 630ac.

[0091] Como el error medio de cuantificación es aproximadamente 0,25 en la mayoría de los casos, el intervalo de noiseVal [0, 0,5] es bastante grande y puede ser optimizado.

[0092] La Fig. 13B representa un algoritmo, que puede estar formado por el rellenador de ruido 770. El algoritmo de la Fig. 13B comprende una primera porción de determinación del valor de ruido (designado con noiseValue” o “noiseVal” - líneas 1 a 4). Una segunda porción del algoritmo comprende una modificación selectiva de un factor de escala (líneas 7 a 9) y un reemplazo selectivo de valores de líneas espectrales por valores de reemplazo de líneas espectrales (líneas 10 a 14).

[0093] Sin embargo, según el algoritmo de la Fig. 13B, el factor de escala (scf) se modifica utilizando la compensación de ruido (noise_offset) siempre y cuando una banda esté cuantificada a cero (véase la línea 7). En esta realización no se hacen diferencias entre las bandas de frecuencias más bajas y las bandas de frecuencias más altas.

[0094] Además, se introduce ruido en las líneas espectrales cuantificadas a cero sólo para las bandas de frecuencias más altas (si la línea está por encima de un cierto umbral predeterminado “noiseFillingStartOffset”).

2.2.5. Conclusión del decodificador

[0095] Para resumir, las realizaciones del decodificador según la presente invención comprenden las siguientes características:

• Comenzando desde una “línea de comienzo de relleno de ruido” (que puede ser una compensación fija o una línea que representa una frecuencia de comienzo) se debe reemplazar cada 0 con un valor de reemplazo; y

• el valor de reemplazo es el valor de ruido indicado (con un signo aleatorio) en el dominio cuantificado y luego se cambia la escala de este “valor de reemplazo” con el factor de escala “scf” transmitido para la banda de factores de escala real;

Además,

• los valores de reemplazo “aleatorios” también pueden obtenerse a partir de, por ejemplo, una distribución de ruido o un conjunto de valores alternativos ponderados con el nivel de ruido señalado.

3. Flujo de audio

3.1 Flujo de audio según las Figs. 14A y 14B

[0096] En lo que viene a continuación, se describirá un flujo de audio que puede utilizarse en relación con realizaciones de la invención. En lo que sigue, se describirá una denominada “cabida útil de tren de bits usac”. La “cabida útil de tren de bits usac” transporta información de cabida útil para representar uno o más canales individuales (cabida útil: “single_channel_element ()) y/o uno o más pares de canales (channel_pair_element ()) tal como puede apreciarse en la Fig. 14A. Una información de canal individual (single_channel_element ()) comprende, entre otra información opcional, un flujo de canal en el dominio de la frecuencia (fd_channel_stream), tal como puede apreciarse en la Figura 14B.

[0097] Una información de pares de canales (channel_pair_element) comprende, además de elementos adicionales, una pluralidad de, por ejemplo, dos flujos de canales en el dominio de la frecuencia (fd_channel_stream), tal como puede apreciarse en la Fig. 14C.

[0098] El contenido de datos de un flujo de canal en el dominio de la frecuencia puede, por ejemplo, depender de si se utiliza o no un relleno de ruido (lo cual se puede señalar en una porción de datos de señalización que no se muestra aquí). En lo que viene a continuación, se supondrá que se utiliza un relleno de ruido. En este caso, el flujo de canal en el dominio de la frecuencia comprende, por ejemplo, los elementos de datos mostrados en la Fig. 14D. Por ejemplo, puede estar presente una información de ganancia global (global_gain), tal como se define en la norma ISO/IEC 14496–3: 2005. Por otra parte, el flujo de canal en el dominio de la frecuencia puede comprender una información de compensación de ruido (noise_offset) y una información de nivel de ruido (noise_level), tal como se describe en este documento. Por ejemplo, la información de compensación de ruido puede codificarse utilizando 3 bits y la información de nivel de ruido puede codificarse utilizando, por ejemplo, 5 bits.

[0100] Además, el flujo de canal en el dominio de la frecuencia puede comprender una información de factor de escala codificado (un scale_factor_data ()) y datos espectrales codificados aritméticamente (AC_spectral_data ()), tal como se describen en este documento y también como se definen en la norma ISO/IEC 14496–3.

[0101] Opcionalmente, el flujo de canal en el dominio de la frecuencia también comprende datos de conformación de ruido temporal (tns_data) ()), tal como se define en la norma ISO/IEC 14496–3.

[0102] Naturalmente, el flujo de canal en el dominio de la frecuencia puede comprender otra información, si es necesario.

3.2 Flujo de audio según la Fig. 15

[0103] La Fig. 15 muestra una representación esquemática de la sintaxis de un flujo de canal que representa un canal individual (individual_channel_stream ()).

[0104] El flujo de canal individual puede comprender una información de ganancia global (global_gain) codificada utilizando, por ejemplo, 8 bits, información de compensación de ruido (noise_offset) codificada utilizando, por ejemplo, 5 bits y una información de nivel de ruido (noise_level) codificada utilizando, por ejemplo, 3 bits.

[0105] El flujo de canal individual comprende además datos de sección (section_data ()), datos de factor de escala (scale_factor ()) y datos espectrales (spectral_data ()).

[0106] Además, el flujo de canal individual puede comprender información opcional individual, tal como puede apreciarse en la Fig. 15.

3.3 Conclusión de flujo de audio

[0107] Para resumir, se describe un tren de bits en el que se utilizan los siguientes elementos de sintaxis de tren de bits.

• Un valor que indica una compensación de factores de escala de ruido para optimizar los bits necesarios para transmitir los factores de escala;
• un valor que indica el nivel de ruido; y/o
• un valor opcional para escoger entre diferentes formas para la sustitución de ruido (ruido distribuido uniforme en lugar de valores constantes o múltiples niveles discretos en lugar de sólo uno).

4. Conclusión

[0108] En la codificación de baja velocidad binaria, puede utilizarse relleno de ruido para dos propósitos:

• Una cuantificación gruesa de los valores espectrales en una codificación de audio de baja velocidad binaria podría conducir a espectros muy dispersos después de la cuantificación inversa, ya que muchas líneas espectrales podrían haber sido cuantificadas a cero. Los espectros de ocupación dispersa tendrán como resultado que la señal decodificada suene aguda o inestable (señales espurias). Reemplazando las líneas cuantificadas a cero con valores "pequeños" en el decodificador, es posible enmascarar o reducir estos artefactos muy obvios sin añadir nuevos artefactos de ruido obvios.
• Si en el espectro original hay partes de la señal similares a ruido, se puede reproducir en el decodificador una representación perceptivamente equivalente de estas partes de señal ruidosa basándose sólo en una información paramétrica pequeña, como la energía de la parte de señal ruidosa. La información paramétrica puede transmitirse con menos bits en comparación con el número de bits necesarios para transmitir la forma de onda codificada.

[0109] El esquema de codificación de relleno de ruido que se acaba de proponer descrito en este documento combina eficientemente los propósitos anteriores en una única aplicación.

[0110] Como comparación, en el audio MPEG-4, se utiliza la sustitución de ruido perceptivo (PNS) sólo para transmitir una información paramétrica de las partes de señal de tipo ruido y para reproducir estas partes de señal de manera perceptivamente equivalente en el decodificador.

[0111] Como comparación adicional, en AMR-WB+, los vectores de cuantificación de vector (vectores VQ) cuantificados a cero se reemplazan por un vector de ruido aleatorio donde cada valor espectral complejo tiene amplitud constante, pero fase aleatoria. La amplitud se controla por un valor de ruido transmitido con el tren de bits.

[0112] Sin embargo, los conceptos de comparación tienen desventajas significativas. La PNS sólo puede utilizarse para rellenar bandas de factores de escala completas, mientras el AMR–WB+ sólo intenta enmascarar los artefactos en la señal decodificada que resultan de partes grandes de la señal que están cuantificadas a cero. En cambio, el esquema de codificación de relleno de ruido propuesto combina eficientemente ambos aspectos del relleno de ruido en una única aplicación.

[0113] Según un aspecto, la presente invención comprende una nueva forma de cálculo de nivel de ruido. El nivel de ruido se calcula en el dominio cuantificado basándose en el error de cuantificación promedio.

[0114] El error de cuantificación en el dominio cuantificado difiere de otras formas de error de cuantificación. El error de cuantificación por línea en el dominio cuantificado está en el intervalo de [–0,5; 0,5] (nivel de cuantificación 1) con un error absoluto promedio de 0,25 (para valores de entrada de distribución normal que normalmente son mayores que 1).

[0115] En lo que viene a continuación, se resumirán algunas ventajas del relleno de ruido en el dominio cuantificado. La ventaja de añadir ruido en el dominio cuantificado es el hecho de que el ruido añadido en el decodificador se cambia de escala, no sólo con la energía promedio en una banda dada, sino también la relevancia psicoacústica de una banda.

[0116] Normalmente, las bandas (tonales) más relevantes perceptivamente serán las bandas cuantificadas con más exactitud, lo que significa que en estas bandas se utilizarán múltiples niveles de cuantificación (valores cuantificados mayores que 1). Ahora, añadir ruido con un nivel del error de cuantificación promedio tendrá sólo una influencia muy limitada sobre la percepción de tal banda.

[0117] Las bandas que no son perceptivamente tan importantes o más similares a ruido, pueden ser cuantificadas con un menor número de niveles de cuantificación. Aunque se cuantificarán a cero muchas más líneas de la banda, el error de cuantificación promedio resultante será el mismo que para las bandas finamente cuantificadas (suponiendo un error de cuantificación de distribución normal en ambas bandas), mientras que el error relativo en la

banda puede ser mucho más alto.

