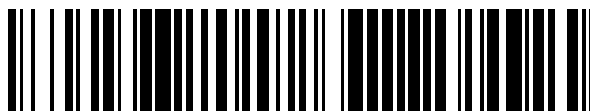


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 526 767**

51 Int. Cl.:

**G10L 19/028** (2013.01)

**G10L 19/00** (2013.01)

**G10L 19/02** (2013.01)

**G10L 19/022** (2013.01)

**G10L 19/035** (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 11157204 (6)**

97 Fecha y número de publicación de la concesión europea: **01.10.2014 EP 2346030**

54 Título: **Codificador de audio, procedimiento para codificar 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:  
**15.01.2015**

73 Titular/es:

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

72 Inventor/es:

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

74 Agente/Representante:

**PONTI SALES, Adelaida**

ES 2 526 767 T3

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

## DESCRIPCIÓN

Codificador de audio, procedimiento para codificar una señal de audio y programa de ordenador

5 Antecedentes de la invención

10 **[0001]** Las formas de realización según la invención se refieren a un codificador para proporcionar una corriente de audio a partir de una representación en el dominio de la transformada de una señal de audio de entrada. Otra realización según la invención proporciona un procedimiento para codificar una señal de audio. Además unas realizaciones según la invención proporcionan programas de ordenador para codificar una señal de audio.

**[0002]** En términos generales, las realizaciones de acuerdo con la invención están relacionadas con un relleno de ruido.

15 **[0003]** Los conceptos de codificación de audio a menudo codifican una señal de audio en el dominio de la frecuencia. Por ejemplo, el llamado concepto de "codificación de audio avanzada" (AAC) codifica el contenido de los diferentes contenedores espectrales (o contenedores de frecuencia), teniendo en cuenta un modelo psicoacústico. Para este propósito, se codifica información de intensidad para diferentes contenedores espectrales. Sin embargo, la resolución utilizada para codificación de intensidades en diferentes contenedores espectrales está adaptada de acuerdo con las relevancias psicoacústicas de los diferentes contenedores espectrales. Por lo tanto, algunos contenedores espectrales, que son considerados como de baja relevancia psicoacústica, se codifican con una resolución de muy baja intensidad, de tal manera que algunos de los contenedores espectrales considerados como de baja relevancia psicoacústica, o incluso un número dominante de estos, se cuantifican a cero. La cuantificación de la intensidad de un contenedor espectral a cero trae consigo la ventaja de que el valor cuantizado a cero se puede codificar de una manera muy económica en cuanto a bits, lo que ayuda a mantener la tasa de bits lo más pequeña posible. Sin embargo, los contenedores espectrales cuantificados a cero a veces dan como resultado artefactos audibles, incluso si el modelo psicoacústico indica que los contenedores espectrales son de baja relevancia psicoacústica.

20 **[0004]** Por lo tanto, hay un deseo de tratar con contenedores espectrales cuantificados a cero, tanto en un codificador de audio como en un decodificador de audio.

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

35 **[0006]** Por ejemplo, el estándar MPEG-4 "AAC" (codificación avanzada de audio) utiliza el concepto de sustitución de ruido perceptivo (PNS). La sustitución de ruido de percepción llena bandas de factor de escala completas con sólo ruido. Los detalles relativos a la MPEG-4 AAC pueden, por ejemplo, encontrarse en la Norma ISO / IEC 14496-3 (Information Technology - Coding of Audio-Visual Objects - Part 3: Audio). Además, el codificador de voz AMR-WB + reemplaza los vectores de cuantificación de vectores (vectores VQ) cuantificado a cero con un vector aleatorio de ruido, donde cada valor espectral complejo tiene una amplitud constante, pero una fase aleatoria. La amplitud está controlada por un valor de ruido transmitido con el flujo de bits. Los detalles sobre el codificador de voz AMR-WB + pueden, por ejemplo, encontrarse 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".

45 **[0007]** Además, el documento EP 1 395 980 B1 describe un concepto de codificación de audio. La publicación describe un medio mediante el cual bandas de frecuencia de información seleccionadas de una señal de audio original, que son audibles, pero que son menos relevantes perceptualmente, no necesitan ser codificadas, sino que se pueden sustituir por un parámetro de relleno de ruido. Esas bandas de señal que tienen contenido, que es perceptualmente más relevante están, por el contrario, completamente codificadas. Los bits de codificación se guardan de esta manera sin dejar huecos en el espectro de frecuencia de la señal recibida. El parámetro de ruido de relleno es una medida del valor RMS de la señal 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 a inyectar en la banda de frecuencia en cuestión.

50 **[0008]** Otros enfoques proporcionan una inserción de ruido no guiada en el decodificador, teniendo en cuenta la tonalidad del espectro transmitido.

60 **[0009]** Sin embargo, los conceptos convencionales típicamente conllevan el problema de que o bien comprenden una resolución pobre con respecto a la granularidad del ruido de llenado, que típicamente degrada la impresión de audición, o requieren una cantidad comparativamente grande de información lateral de ruido de llenado, que requiere una velocidad de bits extra.

65

**[0010]** US 4 956 871 describe una disposición de codificación de discurso sub-banda, que divide el espectro de voz en subbandas y asigna bits para codificar las muestras de intervalos de tiempo de cada sub-banda en respuesta a las energías de voz de las sub-bandas. Las muestras de sub-banda se cuantifican de acuerdo con la asignación de bits de energía sub-banda y muestras cuantificadas de marco de tiempo y se codifican señales de energía de voz. Se genera una señal representativa de la diferencia residual entre cada muestra de voz de intervalo tiempo de la sub-banda y la muestra de voz cuantificada correspondiente de la subbanda. La calidad de la señal de sub-banda codificada se mejora mediante la selección de las sub-bandas con las mayores diferencias residuales, produciendo una señal de vector a partir de la secuencia de señales de diferencia residuales de cada sub-banda seleccionada, y haciendo coincidir la señal de vector de sub-banda con uno de un conjunto de entradas de libro de códigos gaussianos almacenados para generar un código de bits reducido para la señal de vector seleccionado. Las señales cuantificadas de intervalo de tiempo codificadas, las señales de energía de voz y códigos de bits reducidos para las diferencias residuales seleccionadas se combinan para formar una corriente multiplexada para el patrón de voz del intervalo de marco de tiempo.

**[0011]** El documento "3rd Generation Partnership Project: Technical Specification Group Service and System Aspects; Audio Codec Processing Functions; Extended Adaptive Multi-Rate - Wideband (AMR-WB+) Codec; Transcoding Functions (Release 6)" describe un codificador de banda ancha multi-tasa adaptativo extendido en el sistema 3GPP. El documento describe el cartografiado detallado para bloques de entrada de muestras monofónicas o estereofónicas en formato de 16 bit uniforme PCM para bloques codificados de desde bloques codificados a bloques de salida de muestras de audio monofónicas o estereofónicas reconstruidas. El esquema de codificación es una extensión del sistema de codificación AMR-WB y se conoce como AMR-WB o AMR-WB + codec extendido. Comprende todos los modos de códec de voz AMR-WB incluyendo VAD / DTX, así como la funcionalidad extendida para la codificación de señales de audio generales, tales como la música, el habla, mixtas, y otras señales.

**[0012]** El documento "Overview of MPEG-4 Audio and its Applications in Mobile Communications" de J. Herre y B. Grill (Publicado en los procedimientos de the International Conference on Comunicación Technology, China, Agosto 21-25, 2012) describe la norma de codificación MPEG-4, que proporciona un conjunto integrado de codificadores de audio con capacidades específicas, incluyendo escalabilidad de tasa de bits y de anchura de banda. La publicación ofrece una introducción a los conceptos de diseños subyacentes de MPEG-4 y proporciona una visión general de la tecnología de codificación de audio MPEG-4 y sus características.

**[0013]** En vista de lo anterior, existe la necesidad de un concepto mejorado de ruido de llenado, que prevé un mejor compromiso entre la impresión de audición alcanzable y la tasa de bits requerida.

#### Resumen de la invención

**[0014]** Una forma de realización según la invención, tal como se establece en la reivindicación independiente 1, crea un codificador para proporcionar una corriente de audio a partir de una representación en el dominio de la transformada de una señal de audio de entrada. El codificador comprende un calculador de error de cuantificación configurado para determinar un error de cuantificación multi-banda sobre una pluralidad de bandas de frecuencia (por ejemplo, sobre una pluralidad de bandas de factor de escala) de la señal de audio de entrada, para las que hay disponible una información de ganancia de banda separada (por ejemplo, factores de escala separados). El codificador también comprende un proveedor de corriente de audio configurado para proporcionar la corriente de audio de modo que la corriente de audio comprende una información espectral que describe un contenido de audio de las bandas de frecuencia y una información que describe el error de cuantificación multi-banda.

**[0015]** El calculador de error de cuantificación está configurado para determinar que componentes espectrales son errores de cuantificación multi-banda sobre una pluralidad de bandas de frecuencia que comprenden cada una al menos un componente espectral (por ejemplo contenedor de frecuencia) cuantificado a un valor diferente de cero evitando a la vez bandas de frecuencia cuyas componentes espectrales están totalmente cuantificadas a cero. Se ha descubierto que una información de error de cuantificación multi-banda es particularmente significativa si se omiten las bandas de frecuencia totalmente cuantificadas a cero del cálculo. En las bandas de frecuencia totalmente cuantificadas a cero, la cuantificación es típicamente muy basta, de modo que la información de error de cuantificación obtenida de esta banda de frecuencia es típicamente no particularmente significativa. Más bien, el error de cuantificación en las bandas de frecuencia psicoacústicamente más relevantes, que no están totalmente cuantificadas a cero, proporciona una información más significativa, que permite un rellenado de ruido adaptado a la audición humana en el lado descodificador.