[0118] En estas bandas cuantificadas de manera gruesa, el relleno de ruido ayudará a enmascarar perceptivamente los artefactos que resultan de los agujeros espectrales debidos a la cuantificación gruesa.

[0119] Puede alcanzarse una consideración del relleno de ruido en el dominio cuantificado mediante el codificador descrito anteriormente y también mediante del decodificador descrito anteriormente.

5. Implementaciones alternativas

[0120] Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden implementarse en hardware o en software. La implementación puede llevarse a cabo 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 electrónicamente almacenadas en los mismos, que cooperan (o son capaces de cooperar) con un sistema informático programable de modo que se lleva a cabo el procedimiento respectivo.

REIVINDICACIONES

1. Un decodificador (500; 600) para proporcionar una representación decodificada (512, 514; 630b) de una señal de audio basándose en un flujo de audio codificado (510; 610) que representa componentes espectrales de bandas de frecuencias de la señal de audio y un valor de intensidad de ruido multibanda (526), comprendiendo el decodificador:

un rellenador de ruido (520; 770) configurado para introducir ruido en las componentes espectrales de una pluralidad de bandas de frecuencias, a las que se asocia información de ganancia de banda de frecuencia separada, en particular factores de escala, basándose en un valor de intensidad de ruido multibanda común (526), y
donde comenzando desde una línea de comienzo de relleno de ruido, cada línea espectral cuantificada a cero es reemplazada por un valor de reemplazo, que es un valor de ruido indicado, cuya magnitud se determina mediante el valor de intensidad de ruido multibanda, con un signo aleatorio, para realizar un relleno de ruido en un dominio cuantificado;
donde el decodificador está configurado para cambiar de escala un valor de reemplazo con un factor de escala transmitido para una banda de factores de escala real,
para realizar un cambio de escala individual del ruido introducido en distintas bandas de frecuencias basándose en la información de ganancia de banda de frecuencia.

2. Un procedimiento para proporcionar una representación decodificada (512; 514; 630b) de una señal de audio basándose en un flujo de audio codificado (510; 620) que representa componentes espectrales de bandas de frecuencias de la señal de audio y un valor de intensidad de ruido multibanda (526), comprendiendo el procedimiento:

introducir ruido en las componentes espectrales de una pluralidad de bandas de frecuencias, a las que se asocia información de ganancia de banda de frecuencia separada, en particular factores de escala, basándose en un valor de intensidad de ruido multibanda común;
donde comenzando desde una línea de comienzo de relleno de ruido, cada línea espectral cuantificada a cero es reemplazada por un valor de reemplazo, que es un valor de ruido indicado, cuya magnitud se determina mediante el valor de intensidad de ruido multibanda, con un signo aleatorio, para realizar un relleno de ruido en un dominio cuantificado; y
para realizar un cambio de escala individual del ruido introducido en distintas bandas de frecuencias basándose en la información de ganancia de banda de frecuencia.

3. Un programa de ordenador para llevar a cabo un procedimiento de acuerdo con la reivindicación 2 cuando el programa de ordenador se ejecuta en un ordenador.

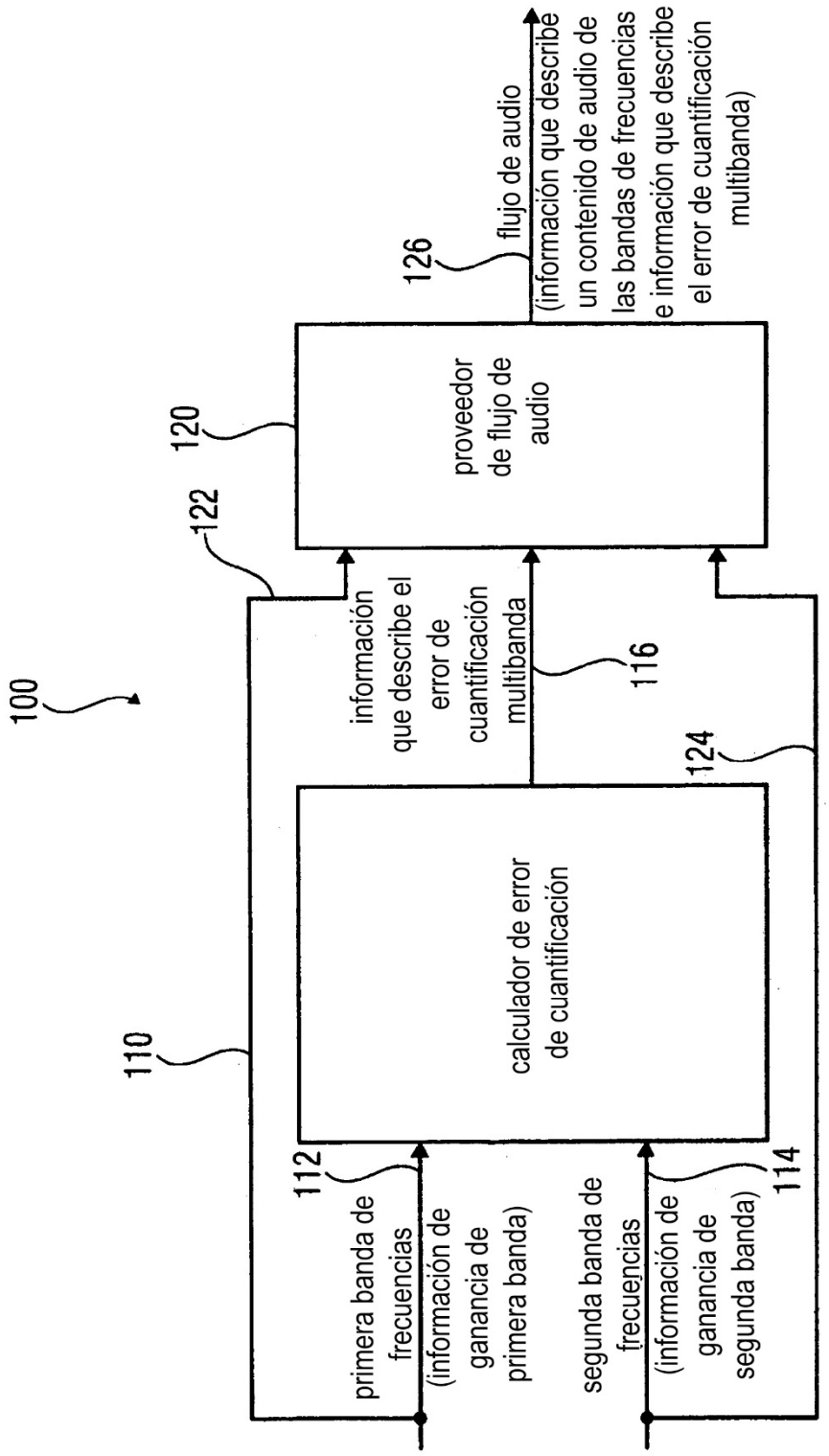

100
110
calculador de error de cuantificación
120
proveedor de flujo de audio
122
información que describe el error de cuantificación multibanda
116
124
126
flujo de audio (información que describe un contenido de audio de las bandas de frecuencias e información que describe el error de cuantificación multibanda)
112
primera banda de frecuencias (información de ganancia de primera banda)
114
segunda banda de frecuencias (información de ganancia de segunda banda)

FIG 1

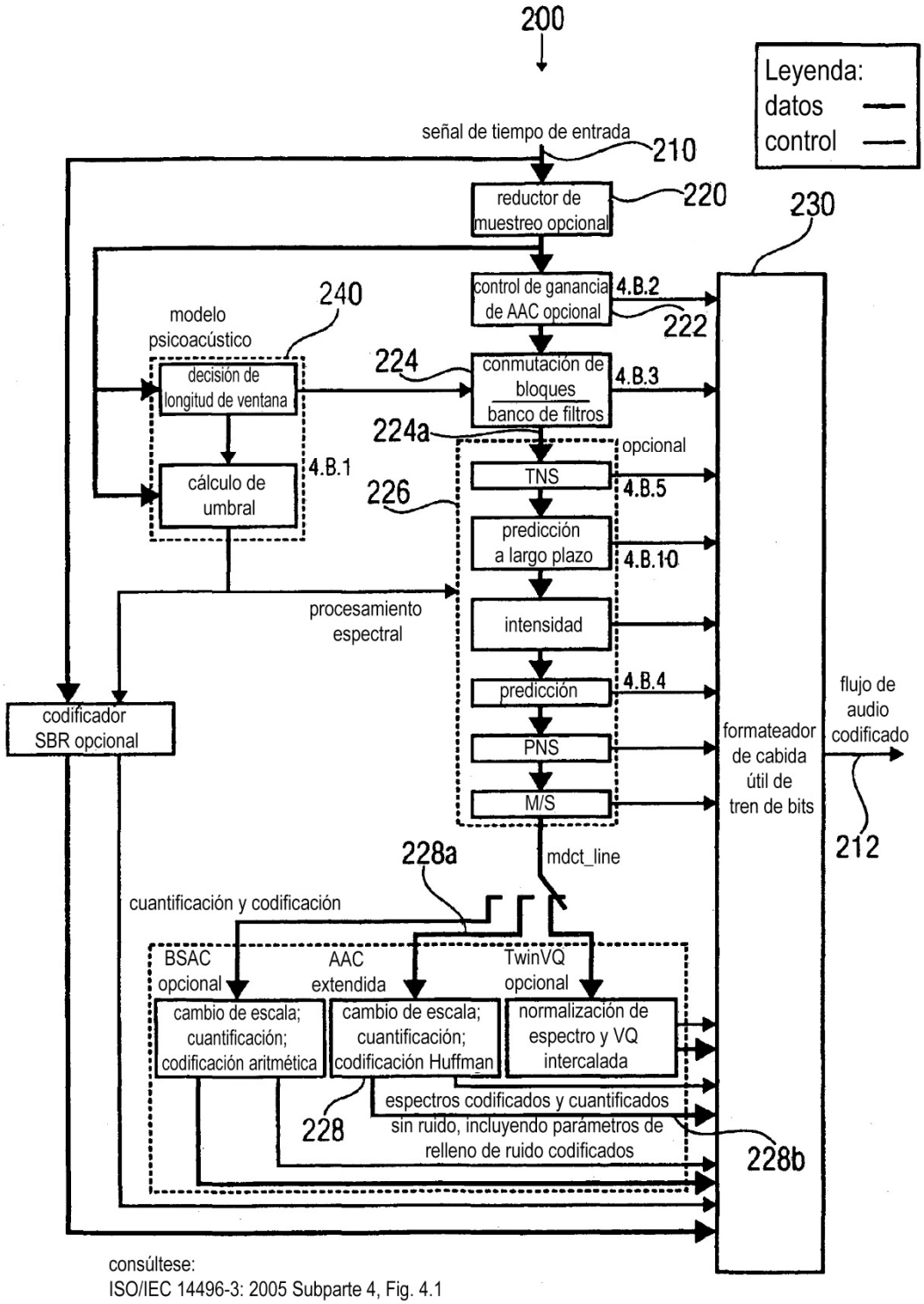
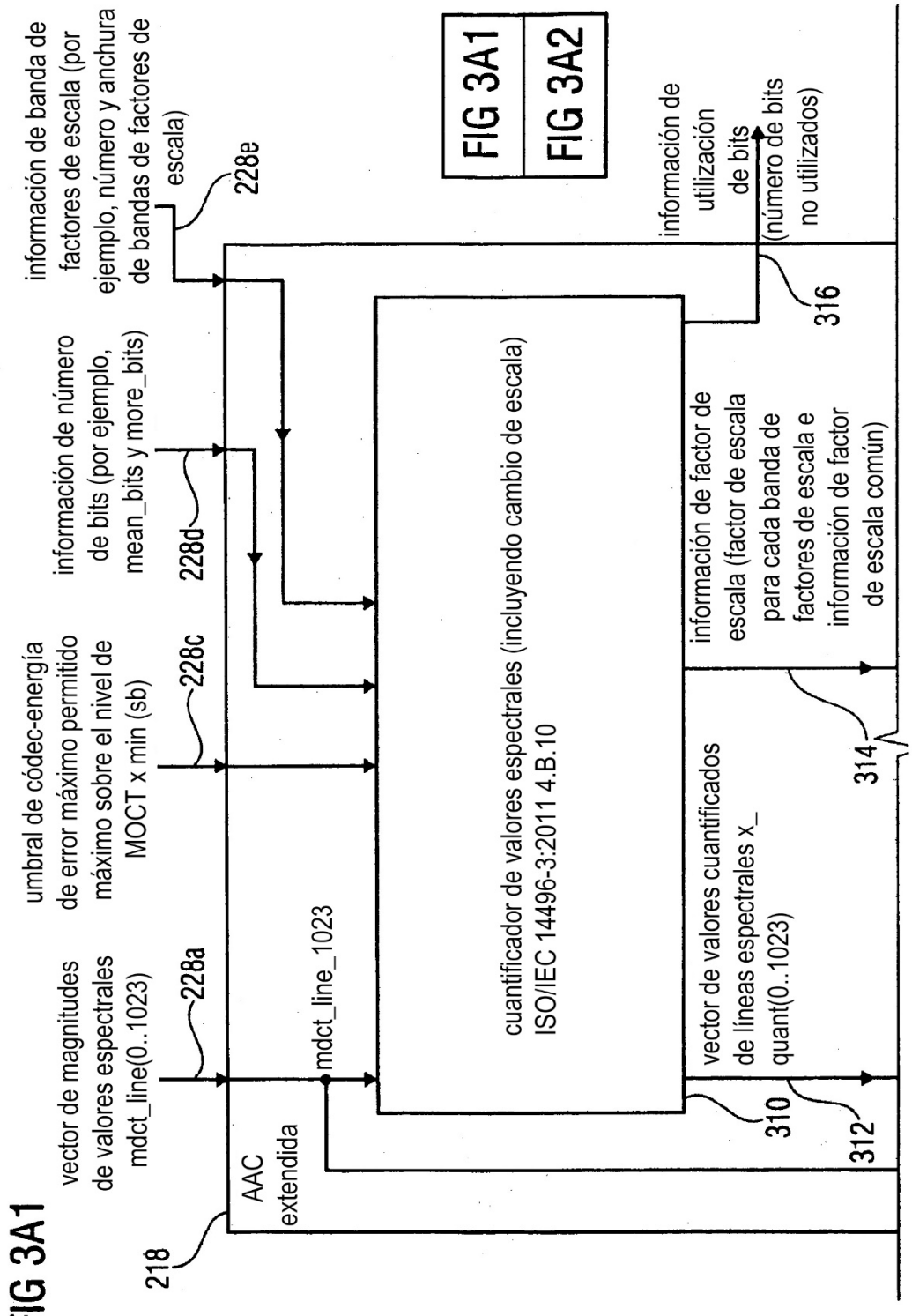

200
Leyenda:
datos
control
señal de tiempo de entrada
210
reductor de muestreo opcional
220
230
control de ganancia de AAC opcional
4.B.2
222
240
modelo psicoacústico
decisión de longitud de ventana
224
conmutación de bloques banco de filtros
4.B.3
224a
opcional
cálculo de umbral
4.B.1
226
TNS
4.B.5
predicción a largo plazo
4.B.10
procesamiento espectral
intensidad
predicción
4.B.4
codificador SBR opcional
PNS
M/S
formateador de cabida útil de tren de bits
flujo de audio codificado
212
228a
mdct_line
cuantificación y codificación
BSAC opcional
AAC extendida
TwinVQ opcional
cambio de escala; cuantificación; codificación aritmética
cambio de escala; cuantificación; codificación Huffman
normalización de espectro y VQ intercalada
espectros codificados y cuantificados sin ruido, incluyendo parámetros de relleno de ruido codificados
228
228b
consúltese:
ISO/IEC 14496-3: 2005 Subparte 4, Fig. 4.1

FIG 2


FIG 3A1
vector de magnitudes de valores espectrales mdct_line(0..1023)
228a
umbral de códec-energía de error máximo permitido máximo sobre el nivel de MOCT x min (sb)
228c
información de número de bits (por ejemplo, mean_bits y more_bits)
228d
información de banda de factores de escala (por ejemplo, número y anchura de bandas de factores de escala)
228e
218
AAC extendida
mdct_line_1023
cuantificador de valores espectrales (incluyendo cambio de escala) ISO/IEC 14496-3:2011 4.B.10
vector de valores cuantificados de líneas espectrales x_quant(0..1023)
310
312
información de factor de escala (factor de escala para cada banda de factores de escala e información de factor de escala común)
314
información de utilización de bits (número de bits no utilizados)
316
FIG 3A1
FIG 3A2