**[0016]** El codificador descrito más arriba se basa en el hallazgo de que la utilización de una información de error de cuantificación multi-banda trae la posibilidad de obtener una buena impresión auditiva a partir de una cantidad de información lateral comparativamente pequeña. En particular, la utilización de una información de error de cuantificación multi-banda, que cubre una pluralidad de bandas de frecuencia para las que hay disponible una información de ganancia de banda separada, permite un escalado en el lado descodificador de valores de ruido, que se basa en el error de cuantificación multi-banda, en función de la información de ganancia de banda. Por consiguiente, como la información de ganancia de banda está típicamente correlacionada con una relevancia

psicoacústica de las bandas de frecuencia o con una precisión de cuantificación aplicada a las bandas de frecuencia, la información de error de cuantificación multi-banda se ha identificado como información lateral, lo que permite una síntesis de ruido de relleno que proporciona una buena impresión auditiva mientras que mantiene un coste de flujo de información lateral bajo.

5 [0017] Según una realización preferida, el codificador comprende un cuantificador configurado para cuantificar componentes espectrales (por ejemplo, coeficientes espectrales) de diferentes bandas de frecuencia de la representación en el dominio de la transformada empleando diferentes precisiones de cuantificación en función de relevancias psicoacústicas de las diferente bandas de frecuencia para obtener componentes espectrales  
10 cuantificados, en el que las diferentes precisiones de cuantificación están reflejadas por la información de ganancia de banda. También, el proveedor de corriente de audio está configurado para proporcionar la corriente de audio de modo que la corriente de audio comprende una información que describe la información de ganancia de banda (por ejemplo, en la forma de factores de escala) y de modo que la corriente de audio también comprende la información que describe el error de cuantificación multi-banda.

15 [0018] Según una realización preferida, el calculador de error de cuantificación está configurado para determinar el error de cuantificación en el dominio cuantificado, de modo que un escalado, en función de la información de ganancia de banda de la componente espectral, que se realiza antes de una cuantificación de valor entero, se tiene en cuenta. Considerando el error de cuantificación en el dominio cuantificado, se considera la relevancia  
20 psicoacústica de los contenedores espectrales cuando se calcula el error de cuantificación multi-banda. Por ejemplo, para bandas de frecuencia de relevancia perceptual menor, la cuantificación puede ser basta, de modo que el error de cuantificación absoluto (en el dominio no cuantificado) es grande. En contraste, para bandas espectrales de gran relevancia psicoacústica, la cuantificación es fina y el error de cuantificación, en el dominio no cuantificado, es pequeño. Con la finalidad de que el error de cuantificación en las bandas de frecuencia de gran relevancia  
25 psicoacústica y de baja relevancia psicoacústica comparable, tal como para obtener a significativa información de error de cuantificación multi-banda, el error de cuantificación se calcula en el dominio cuantificado (antes que en el dominio no cuantificado) según una realización preferida.

30 [0019] En otra realización preferida, el codificador está configurado para ajustar una información de ganancia de banda (por ejemplo, un factor de escala) de una banda de frecuencia, que está completamente cuantificada a cero (donde todos los contenedores espectrales de la banda de frecuencia están cuantificados a cero) a un valor que representa una relación entre una energía de la banda de frecuencia completamente cuantificada a cero y una energía del error de cuantificación multi-banda. Mediante el establecimiento de un factor de escala de una banda de frecuencia que está cuantificado a cero a un valor bien definido, es posible rellenar la banda de frecuencia  
35 cuantificado a cero con un ruido, de modo que la energía del ruido es al menos aproximadamente igual a la energía de señal original de la banda de frecuencia cuantificada a cero. Mediante la adaptación del factor de escala en el codificador, un descodificador puede tratar la banda de frecuencia cuantificado a cero de la misma manera que cualesquiera otras bandas de frecuencia no cuantificadas a cero, de modo que no hay necesidad de un manejo de excepciones complicado (que típicamente requiere una señalización excepcional). Más bien, mediante la adaptación de la información de ganancia de banda (por ejemplo factor de escala), una combinación del valor de ganancia de banda y la información de error de cuantificación multi-banda permiten una determinación conveniente del ruido de relleno. Otra realización según la invención crea un procedimiento para proporcionar una corriente de audio a partir de una representación en el dominio de la transformada de la señal de audio de entrada, tal como se establece en reivindicación independiente 5.

40 [0020] Otra realización según la invención crea un programa de ordenador para realizar el procedimiento mencionado más arriba, tal como se establece en reivindicación independiente 6.

Breve descripción de las figuras

50 [0021] La figura 1 muestra un diagrama de bloques esquemático de un ejemplo de codificador;  
La figura 2 muestra un diagrama de bloques esquemático de otro ejemplo de codificador;  
Las figuras 3a y 3b muestran un diagrama de bloques esquemático de una codificación de audio avanzada extendida (AAC);  
55 Las figuras 4a y 4b muestran listados de programa de pseudo código de algoritmos ejecutados para la codificación de una señal de audio;  
La figura 5 muestra un diagrama de bloques esquemático de un ejemplo de descodificador;  
La figura 6 muestra un diagrama de bloques esquemático de otro ejemplo de descodificador;  
Las figuras 7a y 7b muestra un diagrama de bloques esquemático de un descodificador de AAC extendido  
60 (*advanced audio coding*);  
La figura 8a muestra una representación matemática de una cuantificación inversa, que puede ser realizada en el descodificador AAC extendido de la figura 7;  
La figura 8b muestra un listado de programa de pseudo código de un algoritmo para cuantificación inversa, que puede ser realizado por el descodificador AAC extendido de la figura 7;  
65 La figura 8c muestra una representación de diagrama de flujo de la cuantificación inversa;

La figura 9 muestra un diagrama de bloques esquemático de un rellenedor de ruido y un re-escalador, que se puede usar en el descodificador AAC extendido de la figura 7;

La figura 10a muestra una representación de código de pseudo programa de un algoritmo, que puede ser ejecutado por el rellenedor de ruido mostrado en la figura 7 o por el rellenedor de ruido mostrado en la figura 9;

5 La figura 10b muestra una leyenda de elementos del código de pseudo programa de la figura 10a;

La figura 11 muestra un diagrama de flujo de un procedimiento, que puede ser implementado en el rellenedor de ruido de la figura 7 o en el rellenedor de ruido de la figura 9;

La figura 12 muestra una ilustración gráfica del procedimiento de la figura 11;

10 Las figuras 13a y 13b muestra representaciones de código de pseudo programa de algoritmos, que pueden ser realizados por el rellenedor de ruido de la figura 7 o por el rellenedor de ruido de la figura 9;

Las figuras 14a to 14d muestran representaciones de elementos de corriente de bits de un ejemplo de corriente de audio; y

La figura 15 muestra una representación gráfica de otro ejemplo de corriente de bits.

15 Descripción detallada de las realizaciones

1. Codificador

20 1.1. Codificador Según la figura 1

[0022] La figura 1 muestra un diagrama de bloques esquemático de un codificador para proporcionar una corriente de audio a partir de la representación en el dominio de la transformada de una señal de audio de entrada.

25 [0023] El codificador 100 de la figura 1 comprende un calculador de error de cuantificación 110 y un proveedor de corriente de audio 120. El calculador de error de cuantificación 110 está configurado para recibir una información 112 relativa a una primera banda de frecuencia, para la que hay disponible una primera información de ganancia de banda de frecuencia, y una información 114 acerca de una segunda banda de frecuencia, para la que hay disponible una segunda información de ganancia de banda de frecuencia. El calculador de error de cuantificación está configurado para determinar un error de cuantificación multi-banda sobre una pluralidad de bandas de frecuencia de la señal de audio de entrada, para las que hay disponible una información de ganancia de banda separada. Por ejemplo, el calculador de error de cuantificación 110 está configurado para determinar el error de cuantificación multibanda sobre la primera banda de frecuencia y la segunda banda de frecuencia empleando 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 multi-banda al proveedor de corriente de audio 120. El proveedor de corriente de audio 120 está configurado para recibir también una información 122 que describe la primera banda de frecuencia y una información 124 que describe la segunda banda de frecuencia. Además, el proveedor de corriente de audio 120 está configurado para proporcionar una corriente de audio 126, de modo que la corriente de audio 126 comprende una representación de la información 116 y también una representación del contenido de audio de la primera banda de frecuencia y de la segunda banda de frecuencia.

40 [0024] Por consiguiente, el codificador 100 proporciona una corriente de audio 126, que comprende un contenido de información, que permite una descodificación eficiente del contenido de audio de la banda de frecuencia empleando un relleno de ruido. En particular, la corriente de audio 126 proporcionada por el codificador trae consigo un buen compromiso entre la tasa de bits y la flexibilidad de descodificación- relleno de ruido.

45 1.2. Codificador Según la figura 2

1.2.1. Visión general del codificador

50 [0025] En lo que sigue, se describirá un codificador de audio mejorado, que se basa en el codificador de audio que se describe en la Norma Internacional ISO/IEC 14496-3: 2005(E), Information Technology - Coding of Audio-Visual Objects - Part 3: Audio, Sub-part 4: General Audio Coding (GA) - AAC, Twin VQ, BSAC.

55 [0026] El codificador de audio 200 de acuerdo con la figura 2 se basa específicamente en el codificador de audio descrito en la norma ISO/IEC 14496-3: 2005(E), Parte 3: Audio, Sub-parte 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).

60 [0027] El codificador de audio 200 puede, por ejemplo, estar configurado para recibir una señal de tiempo de entrada 210 y para proporcionar, sobre la base de esta, una corriente de audio codificada 212. Una ruta de procesamiento de señales puede comprender un sub-muestreador opcional 220, un control de ganancia AAC 222 opcional, un banco de filtros de conmutación de bloque 224, un procesamiento de señal opcional 226, un codificador AAC extendido 228 y un formateador de carga útil corriente de bits 230. Sin embargo, el codificador 200 comprende típicamente un modelo psicoacústico 240.

65

**[0028]** En un caso muy simple, el codificador 200 sólo comprende el banco de filtros / conmutación de bloque 224, el codificador AAC extendido 228, el formateador de carga útil de corriente de bits 230 y el modelo psicoacústico 240, mientras que los otros componentes (en particular, los componentes 220, 222, 226) deben considerarse como meramente opcionales.