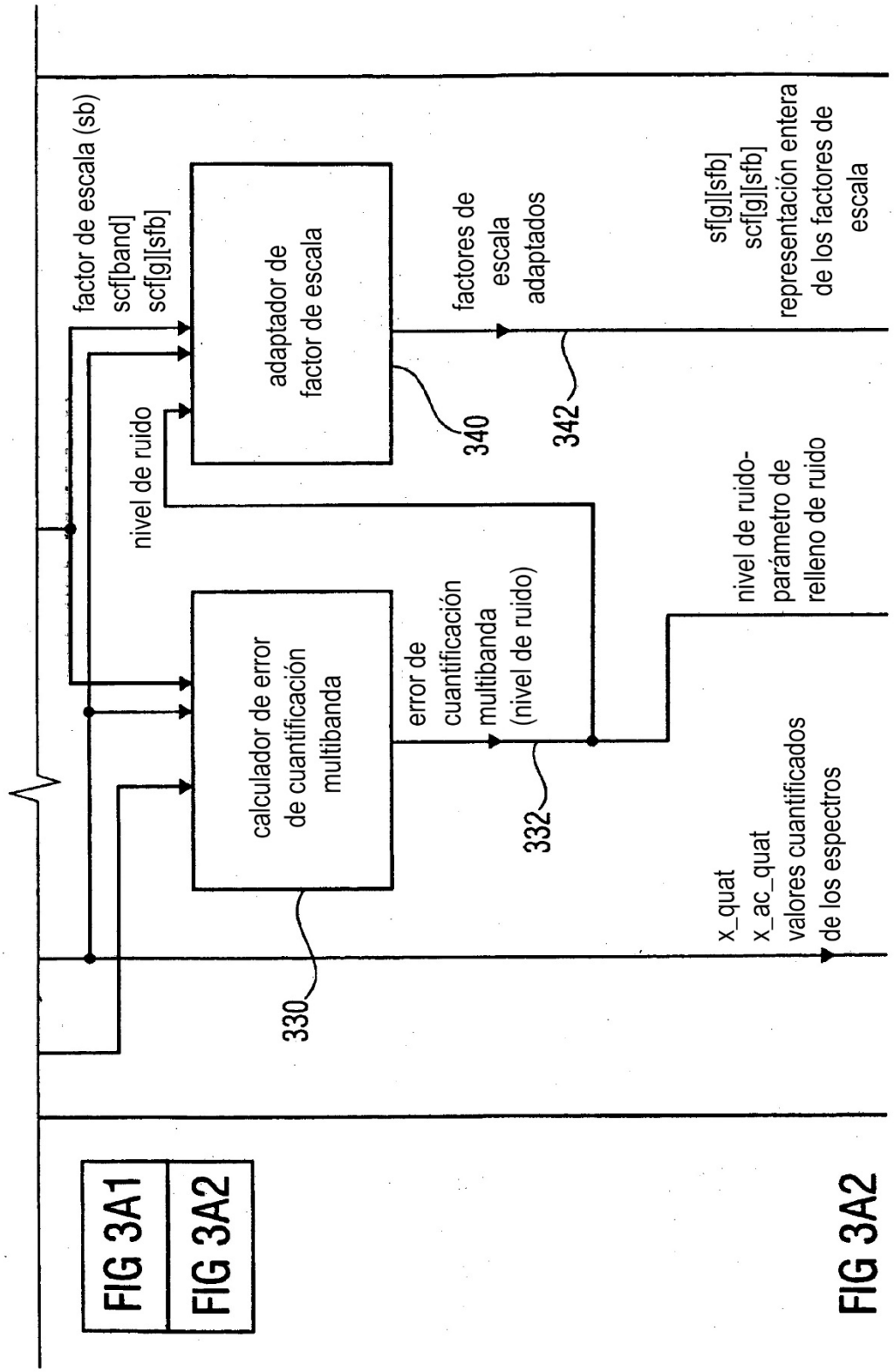


FIG 3A1 | FIG 3A2

FIG 3A2

FIG 3B

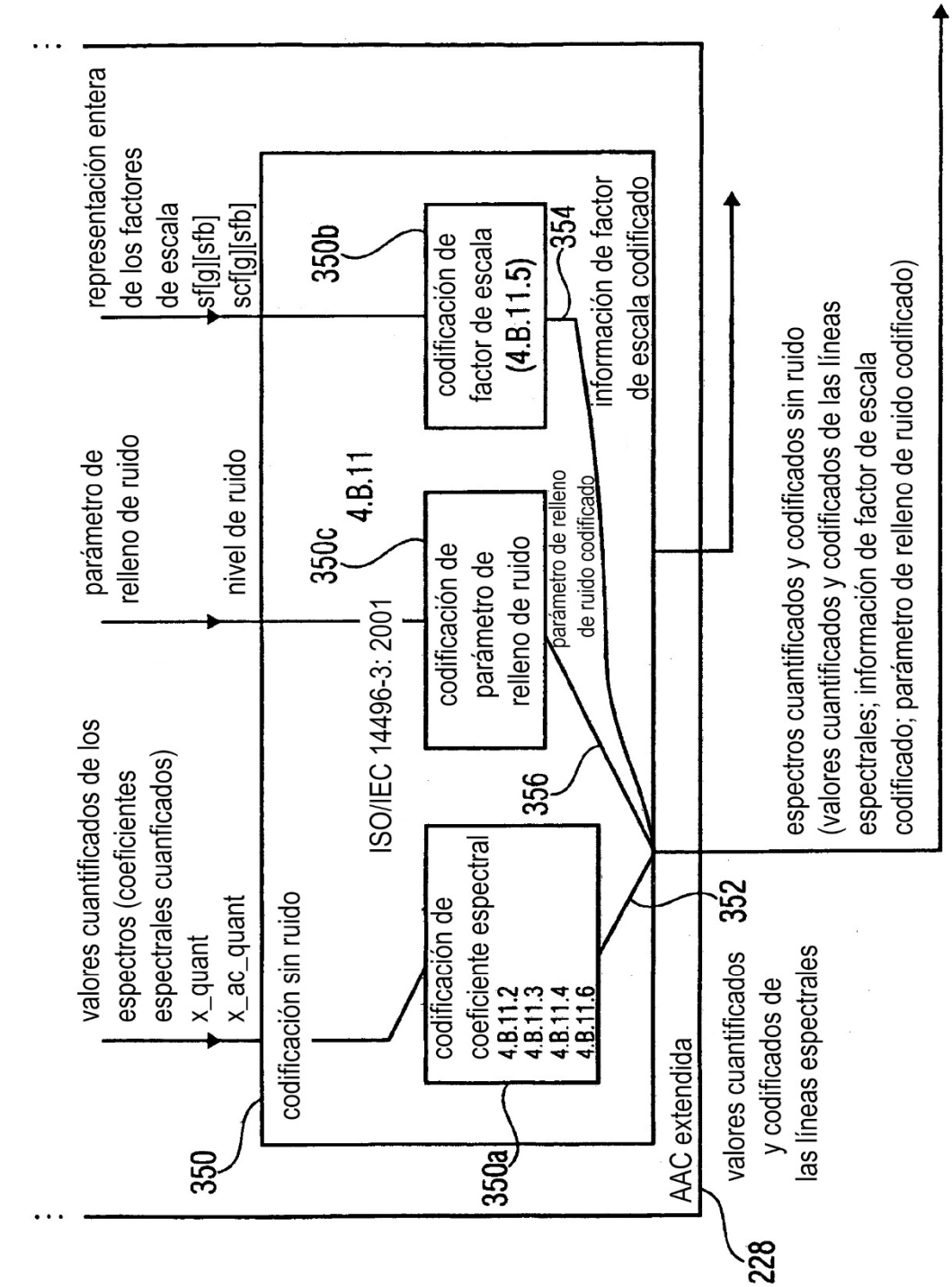

valores cuantificados de los espectros (coeficientes espectrales cuanficados)
x_quant
x_ac_quant
parámetro de relleno de ruido
nivel de ruido
representación entera de los factores de escala
sf[g][sfb]
scf[g][sfb]
codificación sin ruido
ISO/IEC 14496-3: 2001
4.B.11
350
350a
350b
350c
codificación de coeficiente espectral
4.B.11.2
4.B.11.3
4.B.11.4
4.B.11.6
codificación de parámetro de relleno de ruido
codificación de factor de escala
(4.B.11.5)
356
354
parámetro de relleno de ruido codificado
información de factor de escala codificado
AAC extendida
352
228
valores cuantificados y codificados de las líneas espectrales
espectros cuantificados y codificados sin ruido (valores cuantificados y codificados de las líneas espectrales; información de factor de escala codificado; parámetro de relleno de ruido codificado)

FIG 4A

Codificador:

```
1   Calculate Mean Quantization error:
2   nLines = 0;
3   avgError = 0;

4   for (band = all scale factor bands) {
5     for(line = all spectral lines in band) {
6       if(band not quantized to zero) {
7         avgError += fabs ( pow(line, 0.75)*scale factor - (int) pow(line, 0.75)*scale factor)
8         nLines ++;
9       }
10    }
11  }
12  avgError=avgError / nLines;

13  noiseLevel=(int) (14+4*ld(avgError);
14  noiseLevel=max(0, min(7, noiseLevel)

15  Calculate All zero Scale Factor:
16  noiseValue=pow(2.f, ((float)(noiseLevel)-14.f)/4.f)
17  if(noiseLevel>0) {
18  for (band=all scale factor bands){
19    if(band quantized to zero) {
20      scf[band]=(INT) (2.f * log( ((energy)/((float)sfbWidth*noiseVal*noiseVal*noiseVal)))/log(2.f));
21    }
22  }
23  else {
24    scf=don't care
25  }
```

número de línea

valor de magnitud de línea espectral no cuantificada, cambiada de escala

valor de magnitud de línea espectral cuantificada, cambiada de escala

opcional: cuantificación de nivel de ruido

opcional: cuantificación inversa de nivel de ruido

Cuantificación de nivel de ruido

```
noiseLevel = (int) (14 + 4*ld(meanLineError);
noiseLevel = max(0, min(7, noiseLevel)
```

FIG 4B

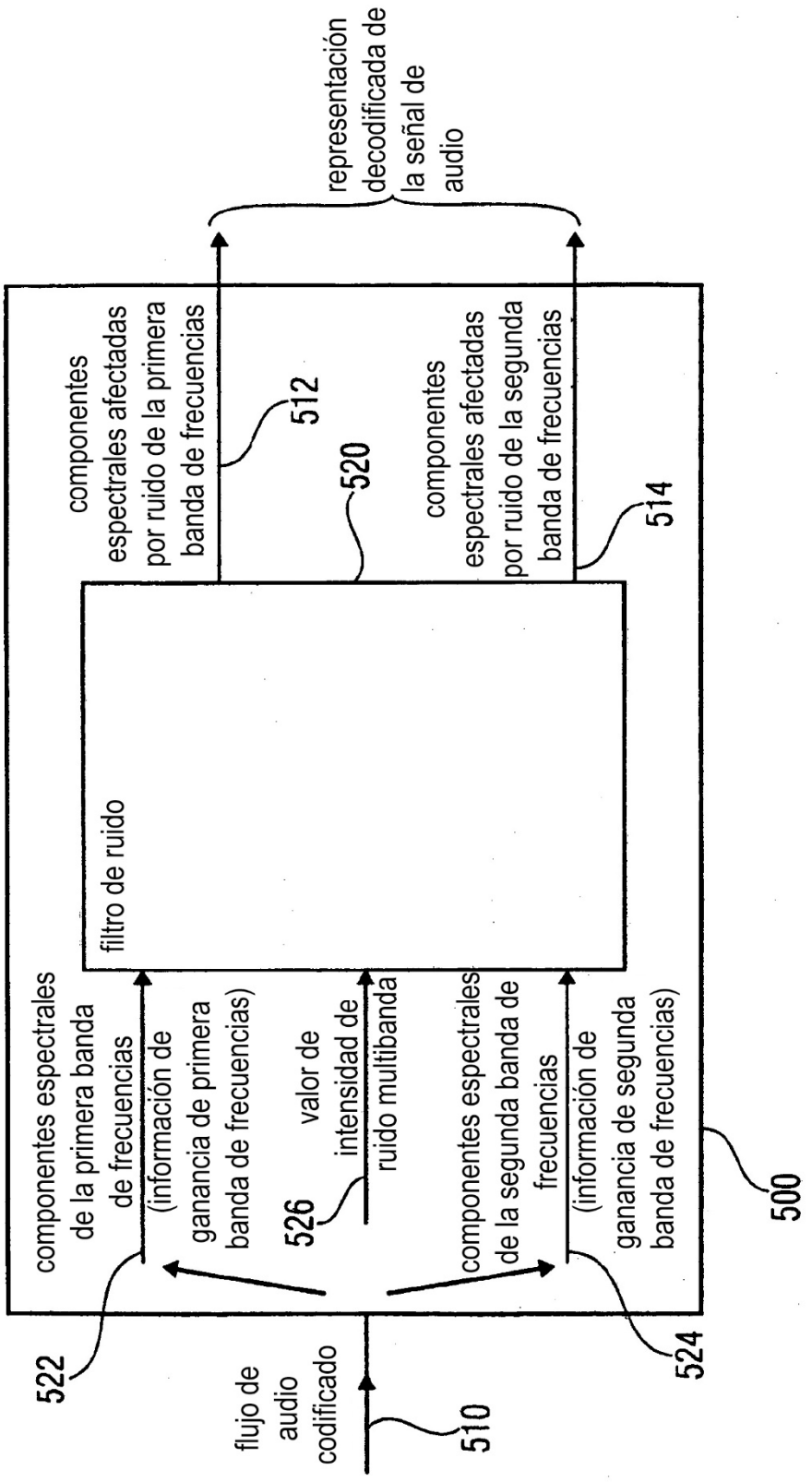

flujo de audio codificado
510
componentes espectrales de la primera banda de frecuencias
522
(información de ganancia de primera banda de frecuencias)
valor de intensidad de ruido multibanda
526
componentes espectrales de la segunda banda de frecuencias
524
(información de ganancia de segunda banda de frecuencias)
500
filtro de ruido
componentes espectrales afectadas por ruido de la primera banda de frecuencias
512
520
componentes espectrales afectadas por ruido de la segunda banda de frecuencias
514
representación decodificada de la señal de audio

FIG 5

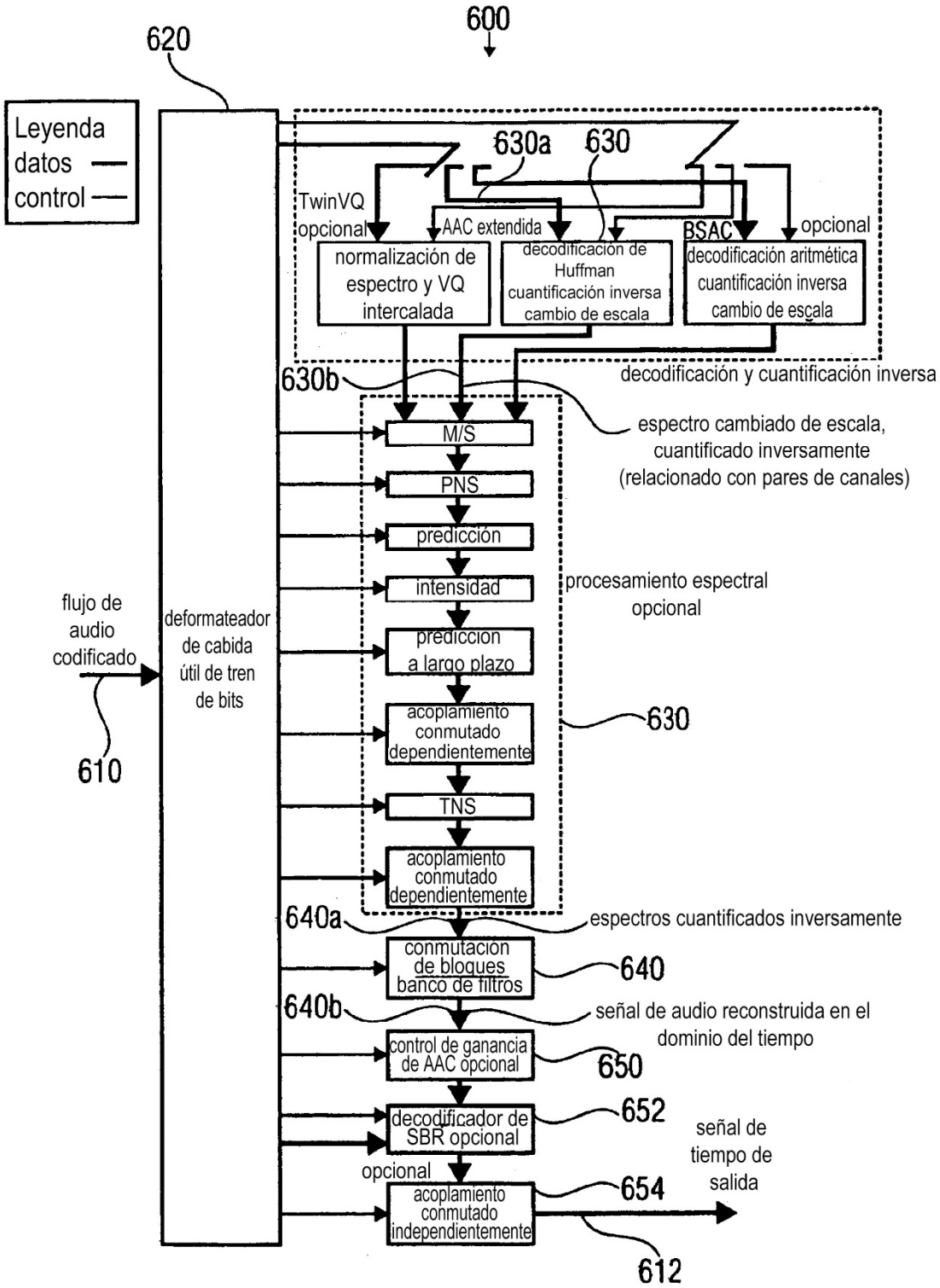

600
620
Leyenda
datos
control
630a
630
TwinVQ
opcional
AAC extendida
BSAC
opcional
normalización de espectro y VQ intercalada
decodificación de Huffman cuantificación inversa cambio de escala
decodificación aritmética cuantificación inversa cambio de escala
decodificación y cuantificación inversa
630b
espectro cambiado de escala, cuantificado inversamente (relacionado con pares de canales)
M/S
PNS
predicción
intensidad
predicción a largo plazo
acoplamiento conmutado dependientemente
TNS
acoplamiento conmutado dependientemente
procesamiento espectral opcional
630
flujo de audio codificado
610
deformateador de cabida útil de tren de bits
640a
espectros cuantificados inversamente
conmutación de bloques banco de filtros
640
640b
señal de audio reconstruida en el dominio del tiempo
control de ganancia de AAC opcional
650
decodificador de SBR opcional
652
opcional
acoplamiento conmutado independientemente
654
señal de tiempo de salida
612