**[0029]** En un caso simple, el bloque de conmutación/ banco de filtros 224, recibe la señal de entrada de tiempo 210 (de manera opcional muestreada hacia abajo por el muestreador descendente 220, y de manera opcional el escalado en ganancia por el controlador de ganancia AAC 222), y proporciona, basándose en este, 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 contenedores espectrales de la señal de entrada de tiempo 210. Por ejemplo, el bloque de conmutación/ banco de filtros 224, puede ser configurado para realizar una transformada de coseno discreta modificada (MDCT) para derivar los valores del dominio de la frecuencia a partir de la señal de entrada de tiempo 210. La representación en el dominio de la frecuencia 224a puede ser lógicamente dividida en diferentes bandas de frecuencia, que también se designan como "bandas de factor de escala". Por ejemplo, se supone que el bloque de conmutación/ banco de filtros 224, proporciona valores espectrales (también designados como valores de contenedores de frecuencia) para un valor elevado de contenedores de frecuencias diferentes. El número de contenedores de frecuencias se determina, entre otros, mediante la longitud de una entrada de ventana en el banco de filtros 224, y también en función de la tasa de muestreo (y de bits). Sin embargo, las bandas de frecuencia o bandas de factor de escala definen sub-conjuntos de los valores espectrales proporcionada por el bloque de conmutación/ banco de filtros. Los detalles relativos a la definición de las bandas de factor de escala son conocidos para el experto en la materia, y también se describen en ISO / IEC 14496-3: 2005 (E), Parte 3, Sub-parte 4.

**[0030]** El codificador AAC extendido 228 recibe los valores espectrales 224a proporcionados por el bloque de conmutación/ banco de filtros 224 a partir de la señal de entrada de tiempo 210 (o una versión pre-procesada de la esta) como una información de entrada 228a. Tal como puede verse a partir de la figura 2, la información de entrada 228a del codificador AAC extendido 228 puede derivarse de los valores espectrales 224a empleando uno o más de las etapas de procesamiento del procesado espectral opcional 226. Para más detalles sobre las etapas de procesamiento opcional del pre-procesado espectral 226, se hace referencia a la norma ISO/IEC 14496-3: 2005(E), y a otras normas referenciadas en esta.

**[0031]** El codificador AAC extendido 228 está configurado para recibir la información de entrada 228a en la forma de valores espectrales para una pluralidad de contenedores espectrales y para proporcionar, basándose en esta, una representación codificada cuantificada y sin ruido 228b del espectro. Con este propósito, el codificador AAC extendido 228 puede, por ejemplo, utilizar información derivada de la señal de audio de entrada 210 (o una versión pre-procesada de esta) empleando el modelo psicoacústico 240. Dicho en términos generales, el codificador AAC extendido 228 puede utilizar una información proporcionada por el modelo psicoacústico 240 para decidir que precisión debería aplicarse a la codificación de diferentes bandas de frecuencia (o bandas de factor de escala) de la información de entrada espectral 228a. Por lo tanto, el codificador AAC extendido 228 puede en general adaptar su precisión de cuantificación a diferentes bandas de frecuencia a las características específicas de la señal de entrada de tiempo 210, y también al número de bits disponible. Por lo tanto, el codificador AAC extendido puede, por ejemplo, ajustar sus precisiones de cuantificación, de modo que la información que representa el espectro codificado cuantificado y sin ruido comprende una tasa de bits apropiada (o tasa de bits promedio).

**[0032]** El formateador de carga útil de corriente de bits 230 está configurado para incluir la información 228b que representa los espectros codificados cuantificados y sin ruido en la corriente de audio codificada 212 según una sintaxis predeterminada.

**[0033]** Para más detalles sobre la funcionalidad de los componentes de codificador que se describen aquí, se hace referencia a la norma ISO/IEC 14496-3: 2005(E) (incluyendo el anexo 4.B de esta), y también a la ISO/IEC 13818-7: 2003.

**[0034]** Además, se hace referencia a ISO/IEC 13818-7: 2005, Sub-cláusulas C1 a C9.

**[0035]** Además, se hace referencia específica en cuanto a la terminología a la norma ISO/IEC 14496-3: 2005(E), Parte 3: Audio, Sub-parte 1: Principal.

**[0036]** Además, se hace referencia específica a la norma ISO/IEC 14496-3: 2005(E), Parte 3: Audio, Sub-parte 4: Codificación General de Audio (GA) - AAC, Twin VQ, BSAC.

### 1.2.2. Detalles del codificador

**[0037]** En lo que sigue, se describirán detalles sobre el codificador haciendo referencia a las figuras 3a, 3b, 4a y 4b.

**[0038]** Las figuras 3a y 3b muestran un diagrama de bloques esquemático de un codificador AAC extendido. El descodificador AAC extendido se designa mediante 228 y puede ocupar el lugar del codificador AAC extendido 228

de la figura 2. El codificador AAC extendido 228 está configurado para recibir, como una información de entrada 228a, un vector de magnitudes de líneas espectrales, en el que el vector de líneas espectrales se designa a veces con  $mdctline$  (0..1023). El codificador AAC extendido 228 también recibe una información de umbral codec 228c, que describe un error de energía permitido máximo en un nivel MDCT. La información de umbral codec 228c se proporciona típicamente individualmente para diferentes bandas de factor de escala y se genera empleando el modelo psicoacústico 240. La información de umbral codec 228 se designa a veces con  $X_{min}(sb)$ , en el que el parámetro  $sb$  indica la dependencia de banda del factor de escala. El codificador AAC extendido 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 promedio (designado con bits promedio) y una información de bits adicional (designada con más bits). El codificador AAC extendido 228 también está configurado para recibir una información de banda de factor de escala 228e, que describe, por ejemplo, un número y anchura de bandas de factor de escala.

**[0039]** El codificador AAC extendido comprende un cuantificador de valor espectral 310, que está configurado para proporcionar un vector 312 de valores cuantificados de líneas espectrales, que también se designa con  $xquant$  (0..1023). El cuantificador de valor espectral 310, que incluye un escalado, también está configurado para proporcionar una información de factor de escala 314, que puede representar un factor de escala para cada banda de factor de escala y también una información común de factor de escala. Además, el cuantificador de valor espectral 310 puede ser 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. Realmente, el cuantificador de valor espectral 310 está configurado para cuantificar diferentes valores espectrales del vector 228a con diferentes precisiones que dependen de la relevancia psicoacústica de los diferentes valores espectrales. Con este propósito, el cuantificador de valor espectral 310 escala los valores espectrales del vector 228a empleando diferentes factores de escala dependientes de las bandas de factor de escala y cuantifica los valores espectrales escalados resultantes. Típicamente, los valores espectrales asociados con bandas de factor de escala importantes desde el punto de vista psico-acústico se escalarán con grandes factores de escala, de modo que los valores espectrales escalados de bandas de factor de escala de importancia desde el punto de vista psico-acústico cubren un gran rango de valores. En contraste, los valores espectrales de bandas de factor de escala menos importantes desde el punto de vista psico-acústico se escalan con factores de escala menores, de modo que los valores espectrales escalados de las bandas de factor de escala menos importantes desde el punto de vista psico-acústico cubren solamente un rango de valores menor. Los valores espectrales escalados se cuantifican entonces, por ejemplo, a un valor integral. En esta cuantificación, varios de los valores espectrales escalados de las bandas de factor de escala menos importantes desde el punto de vista psico-acústico están cuantificados a cero, porque los valores espectrales de las bandas de factor de escala menos importantes desde el punto de vista psico-acústico se escalan solamente con un factor de escala pequeño.

**[0040]** Como resultado, se puede decir que los valores espectrales de bandas de factor de escala más relevantes desde el punto de vista psico-acústico están cuantificados con gran precisión (porque las líneas espectrales escaladas de dichas bandas de factor de escala más relevantes cubren un amplio intervalo de valores y, por lo tanto, varias etapas de cuantificación), mientras que los valores espectrales de las bandas de factor de escala menos importantes desde el punto de vista psico-acústico están cuantificados con menos precisión de cuantificación (porque los valores espectrales escalados de las bandas de factor de escala menos importantes cubren un menor intervalo de valores y están, por lo tanto, cuantificados a menos etapas de cuantificación diferentes).

**[0041]** El cuantificador de valor espectral 310 está típicamente configurado para determinar factores de escalado apropiados empleando el umbral codec 228c y la información de número de bits 228d. Típicamente, el cuantificador de valor espectral 310 también está configurado para determinar los factores de escala apropiados él mismo. Los detalles relativos a una posible implementación del cuantificador de valor espectral 310 se describen en ISO/IEC 14496-3: 2001, Capítulo 4.B.10. Además, la implementación del cuantificador de valor espectral es bien conocida para un experto en la materia de codificación MPEG4.

**[0042]** El codificador AAC extendido 228 también comprende un calculador de error de cuantificación multi-banda 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 factor de escala 314. El calculador de error de cuantificación multi-banda 330 está, por ejemplo, configurado para determinar una desviación entre una versión escalada no-cuantificada de los valores espectrales del vector 228a (por ejemplo, escalada empleando una operación de escalado no lineal y un factor de escala) y una versión escalada y cuantificada (por ejemplo, escalada empleando una operación de escalado no lineal y un factor de escala, y cuantificada empleando una operación de redondeo "entera") de los valores espectrales. Además, el calculador de error de cuantificación multi-banda 330 puede ser configurado para calcular un error de cuantificación promedio sobre una pluralidad de bandas de factor de escala. Hay que destacar que el calculador de error de cuantificación multi-banda 330 preferentemente calcula el error de cuantificación multi-banda in un dominio cuantificado (más precisamente en un dominio escalado psico-acústicamente), de modo que un error de cuantificación en bandas de factor de escala relevantes psico-acústicamente se enfatiza en peso cuando se compara con un error de cuantificación en bandas de factor de escala

psico-acústicamente menos relevantes. Unos detalles relativos al funcionamiento del calculador de error de cuantificación multi-banda se describirán más adelante haciendo referencia a las figuras 4a y 4b.

5 **[0043]** El codificador AAC extendido 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 multi-banda 332, proporcionada por el calculador de error de cuantificación multi-banda 340. El adaptador de factor de escala 340 está configurado para identificar bandas de factor de escala, que están "cuantificadas a cero", es decir bandas de valor de factor de escala para las que todos los valores espectrales (o líneas espectrales) están cuantificados a cero. Para estas bandas de factor de escala cuantificadas totalmente 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 factor de escala cuantificado totalmente a cero a un valor, que representa una relación entre una energía residual (antes de la cuantificación) de la banda de factor de escala respectiva y una energía del error de cuantificación multi-banda 332. Por consiguiente, el adaptador de factor de escala 340 proporciona factores de escala adaptados 342. Hay que destacar que tanto los factores de escala proporcionados por el cuantificador de valor espectral 310 y 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 literatura y también en esta solicitud. Unos detalles relativos al funcionamiento del adaptador de factor de escala 340 se describirán más adelante haciendo referencia a las figuras 4a y 4b.

20 **[0044]** El codificador AAC extendido 228 también comprende un codificador sin ruido 350, que se explica, por ejemplo, en ISO/IEC 14496-3: 2001, Capítulo 4.B.11. En resumen, el codificador sin ruido 350 recibe el vector de valores cuantificados de líneas espectrales (también designados como "valores cuantificados de los espectros") 312, la representación entera 342 de los factores de escala (ya sea tal como se proporcionan por el cuantificador de valor espectral 310, o tal como se adaptan por el adaptador de factor de escala 340), y también un parámetro de relleno de ruido 332 (por ejemplo, en la forma de un información de nivel de ruido) proporcionado por el calculador de error de cuantificación multi-banda 330.

30 **[0045]** El codificador 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. Unos detalles relativos 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 ISO/IEC 14496-3: 2001. El codificador 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 codificada de factor de escala 354. El codificador sin ruido 350 también comprende una codificación de parámetro de relleno de ruido 350c para codificar el o los parámetros de relleno de ruido 332, para obtener uno o más parámetros codificados de relleno de ruido 356. En consecuencia, el codificador AAC extendido proporciona una información que describe los espectros codificados cuantificados y sin ruido, en los que esta información comprende valores cuantificados y codificados de las líneas espectrales, información codificada de factor de escala y información de parámetros de relleno de ruido codificado.

40 **[0046]** A continuación, se describirán la funcionalidad del calculador de error de cuantificación multi-banda 330 y del adaptador de factor de escala 340, que son componentes clave del codificador AAC extendido 228, haciendo referencia a las figuras 4a y 4b. Con este propósito, la figura 4a muestra un listado de programa de un algoritmo realizado por el calculador de error de cuantificación multi-banda 330 y el adaptador de factor de escala 340.

45 **[0047]** Una primera parte del algoritmo, representada por las líneas 1 a 12 del pseudo código de la figura 4a, comprende a cálculo de a mean error de cuantificación, que se realiza por el calculador de error de cuantificación multi-banda 330. El cálculo dl error medio de cuantificación se realiza, por ejemplo, sobre todas las bandas de factor de escala, excepto para aquellas que están cuantificadas a cero. Si una banda de factor de escala está totalmente cuantificada a cero (es decir que todas las líneas espectrales de la banda de factor de escala están cuantificadas a cero), dicha banda de factor de escala se obvia para el cálculo del error medio de cuantificación. Si, sin embargo, una banda de factor de escala no está totalmente cuantificada a cero (es decir comprende al menos un una línea espectral, que no está cuantificada a cero), todas las líneas espectrales de dicha banda de factor 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 precisamente, en un dominio escalado). El cálculo de una contribución al error promedio puede verse en la línea 7 del pseudo código de la figura 4a. En particular, la línea 7 muestra la contribución de una única línea espectral al error promedio, en el que el promedio se realiza sobre todas las líneas espectrales (en el que nLines indica el número del total de líneas consideradas).

60 **[0048]** Como puede verse en la línea 7 del pseudo código, 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 cuantificado y un valor de magnitud de línea espectral cuantificado. En el valor de magnitud de línea espectral no cuantificado, el valor de magnitud "line" (que puede ser igual a mdctline) se escala de manera no lineal empleando una función potencia ( $\text{pow}(\text{line}, 0.75) = \text{line}^{0.75}$ ) y empleando un factor de escala (por ejemplo un factor de escala 314 proporcionado por el cuantificador de valor espectral 310). En el cálculo del valor de magnitud de línea espectral cuantificado, el valor de magnitud de línea espectral "line" puede ser escalado de manera no lineal empleando las funciones de potencia mencionadas y escalado empleando el factor de escala mencionado. El resultado de estos



escalados lineal y no lineal puede ser cuantificado empleando un operador entero "(INT)". Empleando el cálculo tal como se indica en la línea 7 del pseudo código, el impacto diferente de la cuantificación se consideran las bandas de frecuencia psico-acústicamente más importantes y las menos importantes desde el punto de vista psico-acústico.

5 **[0049]** Tras el cálculo del error (promedio) de cuantificación multi-banda (avgError), el error promedio de cuantificación puede de manera opcional ser cuantificado, tal como se muestra en las líneas 13 y 14 del pseudo código. Hay que destacar que la cuantificación del error de cuantificación multi-banda tal como aquí se muestra está específicamente adaptado al intervalo de valores esperado y a las características estadísticas del error de cuantificación, de modo que el error de cuantificación se puede representar con pocos bits. Sin embargo, se pueden  
10 aplicar otras cuantificaciones del error de cuantificación multi-banda.

**[0050]** Una tercera parte del algoritmo, que se representa en las líneas 15 a 25, puede ser ejecutado por el adaptador de factor de escala 340. La tercera parte del algoritmo sirve para ajustar factores de escala de banda de factores de escala de frecuencia, que se han cuantificado a cero totalmente, a un valor bien definido, que permite un simple relleno de ruido, que conlleva una buena impresión auditiva. La tercera parte del algoritmo comprende de manera opcional una cuantificación inversa del nivel de ruido (por ejemplo representado por el error de cuantificación multi-banda 332). La tercera parte del algoritmo también comprende un cálculo de un valor de factor de escala de sustitución para bandas de factor de escala cuantificadas a cero (mientras que los factores de escala de bandas de factor de escala no cuantificados a cero permanecerán sin cambios). Por ejemplo, el valor de factor de escala de sustitución para una determinada banda de factor de escala ("band") se calcula empleando la ecuación mostrada en la línea 20 del algoritmo de la figura 4a. En esta ecuación, "(INT)" representa un operador entero, "2.f" representa el número "2" en una representación de punto flotante, "log" designa un operador logaritmo, "energy" designa una energía de la banda de factor de escala bajo consideración (antes de la cuantificación), "(float)" designa un operador de punto flotante, "sfbAnchura" designa una anchura de la banda de factor de escala determinada en términos de líneas espectrales (o contenedores espectrales), y "ruidoVal" designa un valor de ruido que describe el error de cuantificación multi-banda. En consecuencia, el factor de escala de sustitución describe una relación entre una energía por contenedor de energía promedio (energía/sfbanchura) de las bandas de factor de escala determinadas bajo consideración, y una energía (ruidoVal2) del error de cuantificación multi-banda.

### 30 1.2.3. Conclusión con respecto al Codificador

**[0051]** Las realizaciones según la invención crean un codificador que tiene un nuevo tipo de cálculo de nivel de ruido. En realizaciones preferidas según la invención, tal como se establece en las reivindicaciones dependientes 3 y 4, el error de cuantificación multi-banda que representa el nivel de ruido se calcula en el dominio cuantificado.

**[0052]** El cálculo del error de cuantificación en el dominio cuantificado conlleva ventajas significativas, por ejemplo, porque se considera la relevancia psicoacústica de diferentes bandas de frecuencia (bandas de factor de escala). El error de cuantificación por línea(es decir per línea espectral, o contenedor espectral) en el dominio cuantificado está típicamente en el intervalo [-0.5; 0.5] (1 nivel de cuantificación) con un error absoluto promedio de 0.25 (para valores de entradas con distribución normal que suelen ser mayores que 1). Empleando un codificador, que proporciona información sobre un error de cuantificación multi-banda, las ventajas de relleno de ruido en el dominio cuantificado pueden explotarse en un codificador, tal como se describirá más adelante.

**[0053]** El cálculo del nivel de ruido y detección de sustitución de ruido en el codificador puede comprender las siguientes etapas:

- Detectar y marcar bandas espectrales que pueden ser reproducidas de manera perceptualmente equivalente en el descodificador por sustitución de ruido. Por ejemplo, pueden ser verificadas una tonalidad o una medida de planitud espectral para este fin;
- 50 • Calcular y cuantificar el error de cuantificación media (que puede calcularse sobre todas las bandas de factor de escala no cuantificadas a cero); y
- Calcular el factor de escala (scf) para una band cuantificada a cero de modo que el ruido introducido (descodificador) coincide con la energía original.

55 **[0054]** Una cuantificación de nivel de ruido apropiada puede ayudar a producir el número de bits requerido para transportar la información que describe el error de cuantificación multi-banda. Por ejemplo, el nivel de ruido puede ser cuantificado en 8 niveles de cuantificación en el dominio logarítmico, considerando la percepción humana de los bajos. Por ejemplo, puede utilizarse el algoritmo mostrado en la figura 4b, en el que "(INT)" designa un operador entero, en el que "LD" designa una operación logarítmica para una base de 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. Descodificador

### 65 2.1. Descodificador según la figura 5

**[0055]** La figura 5 muestra un diagrama de bloques esquemático de un ejemplo de descodificador. El descodificador 500 está configurado para recibir una información de audio codificada, por ejemplo, en la forma de una corriente de audio codificada 510, y para proporcionar, basándose en esta, una representación descodificada de la señal de audio, por ejemplo, a partir de componentes espectrales 522 de una primera banda de frecuencia y componentes espectrales 524 de una segunda banda de frecuencia. El descodificador 500 comprende un relleno de ruido 520, que está configurado para recibir una representación 522 de componentes espectrales de una primera banda de frecuencia, a la que una primera información de ganancia de banda de frecuencia está asociada, y una representación 524 de componentes espectrales de una segunda banda de frecuencia, a la que una segunda información de ganancia de banda de frecuencia está asociada. Además, el relleno de ruido 520 está configurado para recibir una representación 526 de un valor de intensidad de ruido multi-banda. Además, el relleno de ruido está configurado para introducir ruido en los componentes espectrales (por ejemplo en valores de línea espectral o valores de contenedor espectral) de una pluralidad de bandas de frecuencia a la que está asociada una información de ganancia de banda de frecuencia separada (por ejemplo en la forma de factores de escala) a partir del valor de intensidad de ruido multi-banda común 526. Por ejemplo, el relleno de ruido 520 puede ser configurado para introducir ruido en las componentes espectrales 522 de la primera banda de frecuencia para obtener los componentes espectrales 512 de la primera banda de frecuencia afectados por ruido, y también para introducir ruido en las componentes espectrales 524 de la segunda banda de frecuencia para obtener las componentes espectrales 514 de la segunda banda de frecuencia afectadas por ruido.