FIG 6

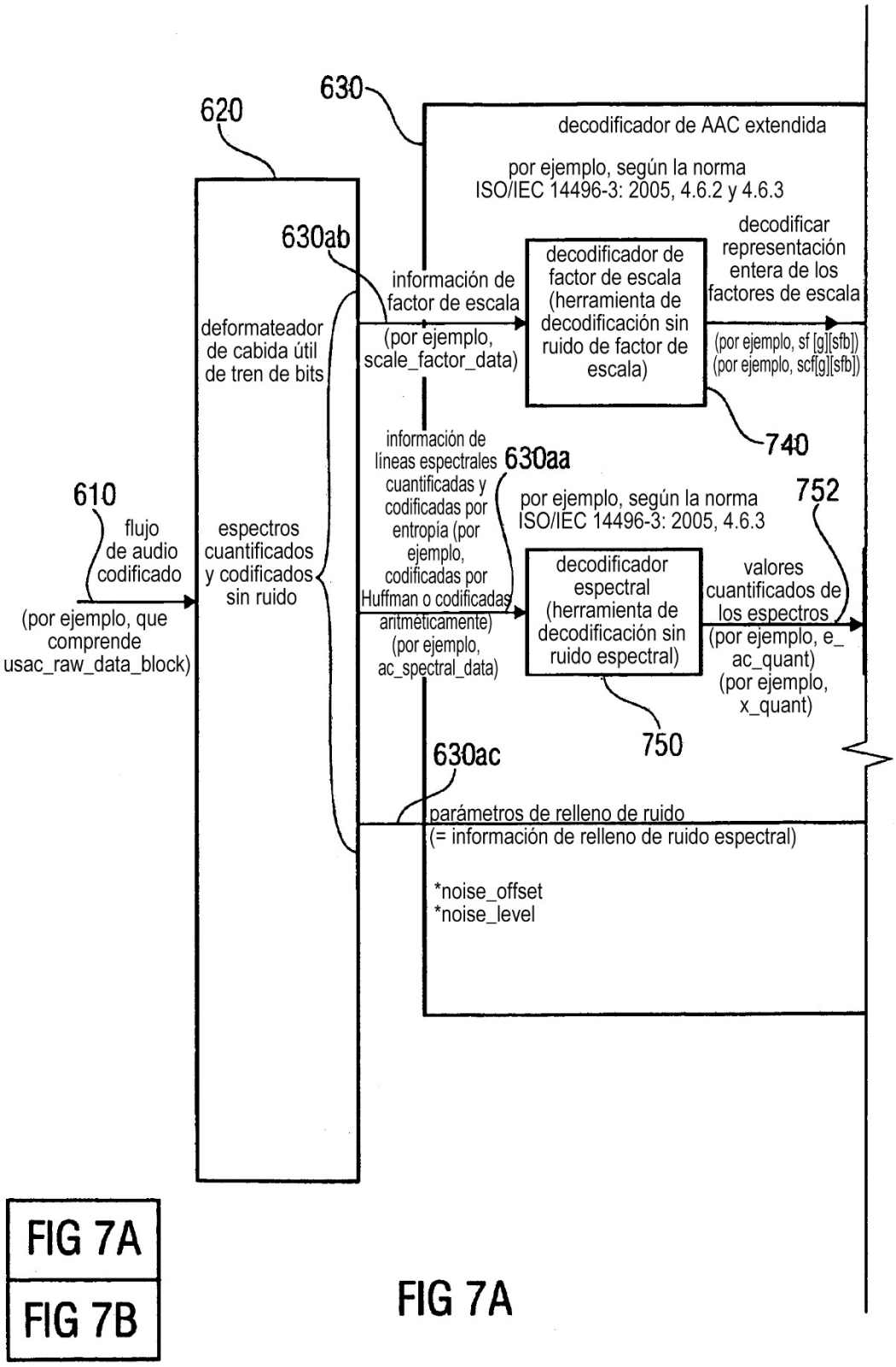


FIG 7A

FIG 7B

FIG 7A

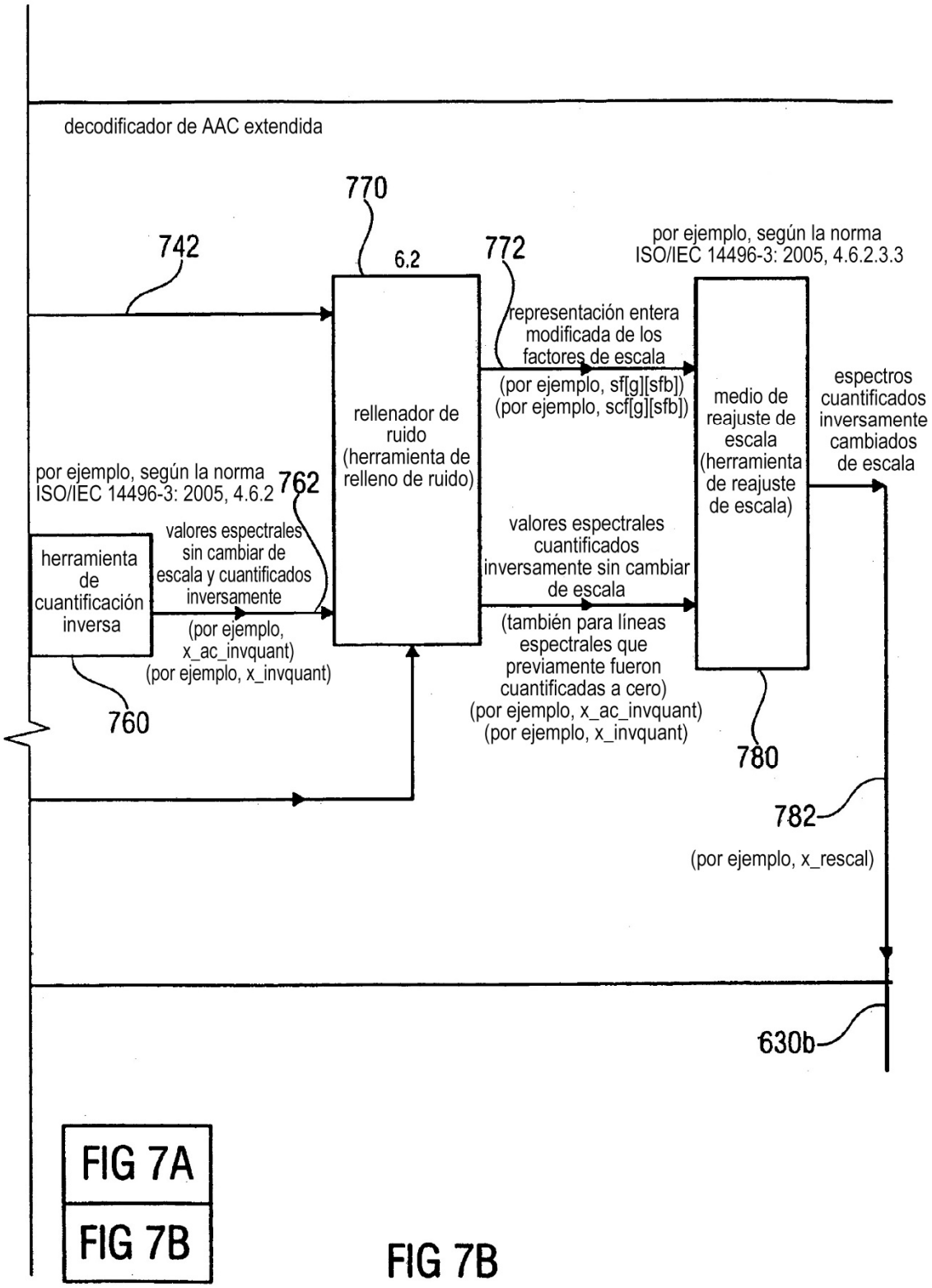


FIG 7A

FIG 7B

FIG 7B

$$x_invquant = Sign(x_quant) \cdot |x_quant|^{\frac{4}{3}}$$

or:

$$x_ac_invquant = Sign(x_ac_quant) \cdot |x_ac_quant|^{\frac{4}{3}}$$

FIG 8A

```
for (g = 0; g< num_window_groups; g++) {
  for (sfb = 0; sfb < max_sfb; sfb++) {
    width = (swb_offset [sfb+1] - swb_offset [sfb]);
    for (win - 0; win < window_group_len[g]; win++) {
      for (bin = 0; bin < width; bin++) {
        x_ac_invquant [g] [sfb] [sfb] [bin] =
sign(x_ac_quant[g] [win] [sfb] [bin])*abs(x_ac_quant[g] [win] [sfb] [bin]) ^ (4/3);
      }
    }
  }
}
```