**[0056]** Aplicando el ruido descrito por un único valor de intensidad de ruido multi-banda 526 a componentes espectrales de diferentes bandas de frecuencia a la que hay asociada información diferente de ganancia de banda de frecuencia, se puede introducir ruido en las diferentes bandas de frecuencia de manera muy precisa, considerando la diferente relevancia psico-acústica de diferentes bandas de frecuencia, que se expresa mediante la información de ganancia de banda de frecuencia. Por lo tanto, el descodificador 500 es capaz de realizar un relleno de ruido ajustado en el tiempo a partir de una información lateral de relleno de ruido muy pequeña (eficiente en bits).

## 2.2. Descodificador según la figura

### 6.2.2.1. Visión general del descodificador

**[0057]** La figura 6 muestra un diagrama de bloques esquemático de otro ejemplo de descodificador 600.

**[0058]** El descodificador 600 es similar al descodificador descrito en ISO/IEC 14496.3: 2005 (E), de modo que se hace referencia a esta Norma Internacional. El descodificador 600 está configurado para recibir una corriente de audio codificada 610 y para proporcionar, basándose en esta, señales temporales de salida 612. La corriente de audio codificada puede comprender parte o toda la información descrita en ISO/IEC 14496.3: 2005 (E), y de manera adicional comprende información que describe un valor de intensidad de ruido multi-banda. El descodificador 600 también comprende un desformateador de carga útil de corriente de bits 620, que está configurado para extraer de la corriente de audio codificada 610 una pluralidad de parámetros de audio codificados, algunos de los cuales se describirán con más detalle más adelante. El descodificador 600 también comprende un descodificador “de codificación de audio avanzada” extendido (AAC) 630, cuya funcionalidad se describirá en detalle, haciendo referencia a las figuras 7a, 7b, 8a to 8c, 9, 10a, 10b, 11, 12, 13a y 13b. El descodificador AAC extendido 630 está configurado para recibir una información de entrada 630a, que comprende, por ejemplo, una información de línea espectral cuantificada y codificada, una información codificada de factor de escala y una información de parámetros de relleno de ruido codificado. Por ejemplo, la información de entrada 630a del codificador AAC extendido 630 puede ser idéntica a la información de salida 228b proporcionada por el codificador AAC extendido 220a descrito con referencia a la figura 2.

**[0059]** El descodificador AAC extendido 630 puede ser configurado para proporcionar, a partir de la información de entrada 630a, una representación 630b de un espectro escalado e inversamente cuantificado, por ejemplo, en la forma de valores de línea espectral escalados e inversamente cuantificados para una pluralidad de contenedores de frecuencias (por ejemplo, para 1024 contenedores de frecuencias).

**[0060]** 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 puede ser utilizado como alternativa al decodificador de espectro extendido AAC 630 en algunos casos.

**[0061]** El decodificador 600 puede comprender opcionalmente un procesamiento de espectro 640, que está configurado para procesar la información de salida 630b del descodificador AAC extendido 630 con el fin de obtener una información de entrada 640a de un conmutador de bloque / banco de filtros 640. El procesamiento espectral opcional 630 puede comprender uno o más, o incluso la totalidad, de las funcionalidades M/S, PNS, predicción, intensidad, predicción a largo plazo, acoplamiento dependiente de la conmutación, TNS, acoplamiento dependiente de la conmutación, cuyas funcionalidades se describen detalle en la norma ISO/IEC 14493.3: 2005 (E) y los documentos ahí referenciados. Si, sin embargo, el procesamiento espectral 630 se omite, la información de salida 630b del descodificador AAC extendido 630 puede servir directamente como información de entrada 640a del bloque

de conmutación/ banco de filtros 640. Por lo tanto, el descodificador AAC extendido 630 puede proporcionar, como información de salida 630b, espectros escalados e inversamente cuantificados. El bloque de conmutación/ banco de filtros 640 utiliza, como información de entrada 640a, los espectros inversamente cuantificados (de manera opcional pre-procesados) y proporciona, basándose en estos, una o más señales de audio reconstruidas en el dominio del tiempo como una información de salida 640b. El conmutador de bloque / banco de filtros puede, por ejemplo, estar configurado para aplicar la inversa de del cartografiado de frecuencia que se llevó a cabo en el codificador (por ejemplo, in el bloque de conmutación/ banco de filtros 224). Por ejemplo, se puede utilizar una transformada inversa de coseno discreta modificada (IMDCT) por el banco de filtros. Por ejemplo, el IMDCT puede ser configurado para soportar ya sea 120, 128, 480, 512, 960 o 1024, o cuatro conjuntos de 32 o 256 coeficientes espectrales.

**[0062]** Para más detalles, se hace referencia, por ejemplo, a la norma internacional ISO/IEC 14496-3: 2005 (E). El descodificador 600 puede de manera opcional comprender además un control de ganancia AAC 650, un descodificador SBR 652 y un acoplamiento conmutado de manera independiente 654, para derivar la señal temporal de salida 612 de la señal de salida 640b del bloque de conmutación/ banco de filtros 640.

**[0063]** Sin embargo, la señal de salida 640b del bloque de conmutación/ banco de filtros 640 también puede servir como la señal temporal de salida 612 en ausencia de la funcionalidad 650, 652, 654.

### 2.2.2. Detalles del Descodificador Extendido AAC

**[0064]** A continuación, se describirán unos detalles relativos al descodificador AAC extendido, haciendo referencia a las figuras 7a y 7b. Las figuras 7a y 7b muestran un diagrama de bloques esquemático de del descodificador AAC 630 de la figura 6 en combinación con el desformateador de carga útil de corriente de bits 620 de la figura 6.

**[0065]** El desformateador de carga útil de corriente de bits 620 recibe una corriente de audio descodificada 610, que puede, por ejemplo, comprender una corriente de datos de audio codificada que comprende un elemento de sintaxis titulado "acrawdata block", que es un bloque de datos en bruto de codificador de audio. Sin embargo, el formateador de carga útil de corriente de bits 620 está configurado para proporcionar al descodificador AAC extendido 630 un espectro codificado cuantificado y sin ruido o una representación, que comprende una información de línea espectral cuantificada y codificada aritméticamente 630aa (por ejemplo designada como acspectraldata), una información de factor de escala 630ab (por ejemplo designada como scalefactor data) y una información de parámetro de rellenado de ruido 630ac. La información de parámetro de rellenado de ruido 630ac comprende, por ejemplo, un valor de ruido de desplazamiento (designado como noiseoffset) y un nivel de valor de ruido (designado con noiselevel).

**[0066]** En relación con el descodificador AAC extendido, hay que destacar que el descodificador AAC extendido 630 es muy similar al descodificador AAC de la Norma Internacional ISO/IEC 14496-3: 2005 (E), de modo que se hace referencia a la descripción detallada en dicha norma.

**[0067]** El descodificador AAC extendido 630 comprende un descodificador de factor de escala 740 (también designado como herramienta de descodificación sin ruido de factor de escala), que está configurado para recibir la información de factor de escala 630ab y para proporcionar basándose en esta, una representación descodificada entera 742 de los factores de escala (que también se designa como sf[g][sfb] o scf[g][sfb]). Relativa al descodificador de factor de escala 740, se hace referencia a ISO/IEC 14496-3: 2005, Capítulos 4.6.2 y 4.6.3. Hay que destacar que la representación descodificada entera 742 de los factores de escala refleja una precisión de cuantificación con la que diferentes bandas de frecuencia (también designados como bandas de factor de escala) de una señal de audio están cuantificadas. Unos factores de escala mayores indican que las bandas de factor de escala se han cuantificado con alta precisión, y factores de escala menores indican que las bandas de factor de escala correspondientes se han cuantificado con baja precisión.

**[0068]** El descodificador AAC extendido 630 también comprende un descodificador espectral 750, que está configurado para recibir la información de línea espectral cuantificada y codificada por entropía 630aa (por ejemplo codificada mediante Huffman o codificada aritméticamente) y para proporcionar, basándose en esta, valores cuantificados 752 de los uno o más espectros (por ejemplo designados como xacquant o xquant). En relación con el descodificador espectral, se hace referencia, por ejemplo, a la sección 4.6.3 de la Norma Internacional mencionada más arriba. Sin embargo, unas alternativas de implementaciones de los espectros descodificador pueden aplicarse naturalmente. Por ejemplo, el descodificador Huffman de la ISO/IEC 14496-3: 2005 puede ser sustituido por un descodificador aritmético si la información de línea espectral 630aa está codificada aritméticamente.

**[0069]** El descodificador AAC extendido 630 también comprende un cuantificador inverso 760, que puede ser un cuantificador inverso no-uniforme. Por ejemplo, el cuantificador inverso 760 puede proporcionar valores espectrales inversamente cuantificados no escalados 762 (por ejemplo, designados mediante xacinvquant, o xinquant). Por ejemplo, el cuantificador inverso 760 puede comprender la funcionalidad descrita en ISO/IEC 14496-3: 2005, Capítulo 4.6.2. De manera alternativa, el cuantificador inverso 760 puede comprender la funcionalidad descrita con referencia a las figuras 8a to 8c.