FIG 8B

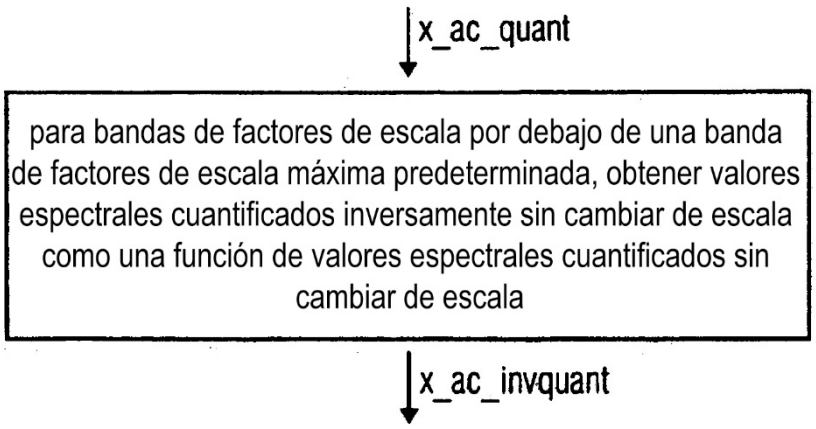


FIG 8C

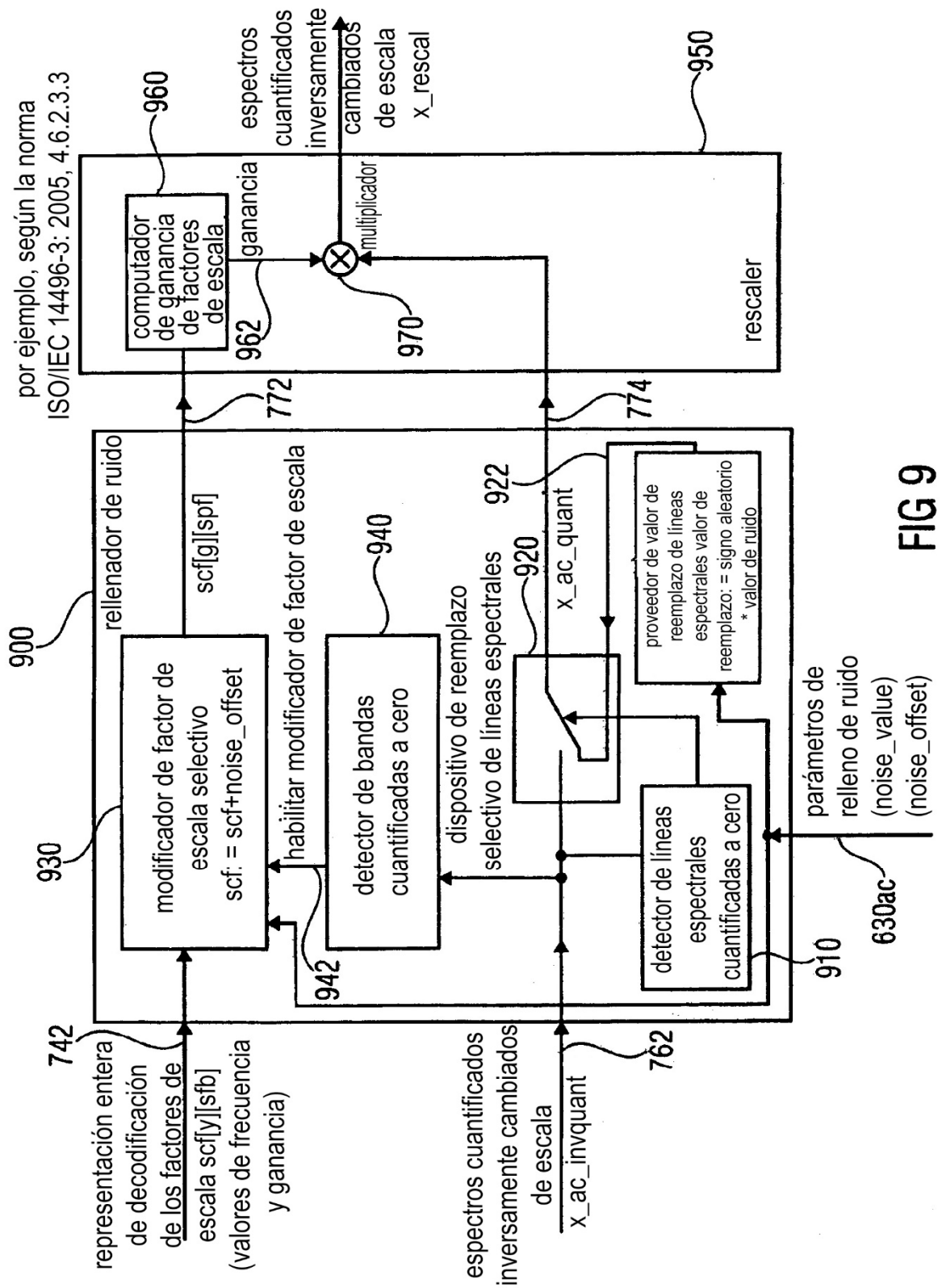

por ejemplo, según la norma ISO/IEC 14496-3: 2005, 4.6.2.3.3
960
computador de ganancia de factores de escala
ganancia
962
multiplicador
970
espectros cuantificados inversamente cambiados de escala
x_rescal
950
rescaler
772
774
900
rellenador de ruido
930
modificador de factor de escala selectivo
scf: = scf+noise_offset
scf[g][spf]
habilitar modificador de factor de escala
942
940
detector de bandas cuantificadas a cero
dispositivo de reemplazo selectivo de líneas espectrales
920
922
x_ac_quant
proveedor de valor de reemplazo de líneas espectrales valor de reemplazo: = signo aleatorio * valor de ruido
detector de líneas espectrales cuantificadas a cero
910
parámetros de relleno de ruido
(noise_value)
(noise_offset)
630ac
742
representación entera de decodificación de los factores de escala scf[y][sfb]
(valores de frecuencia y ganancia)
762
espectros cuantificados inversamente cambiados de escala
x_ac_invquant

FIG 9

Proceso de decodificación: Proceso de relleno de ruido

FIG 10A

número de línea

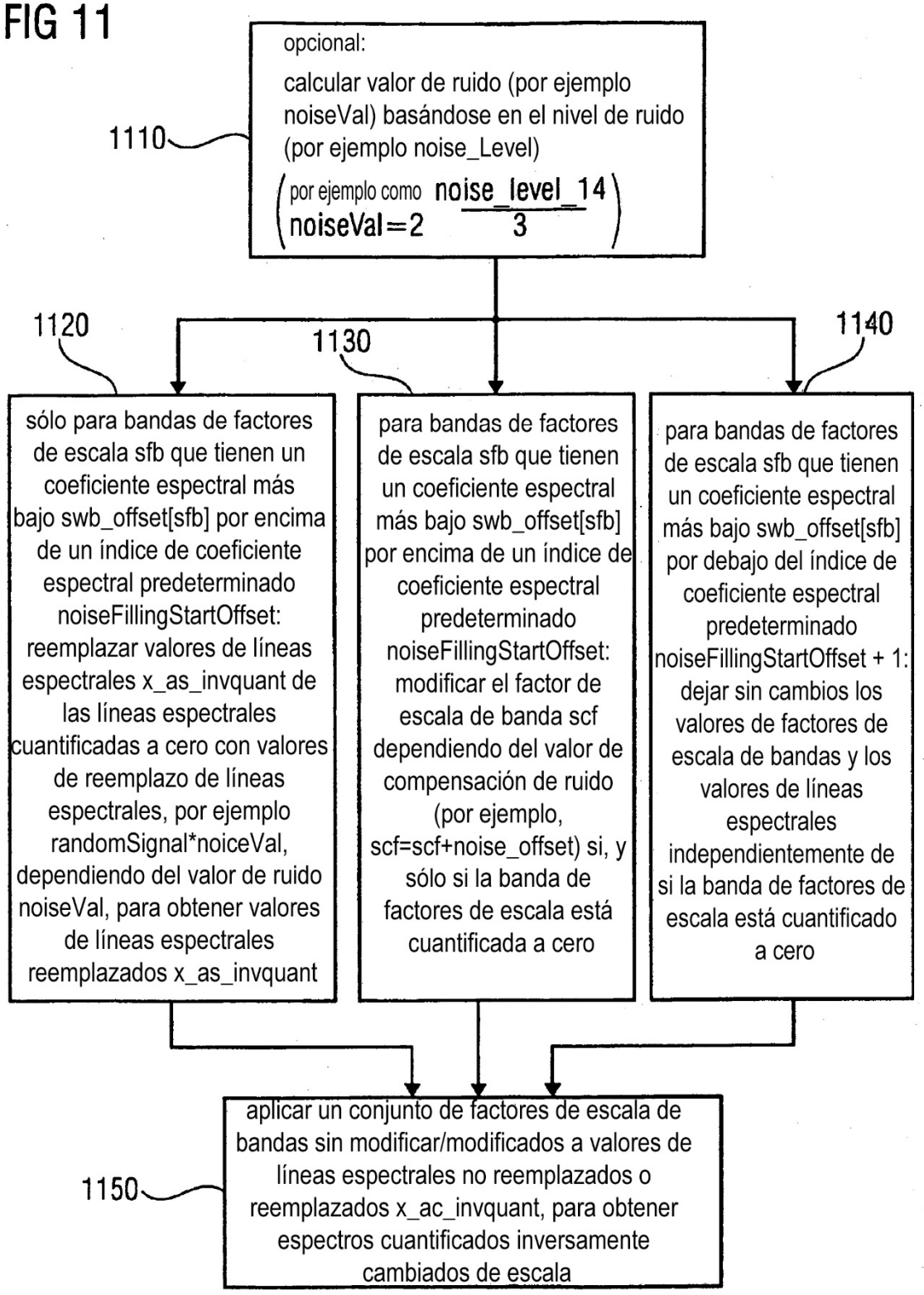
```
 1 if (noise_level !=0) {
 2   noiseVal = pow(2, (noise_level-14)/3) ;
 3   noise_offset = noise_offset - 16;
 4 }
 5 else {
 6   noiseVal = 0;
 7   noise_offset = 0;
 8 }
 9 for (g=0; g<num_window_groups; g++) {
10   for (sfb=0; sfb < max_sfb; sfb++) {
11     band_quantized_to_zero=1;                          -- se supone que una banda está cuantificada a cero
12     width=(swb_offset [sfb+1] - swb_offset [sfb]);
13     if(swb_offset [sfb] > noiseFillingStartOffset) {
14       for (win=0); win < window_group_len[g]; win++) {
15         for (bin=0; bin < width; bin++) {
16           if (x_ac_invquant [g] [win] [sfb] [bin] ==0) {
17             x_ac_invquand [g] [win] [sfb] [bin] = randomSign () *noiseVal;   -- para bandas de factores de escala, comenzar por encima de noiseFillingStartOffset;
18           }                                                                  -- añadir ruido de amplitud noiseVal a las líneas espectrales cuantificadas a cero;
19           else {
20             band_quantized_to_zero=0;                                        -- para bandas de factores de escala, comenzar por encima de noiseFillingStartOffset,
21           }                                                                  -- si un solo segmento de una banda de factores de escala es ≠0, entonces la banda
22         }                                                                    -- no está cuantificada a cero;
23       }
24     }
24a    else {
24b      band_quantized_to_zero=0;                          -- para bandas de factores de escala, comenzar por debajo de noiseFillingStartOffset,
24c    }                                                    -- si siempre se supone que la banda no está cuantificada a cero;

25     if(band_quantized_to_zero) {
26       scf[g] [sfb] = scf[g] [sfb] + noise_offset;        -- para bandas de factores de escala cuantificadas a cero, modifica el factor de
27     }                                                    -- escala de banda dependiendo del valor de compensación de ruido
28   }
29 }
```