**[0070]** El descodificador AAC extendido 630 también comprende un rellenedor de ruido 770 (también designado como herramienta de relleno de ruido), que recibe la representación descodificada entera 742 de los factores de escala a partir del descodificador de factor de escala 740, los valores espectrales cuantificados inversamente no escalados 762 del cuantificador inverso 760 y el información de parámetros de relleno de ruido 630ac del desformateador de carga útil de corriente de bits 620. El rellenedor de ruido está configurado para proporcionar, basándose en este, la representación modificada (típicamente entera) 772 de los factores de escala, que también se designan aquí con  $sf[g][sfb]$  o  $scf[g][sfb]$ . El rellenedor de ruido 770 también está configurado para proporcionar valores espectrales inversamente cuantificados no escalados 774, también designados como  $xacinvquant$  o  $xinvquant$  a partir de su información de entrada. Unos detalles relativos a la funcionalidad del rellenedor de ruido se describirán más adelante, haciendo referencia a las figuras 9, 10a, 10b, 11, 12, 13a y 13b.

**[0071]** El descodificador AAC extendido 630 también comprende un re-escalador 780, que está configurado para recibir la representación modificada entera de los factores de escala 772 y valores espectrales cuantificados inversamente no escalados 774, y para proporcionar, basándose en este, valores espectrales cuantificados inversamente escalados 782, que también se pueden designar como  $xrescal$ , y que pueden servir como la información de salida 630b del descodificador AAC extendido 630. El re-escalador 780 puede, por ejemplo, comprender la funcionalidad tal como se describe en ISO/IEC 14496-3: 2005, Capítulo 4.6.2.3.3.

### 2.2.3. Cuantificador inverso

**[0072]** A continuación, se describirá la funcionalidad del cuantificador inverso 760, haciendo referencia a las figuras 8a, 8b y 8c. La figura 8a muestra una representación se describirá valores espectrales cuantificados inversamente no escalados 762 a partir de los valores espectrales cuantificados 752. En las ecuaciones alternativas de la figura 8a, "sign(.)" designa un operador de signo, y "." designa un operador de valor absoluto. La figura 8b muestra un código de pseudo programa que representa la funcionalidad del cuantificador inverso 760. Como puede verse, la cuantificación inversa de acuerdo con la regla de asignación matemática que se muestra en la figura 8a se lleva a cabo para todos los grupos de ventanas (designadas por la variable que se desplaza  $g$ ), para todas las bandas de factor de escala (designadas por la variable que se desplaza  $sfb$ ), para todas las ventanas (designadas por el índice que se desplaza  $win$ ) y todas las líneas espectrales (o contenedores espectrales) (designados por la variable que se desplaza  $bin$ ). La figura 8c muestra una representación de diagrama de flujo del algoritmo de la figura 8b. Para bandas de factor de escala por debajo de una banda de factor de escala máxima predeterminada (designada con  $maxsfb$ ), unos valores espectrales cuantificados inversamente no escalados se obtienen como una función de valores espectrales cuantificados no escalados. Se aplica una regla de cuantificación inversa no lineal.

### 2.2.4 Rellenedor de ruido

#### 2.2.4.1. Rellenedor de ruido según las figuras 9 a 12

**[0073]** La figura 9 muestra un diagrama de bloques esquemático de un rellenedor de ruido 900. El rellenedor de ruido 900 puede, por ejemplo, tomar el lugar del rellenedor de ruido 770 descrito con referencia a las figuras 7A y 7B.

**[0074]** El rellenedor de ruido 900 recibe la representación descodificada entera 742 de los factores de escala, que pueden ser considerados como valores de ganancia de banda de frecuencia. El rellenedor de ruido 900 también recibe valores espectrales cuantificados inversamente no escalados 762. Además, el rellenedor de ruido 900 recibe el información de parámetros de relleno de ruido 630ac, por ejemplo, que comprende parámetros de relleno de ruido  $noisevalue$  y  $noiseoffset$ . El rellenedor de ruido 900 proporciona además la representación modificada entera 772 de los factores de escala y valores espectrales cuantificados inversamente no escalados 774. El rellenedor de ruido 900 comprende un detector de líneas espectrales cuantificadas a cero 910, que está configurado para determinar si una línea espectral (o contenedor espectral) está cuantificado a cero (y que posiblemente satisface además requisitos de relleno de ruido). Con este propósito, el detector de líneas espectrales cuantificadas a cero 910 recibe directamente los espectros inversamente cuantificados no escalados 762 como información de entrada. El rellenedor de ruido 900 también comprende sustituidor de línea espectral selectivo 920, que está configurado para sustituir de manera selectiva valores espectrales de la información de entrada 762 por valores de sustitución de línea espectral 922 en función de la decisión del detector de líneas espectrales cuantificadas a cero 910. Por lo tanto, si el detector de líneas espectrales cuantificadas a cero 910 indica que una determinada línea espectral de la información de entrada 762 debería ser sustituida por un valor de sustitución, entonces el sustituidor de línea espectral selectivo 920 sustituye la determinada línea espectral por el valor de sustitución de línea espectral 922 para obtener la información de salida 774. De otro modo, el sustituidor de línea espectral selectivo 920 reenvía el determinado valor de línea espectral sin cambios para obtener la información de salida 774. El rellenedor de ruido 900 también comprende un modificador de factor de escala selectivo 930, que está configurado para modificar de manera selectiva factores de escala de la información de entrada 742. Por ejemplo, el modificador de factor de escala selectivo 930 está configurado para aumentar factores de escala de banda de frecuencia de factor de escala, que se han cuantificado a cero en un valor predeterminado, que se designa como "noiseoffset". Por lo tanto, en la información de salida 772, los factores de escala de bandas de frecuencia cuantificadas a cero aumentan cuando se comparan con los valores de factor de escala correspondientes dentro de la información de entrada 742. En

contraste, los valores correspondientes de factor de escala de bandas de frecuencia de factor de escala, que no están cuantificados a cero, son idénticos en la información de entrada 742 y en la información de salida 772.

[0075] Para determinar si una banda de frecuencia de factor de escala está cuantificado a cero, el relleno de ruido 900 también comprende un detector de banda cuantificada a cero 940, que está configurado para controlar el modificador de factor de escala selectivo 930 proporcionando una señal de "habilitar modificación factor de escala" o etiqueta 942 a partir de la información de entrada 762. Por ejemplo, el detector de banda cuantificada a cero 940 puede proporcionar una señal o etiqueta que indica la necesidad de un aumento de un factor de escala al modificador de factor de escala selectivo 930 si todos los contenedores de frecuencias (también designados como contenedores espectrales) de una banda de factor de escala están cuantificados a cero.

[0076] Cabe señalar aquí que el modificador de factor de escala selectivo también puede tomar la forma de un sustituidor de factor de escala selectivo, que está configurado para establecer los factores de escala de las bandas de factor de escala cuantificadas completamente a cero a un valor predeterminado, independientemente de la información de entrada 742.

[0077] En lo que sigue, se describirá un re-escalador 950, que puede tomar la función del re-escalador 780. El re-escalador 950 está configurado para recibir la representación entera modificada 772 de los factores de escala proporcionados por el relleno de ruido y también para los valores espectrales inversamente cuantificados no escalados 774 proporcionados por el relleno de ruido. El re-escalador 950 comprende un calculador de ganancia de factor de escala 960, que está configurado para recibir una representación entera del factor de escala por banda de factor de escala y para proporcionar un valor de ganancia por banda de factor de escala. Por ejemplo, el calculador de ganancia de factor de escala 960 puede ser configurado para calcular un valor de ganancia 962 para banda de frecuencia *i*-ésima a partir de una representación modificada entera 772 del factor de escala para la banda de factor de escala *i*-ésima. Por lo tanto, el calculador de ganancia de factor de escala 960 proporciona valores de ganancia individuales para las diferentes bandas de factor de escala. El re-escalador 950 también comprende un multiplicador 970, que está configurado para recibir los valores de ganancia 962 y los valores espectrales inversamente cuantificados no escalados 774. Hay que destacar que cada uno de los valores espectrales inversamente cuantificados no escalados 774 está asociado a una banda de frecuencia de factor de escala (sfb). Por consiguiente, el multiplicador 970 está configurado para escalar cada uno de los valores espectrales inversamente cuantificados no escalados 774 con un valor de ganancia correspondiente asociado con la misma banda del factor de escala. En otras palabras, todos los valores espectrales inversamente cuantificados no escalados 774 asociados con una banda de factor de escala determinada se escalan con el valor de ganancia asociado con el banda de factor de escala determinado. Por consiguiente, los valores espectrales inversamente cuantificados no escalados asociados con diferentes bandas de factor de escala se escalan típicamente con diferentes valores de ganancia asociados con las diferentes bandas de factor de escala.

[0078] Por lo tanto, diferentes de los valores espectrales inversamente cuantificados no escalados se escalan con valores de ganancia diferentes que dependen de qué bandas de factor de escala tienen asociados.

#### Representación de Código de pseudo programa

[0079] A continuación, se describirá la funcionalidad del relleno de ruido 900 haciendo referencia a las figuras 10A y 10B, que muestra una representación de código de pseudo programa (figura 10A) y una leyenda correspondiente (figura 10B). Los comentarios comienzan con "--".

[0080] El algoritmo de relleno de ruido representado por el listado de programa de pseudo código de la figura 10 comprende una primera parte (líneas 1 a 8) de derivación de un valor de ruido (noiseVal) a partir de una representación de nivel de ruido (noiselevel). Además, se deriva un desplazamiento de ruido (noiseoffset). La derivación del valor de ruido a partir del nivel de ruido comprende un escalado no lineal, en el que el valor de ruido se calcula según

$$\text{noiseVal} = 2^{(\text{noise\_level}-14)/3}$$

Además, se realiza un desplazamiento del intervalo de valores de desplazamiento de ruido de modo que el valor de desplazamiento de ruido del intervalo desplazado puede tomar valores positivos y negativos.