Elementos de datos	
noise_offset	compensación adicional para modificar el factor de escala de las bandas cuantificadas a cero
noise_level	número entero que representa el ruido de cuantificación que ha de añadirse para cada línea espectral cuantificada a cero
Elementos de ayuda	
x_ac_invquant[g][win][sfb][bin]	coeficiente espectral de AAC para el grupo g, ventana win, banda de factores de escala sfb, segmento de coeficientes después de la cuantificación inversa
noiseFillingStartOffset[win]	una compensación general o frecuencia de comienzo de relleno de ruido. La compensación está definida para ser 20 para una corta (window_sequence = = EIGHT_SHORT_SEQUENCE) y 160 en otros casos.
noiseVal	El valor de ruido absoluto que reemplaza cada segmento cuantificado a cero
randomSign	Signo (-1, 1) aleatorio multiplicado con noiseVal
band_quantized_to_zero[	indicador para señalar si una sfb está completamente cuantificada a cero
swb_offset[sfb]	índice del coeficiente espectral más bajo de la banda de factores de escala sfb
num_window_groups	número de grupos de ventanas que comparten un conjunto de factores de escala
max_sfb	número de bandas de factores de escala por grupo
window_group_len	número de ventanas en cada grupo
g	índice de grupo
win	índice de ventana dentro del grupo
sfb	índice de banda de factores de escala dentro del grupo
swb	índice de banda de ventana de factores de escala dentro de la ventana
bin	índice de coeficiente
num_windows	número de ventanas de la secuencia de ventanas real

FIG 10B

FIG 11


opcional:
calcular valor de ruido (por ejemplo noiseVal) basándose en el nivel de ruido (por ejemplo noise_Level)
(por ejemplo como noiseVal=2 noise_level_14 / 3)
1110
1120
1130
1140
sólo para bandas de factores de escala sfb que tienen un coeficiente espectral más bajo swb_offset[sfb] por encima de un índice de coeficiente espectral predeterminado noiseFillingStartOffset: reemplazar valores de líneas espectrales x_as_invquant de las líneas espectrales cuantificadas a cero con valores de reemplazo de líneas espectrales, por ejemplo randomSignal*noiceVal, dependiendo del valor de ruido noiseVal, para obtener valores de líneas espectrales reemplazados x_as_invquant
para bandas de factores de escala sfb que tienen un coeficiente espectral más bajo swb_offset[sfb] por encima de un índice de coeficiente espectral predeterminado noiseFillingStartOffset: modificar el factor de escala de banda scf dependiendo del valor de compensación de ruido (por ejemplo, scf=scf+noise_offset) si, y sólo si la banda de factores de escala está cuantificada a cero
para bandas de factores de escala sfb que tienen un coeficiente espectral más bajo swb_offset[sfb] por debajo del índice de coeficiente espectral predeterminado noiseFillingStartOffset + 1: dejar sin cambios los valores de factores de escala de bandas y los valores de líneas espectrales independientemente de si la banda de factores de escala está cuantificado a cero
aplicar un conjunto de factores de escala de bandas sin modificar/modificados a valores de líneas espectrales no reemplazados o reemplazados x_ac_invquant, para obtener espectros cuantificados inversamente cambiados de escala
1150

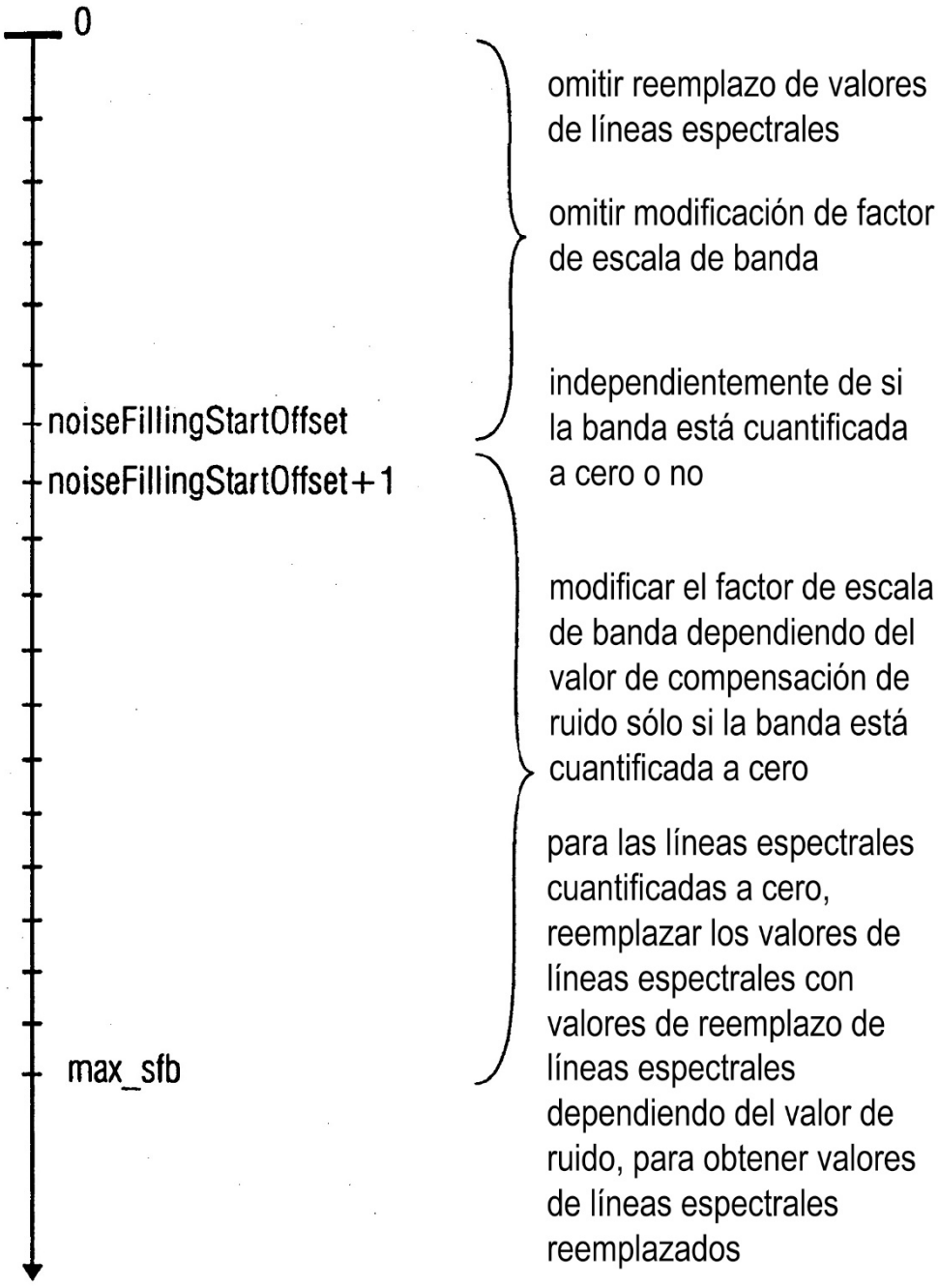

0
noiseFillingStartOffset
noiseFillingStartOffset+1
max_sfb
omitir reemplazo de valores de líneas espectrales
omitir modificación de factor de escala de banda
independientemente de si la banda está cuantificada a cero o no
modificar el factor de escala de banda dependiendo del valor de compensación de ruido sólo si la banda está cuantificada a cero
para las líneas espectrales cuantificadas a cero, reemplazar los valores de líneas espectrales con valores de reemplazo de líneas espectrales dependiendo del valor de ruido, para obtener valores de líneas espectrales reemplazados
segmento de comienzo de banda de factores de escala

FIG 12

número de línea	Valor de ruido en el decodificador
1	if(noiseLevel !=0)
2	noiseValue=pow(2.f((float)(noiseLevel)-14.f)/4.f)
3	else
4	noiseVal=0

FIG 13A

número de línea	Decodificador: "Cálculo de "Valor de reemplazo de cero":
1	if(noiseLevel !=0)
2	noiseValue=pow(2.f((float)(noiseLevel)-14.f)/4.f)
3	else
4	noiseVal=0
5	for(band=all scale factor bands) {
6	for(line=all spectral lines in band) {
7	if(band quantized to zero) {
8	scf=scf + noise_Offset {
9	}
10	if(line>noiseFillingStartOffset) {
11	if(quantizedSpec[line]==0) {
12	quantizedSpec[line]=randomSign()*noiseValue;
13	}
14	}
15	}
16	}

FIG 13B

Cabida útil de tren de bits USAC (Tabla 4.3)

```
usac_raw_data_block ()
{
  single_channel_element () ; and/or
  channel_pair_element () ;
  opcional: elementos de canal adicionales
}
```

FIG 14A

single_channel_element () (Tabla 4.4)

```
{
  fd_channel_stream (*,*, relleno de ruido)
}
```

FIG 14B

channel_pair_element (Tabla 4.5)

```
{
  fd_channel_stream (*,*, relleno de ruido) (1er canal); y/o
  fd_channel_stream (*,*, relleno de ruido) (2º canal)
}
```

FIG 14C

fd_channel_stream () (Tabla 4.8)

```
{
  global_gain;                  por ejemplo 8 bits
  noise_offset;                 por ejemplo 3 bits
  noise_level;                  por ejemplo 5 bits

  scale_factor_data ();

  tus_data ();                  opcional

  ac_spectral_data ()
}
```

FIG 14D

Sintaxis de individual_channel_stream()

Sintaxis	Nº de bits	Mnemotécnica
individual_channel_stream(common_window)		
{		
global_gain;	8	uimsbf
noise_Offset	5	uimsbf
noise_Level	3	uimsbf
if (!common_window)		
ics_info();		
section_data();		
scale_factor_data();		
pulse_data_present;	1	uimsbf
if (pulse_data_present) {		
pulse_data();		
}		
tns_data_present;	1	uimsbf
if(tns_data_present){		
tns_data();		
}		
gain_control_data_present;	1	uimsbf
if (gain_control_data_present) {		
gain_control_data();		
}		
spectral_data()		
}		

FIG 15