[0081] Una segunda parte del algoritmo (líneas 9 a 29) es responsable de una sustitución selectiva de valores espectrales inversamente cuantificados no escalados con valores de sustitución de línea espectral y para una modificación selectiva de los factores de escala. Como puede verse a partir del código de pseudo programa, el algoritmo puede ser ejecutado para todos los grupos de ventanas disponibles bucle *for* de las líneas 9 a 29). Además, todas las bandas de factor de escala entre cero y una banda de factor de escala máxima (maxsfb) pueden

ser procesadas a pesar de que el procesamiento puede ser diferente para diferentes bandas de factor de escala (bucle *for* entre las líneas 10 y 28). Un aspecto importante es el hecho de que se supone en general que una banda de factor de escala está cuantificado a cero a menos que se encuentre que la banda de factor de escala no está cuantificado a cero (línea 11). Sin embargo, la verificación de si una banda de factor de escala está cuantificado a cero o no solamente se ejecuta para bandas de factor de escala, una línea de frecuencia de inicio (`swboffset[sfb]`) que está por encima de un índice de coeficiente espectral predeterminado (`noiseFillingStartOffset`). Una rutina condicional entre las líneas 13 y 24 sólo se ejecuta si un índice de los coeficientes espectrales de banda de factor de escala menores `sfb` es mayor que un desplazamiento a inicio de relleno de ruido. En contraste, para cualquier banda de factor de escala para la que un índice del coeficiente espectral más bajo (`swboffset[sfb]`) es menor que o igual a un valor predeterminado (`noiseFillingStartOffset`), se supone que las bandas no están cuantificadas a cero, independiente de los valores de línea espectral reales (ver líneas 24a,24b y 24c).

**[0082]** Si, sin embargo, el índice de los coeficientes espectrales más bajos de una determinada banda de factor de escala es mayor que el valor predeterminado (`noiseFillingStartOffset`), a continuación, la banda de factor de escala determinada se considera como estando cuantificada a cero sólo si todas las líneas espectrales del factor de escala de la banda están cuantificadas a cero (se restablece la etiqueta "bandquantizedtozero" por el bucle *for* de entre las líneas 15 y 22, si un solo contenedor espectral de la banda de factor de escala no se cuantifica a cero.

**[0083]** En consecuencia, un factor de escala de una determinada banda de factor de escala se modifica empleando el desplazamiento de ruido si la etiqueta "bandquantizedtozero", que se establece inicialmente por defecto (línea 11) no se elimina durante la ejecución del código de programa entre las líneas 12 y 24. Como se mencionó anteriormente, un restablecimiento de la etiqueta sólo puede ocurrir para bandas de factor de escala para las que un índice de coeficiente espectral más bajo es superior al valor predeterminado (`noiseFillingStartOffset`). Además, el algoritmo de la figura 10A comprende una sustitución de valores de línea espectral con valores de sustitución de línea espectral si la línea espectral está cuantificada a cero (condición de la línea 16 y operación de sustitución de la línea 17). Sin embargo, dicha sustitución se realiza sólo para las bandas de factor de escala para las que un índice de coeficiente espectral más bajo es superior al valor predeterminado (`noiseFillingStartOffset`). Para las bandas de frecuencias espectrales más bajas, se omite la sustitución de los valores espectrales cuantificados a cero por valores espectrales de sustitución.

**[0084]** Además, debe tenerse en cuenta que los valores de sustitución podrían calcularse de una manera simple añadiendo una señal aleatoria o pseudo-aleatoria al valor de ruido (`noiseVal`) calculado en la primera parte del algoritmo (línea 17).

**[0085]** Hay que destacar que la figura 10B muestra una leyenda de los símbolos relevantes utilizados en el código de pseudo programa de la figura 10A para facilitar una mejor comprensión del código de pseudo programa.

**[0086]** Unos aspectos importantes de la funcionalidad del relleno de ruido se ilustran en la figura 11. Como puede verse, la funcionalidad del relleno de ruido de manera opcional comprende calcular 1110 un valor de ruido a partir del nivel de ruido. La funcionalidad del relleno de ruido también comprende sustituir 1120 valores de línea espectral de líneas espectrales cuantificadas a cero con valores de sustitución de línea espectral en función del valor de ruido para obtener valores de línea espectral. Sin embargo, la sustitución 1120 solamente se realiza para bandas de factor de escala que tienen un coeficiente espectral menor debajo del índice de coeficiente espectral predeterminado.

**[0087]** La funcionalidad del relleno de ruido también comprende modificar 1130 un factor de escala de banda en función del valor de desplazamiento de ruido si, y solamente si, la banda de factor de escala está cuantificada a cero. Sin embargo, la modificación 1130 se ejecuta de esta manera bandas de factor de escala que tienen un coeficiente espectral menor debajo del índice de coeficiente espectral predeterminado.

**[0088]** El relleno de ruido también comprende una funcionalidad 1140 para dejar factores de escala de sin cambiar, independiente de si la banda de factor de escala está cuantificada a cero, para bandas de factor de escala que tienen un coeficiente espectral menor debajo del índice de coeficiente espectral predeterminado.

**[0089]** Además, el re-escalador comprende una funcionalidad 1150 para aplicar (si está disponible) factores de escala de banda no modificados o modificados a valores de línea espectral no sustituidos o sustituidos (si está disponible) para obtener espectros escalados e inversamente cuantificados.

**[0090]** La figura 12 muestra una representación esquemática del concepto descrito con referencia a las figuras 10A, 10B y 11. En particular, las diferentes funcionalidades están representados en función de una banda de contenedor de inicio de factor de escala.

#### 2.2.4.2 Rellenador de ruido según las figuras 13A y 13B

**[0091]** Las figuras 13A y 13B muestran listados de programa de pseudo código de algoritmos, que pueden realizarse en una alternativa de implementación del relleno de ruido 770. La figura 13A describe un algoritmo para derivar

un valor de ruido (para utilizar dentro de el rellenedor de ruido) a partir de una información de nivel de ruido, que puede ser representada por la información de parámetros de relleno de ruido 630ac.

5 **[0092]** Como el error medio de cuantificación es de aproximadamente 0,25 la mayor parte del tiempo, la gama noiseVal [0, 0,5] es bastante grande y se puede optimizar.

10 **[0093]** La figura 13B representa un algoritmo, que puede ser formado por el rellenedor de ruido 770. El algoritmo de la figura 13B comprende una primera parte de determinación del valor de ruido (designada con "noiseValue" o "noiseVal" - líneas 1 a 4). Una segunda parte del algoritmo comprende una modificación selectiva de un factor de escala (líneas 7 a 9) y una sustitución selectiva de valores de línea espectral con valores de sustitución de línea espectral (líneas 10 a 14).

15 **[0094]** Sin embargo, según o el algoritmo de la figura 13B, el factor de escala (scf) se modifica empleando el desplazamiento de ruido (noiseoffset) siempre que una banda esté cuantificado a cero (ver al línea 7). No se hace diferencia entre las bandas de frecuencia inferiores y las bandas de frecuencia más altas en esta realización.

**[0095]** Además, el ruido se introduce en líneas espectrales cuantificadas a cero sólo para las bandas de frecuencia más alta (si la línea está por encima de un cierto umbral predeterminado "noiseFillingStartOffset").

### 20 2.2.5. Conclusión Descodificador

**[0096]** Para resumir, los ejemplos de descodificador de acuerdo pueden comprender una o más de las siguientes características:

- 25 • Empezar de una "línea de comienzo de relleno de ruido " (que puede ser un desplazamiento fijado una línea que representa una sustitución de frecuencia de inicio cada 0 con un valor de sustitución
- el valor de sustitución es el valor de ruido indicado (con un signo aleatorio) en el dominio cuantificado y entonces se escala este " valor de sustitución " con el factor de escala "scf") transmitido para la banda de factor de escala real; y
- 30 • los valores de sustitución "aleatorios" también se pueden derivar de por ejemplo una distribución de ruido o un conjunto de valores alternados ponderados con el nivel de ruido indicado.

## 3. Corriente de audio

### 35 3.1. Corriente de audio según las figuras 14A y 14B

40 **[0097]** A continuación, se describirá un ejemplo de corriente de audio. A continuación, se describirá una así llamada "usac bitstream payload". La "usac bitstream payload" lleva información de carga útil para representar uno o más canales individuales (carga útil "singlechannelement ()") y/o uno o más pares de canales (channelpairelement ()), tal como puede verse a partir de la figura 14A. Una información de canal único (singlechannelement ()) comprende, entre otra información opcional, una corriente de canal en el dominio de la frecuencia (fdchannelstream), tal como puede verse a partir de la figura 14B.

45 **[0098]** Una información de par de canal (channelpairelement) comprende, además de elementos adicionales, una pluralidad de, por ejemplo, dos corrientes de canal de dominio de la frecuencia (fdchannelstream), como puede verse en la figura 14C.

50 **[0099]** El contenido de los datos de una corriente de canal en el dominio de la frecuencia puede, por ejemplo, depender de si se utiliza o no un relleno de ruido (que se puede señalar en una porción de datos de señalización que no se muestra aquí). En lo que sigue, se supondrá que se utiliza un relleno de ruido. En este caso, la corriente de canal en el dominio de la frecuencia comprende, por ejemplo, los elementos de datos que se muestran en la figura 14D. Por ejemplo, puede haber una información de ganancia global (global gain), tal como se define en la norma ISO/IEC 14496-3: 2005. Además, la corriente de canal en el dominio de la frecuencia puede comprender una información de desplazamiento de ruido (noiseoffset) y una información de nivel de ruido (noiselevel), tal como se describe aquí. La información de desplazamiento de ruido puede, por ejemplo, codificarse empleando 3 bits y la información de nivel de ruido pueden, por ejemplo, codificarse empleando 5 bits.

60 **[0100]** Además, la corriente de canal de dominio de la frecuencia puede comprender información codificada de factor de escala (scalefactordata ()) y los datos espectrales codificados aritméticamente (ACspectraldata ()) tal como se describe en este documento y como también se define en la norma ISO / IEC 14496-3.

**[0101]** Opcionalmente, la corriente de canal de dominio de la frecuencia comprende también datos de conformación de ruido temporal (tnsdata ()) , tal como se define en ISO / IEC 14496-3.

65 **[0102]** Naturalmente, la corriente de canal de dominio de la frecuencia puede comprender otra información, si se requiere.

3.2. Corriente de audio según la figura 15

5 **[0103]** La figura 15 muestra una representación esquemática de la sintaxis de una corriente de canal que representa un canal individual (individual\_channel\_stream ()).

10 **[0104]** La corriente de canal individual puede comprender una información de ganancia global (globalgain) codificada utilizando, por ejemplo, 8 bits, información de desplazamiento de ruido (noiseoffset) codificada utilizando, por ejemplo, 5 bits y una información de nivel de ruido (noiselevel) codificada utilizando, por ejemplo, 3 bits.

15 **[0105]** La corriente de canal individual comprende además datos de la sección (sección data ()), datos sobre factores de escala (scalefactordata ()) y datos espectrales (spectraldata ()).

20 **[0106]** Además, la corriente de canal individual puede comprender información opcional adicional, como puede verse en la figura 15.

3.3. Conclusión Corriente de audio

25 **[0107]** Para resumir lo anterior, en algunas formas de realización de acuerdo con la invención, se utilizan los siguientes elementos de sintaxis de flujo de bits:

- Valor que indica un desplazamiento de factor de escala de ruido para optimizar los bits necesarios para transmitir los factores de escala;
- valor que indica el nivel de ruido; y/o
- valor opcional para elegir entre diferentes formas para la sustitución de ruido (ruido uniformemente distribuido en lugar de valores constantes o múltiples niveles discretos en lugar de sólo uno).

4. Conclusión

30 **[0108]** En la codificación de baja velocidad de bits, se puede utilizar ruido de re llenado para dos propósitos:

- La cuantificación aproximada de valores espectrales de codificación de audio de baja tasa de bits podría dar lugar a espectros muy escasos después de la cuantificación inversa, ya que muchas líneas espectrales podrían haber sido cuantificadas a cero. Los espectros escasamente poblados resultarán en una señal decodificada que sonará aguda o inestable (chirridos). Mediante la sustitución de las líneas puestos a cero por los valores "pequeños" en el decodificador, es posible enmascarar o reducir estos artefactos muy obvios sin añadir nuevos artefactos de ruido obvios.
- Si hay partes de señal de similares al ruido en el espectro original, puede reproducirse una representación perceptivamente equivalente de estas partes de señal de ruido en el decodificador a partir solamente de poca información paramétrica, tal como la energía de la parte de señal de ruido. La información paramétrica se puede transmitir con menos bits en comparación con el número de bits necesarios para transmitir la forma de onda codificada.

35 **[0109]** La nueva propuesta de esquema de codificación de relleno de ruido descrito en el presente documento combina de manera eficiente los fines antes señalados, en una sola aplicación.

40 **[0110]** Como comparación, en MPEG-4 de audio, la sustitución por ruido perceptivo (PNS) se utiliza sólo para transmitir una información parametrizada de partes de señal similares a ruido y para reproducir estas partes de señal de manera equivalente perceptualmente en el descodificador.

45 **[0111]** Como comparación adicional, en AMR-WB +, los vectores de cuantificación vectorial (VQ-vectores) cuantificados a cero se sustituyen con un vector de ruido aleatorio en el que cada valor espectral complejo tiene una amplitud constante, pero de fase aleatoria. La amplitud está controlada por un valor de ruido transmitido con el flujo de bits.

50 **[0112]** Sin embargo, los conceptos de comparación proporcionan desventajas significativas. PNS sólo se puede utilizar para llenar bandas completas de factor de escala con ruido, mientras que AMR-WB + sólo intenta enmascarar artefactos en la señal decodificada resultante de grandes partes de la señal que se cuantifica a cero. En contraste, el esquema de codificación de relleno ruido propuesto combina eficientemente ambos aspectos de ruido de relleno en una sola aplicación.

55 **[0113]** Según un aspecto, la presente invención comprende una nueva forma de cálculo del nivel de ruido. El nivel de ruido se calcula en el dominio cuantificado basado en el error de cuantificación medio.

60 **[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 [-0.5; 0.5] (1 nivel de cuantificación) con un



error absoluto promedio de 0.25 (para valores de entrada con distribución normal que son generalmente más grandes que 1).

5 **[0115]** A continuación, se resumen algunas de las ventajas del rellenado de ruido en el dominio cuantificado. La ventaja de la adición de ruido en el dominio cuantificado es el hecho de que el ruido añadido en el descodificador se escala, no sólo con la energía media en una banda dada, sino también la relevancia psicoacústica de una banda.

10 **[0116]** Por lo general, las bandas perceptualmente más relevantes (tonales) serán las bandas cuantificadas con mayor precisión, es decir, se utilizarán múltiples niveles de cuantificación (valores cuantificados más grandes que 1) en estas bandas. Ahora la adición de ruido con un nivel del error medio de cuantificación en estas bandas tendrá una influencia muy limitada en la percepción de esta banda.

15 **[0117]** Las bandas que no son tan perceptivamente relevantes o que son más similares al ruido, pueden ser cuantificadas con un menor número de niveles de cuantificación. Aunque muchas más líneas espectrales en la banda se cuantificarán a cero, el error de cuantificación promedio resultante será el mismo que para las bandas cuantificadas finas (asumiendo un error de cuantificación distribuido de forma normal en ambos grupos), mientras que el error relativo en la banda puede ser mucho más alto.

20 **[0118]** En estas bandas cuantificadas aproximadas, el rellenado de ruido de llenado ayudará a enmascarar perceptualmente los artefactos resultantes de los agujeros espectrales debidos a la cuantificación aproximada.

**[0119]** Una consideración del rellenado de ruido en el dominio cuantificado se puede lograr por el codificador descrito anteriormente, y también por el decodificador anteriormente descrito.

25 5. Alternativas de implementación

30 **[0120]** Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden ser implementadas en hardware o en software. La implementación puede realizarse utilizando un medio de almacenamiento digital, por ejemplo un disquete, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene señales de control legibles electrónicamente almacenadas en el mismo, que cooperan (o son capaces de cooperar) con un sistema informático programable de tal manera que se lleva a cabo el procedimiento respectivo.

35 **[0121]** Algunas formas de realización de acuerdo con la invención comprenden un portador de datos que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema informático programable, de modo que se lleva a cabo uno de los procedimientos descritos en este documento.

40 **[0122]** Generalmente, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa, siendo el código de programa operativo para llevar a cabo uno de los procedimientos, cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede, por ejemplo ser almacenado en un soporte legible por máquina.

45 **[0123]** Otras realizaciones comprenden el programa de ordenador para realizar uno de los procedimientos descritos en este documento, almacenado en un soporte legible por máquina.

**[0124]** En otras palabras, una realización del procedimiento de la invención es, por lo tanto, un programa informático que tiene un código de programa para realizar uno de los procedimientos descritos en este documento, cuando el programa informático se ejecuta en un ordenador.

50 **[0125]** Una realización adicional es, por lo tanto, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, registrado en este, el programa de ordenador para realizar uno de los procedimientos descritos en este documento.

55 **[0126]** Una realización adicional del procedimiento de la invención es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa de ordenador para realizar uno de los procedimientos descritos en este documento. El flujo de datos o la secuencia de señales pueden, por ejemplo estar configurados para ser transferidos a través de una conexión de comunicación de datos, por ejemplo a través de Internet.

60 **[0127]** Una forma de realización comprende además un medio de procesamiento, por ejemplo un ordenador, o un dispositivo lógico programable, configurado o adaptado para realizar uno de los procedimientos descritos en este documento.

**[0128]** Una forma de realización comprende además un ordenador que tiene instalado el programa informático para realizar uno de los procedimientos descritos en este documento.

## REIVINDICACIONES

- 5 1. Un codificador (100; 228) para proporcionar una corriente de audio (126; 212) a partir de una representación en el dominio de la transformada (112; 114; 228a) de una señal de audio de entrada, comprendiendo el codificador: un  
 10 calculador de error de cuantificación (110; 330) configurado para determinar un error de cuantificación multi-banda (116; 332) sobre una pluralidad de bandas de frecuencia de la señal de audio de entrada, para las que hay disponible una información de ganancia de banda separada (228a); y un proveedor de corriente de audio (120; 230) configurado para proporcionar la corriente de audio (126; 212) de modo que la corriente de audio comprende una información espectral que describe un contenido de audio de las bandas de frecuencia y una información que describe el error de cuantificación multi-banda; en el que el calculador de error de cuantificación (330) está configurado para determinar el error de cuantificación multi-banda (332) sobre una pluralidad de bandas de frecuencia que comprenden cada una al menos una componente espectral cuantificada a un valor diferente de cero evitando a la vez bandas de frecuencia, cuyas componentes espectrales están totalmente cuantificadas a cero.
- 15 2. El codificador (100; 228) según la reivindicación 1, en el que el codificador comprende un cuantificador (310) configurado para cuantificar componentes espectrales de diferentes bandas de frecuencia de la representación en el dominio de la transformada (228a) empleando diferentes precisiones de cuantificación en función de relevancias psicoacústicas (228c) de las diferente bandas de frecuencia, para obtener componentes espectrales cuantificados, en el que las diferentes precisiones de cuantificación están reflejadas por la información de ganancia de banda; y en el que el proveedor de corriente de audio (212) está configurado para proporcionar la corriente de audio de modo que la corriente de audio comprende una información que describe la información de ganancia de banda y de modo que la corriente de audio también comprende la información que describe el error de cuantificación multi-banda.
- 25 3. El codificador (100; 228) según la reivindicación 2, en el que el cuantificador (310) está configurado para realizar un escalado de la componente espectral en función de la información de ganancia de banda y para realizar una cuantificación de valor entero de las componentes espectrales escaladas; y en el que el calculador de error de cuantificación (330) está configurado para determinar el error de cuantificación multi-banda (332) en el dominio cuantificado, de modo que el escalado de las componentes espectrales, que se realiza antes de la cuantificación de valor entero, se tiene en cuenta en el error de cuantificación multi-banda.
- 30 4. El codificador (100; 228) según cualquiera de las reivindicaciones 1 a 3, en el que el codificador está configurado para ajustar una información de ganancia de banda de una banda de frecuencia, que está completamente cuantificada a cero, a un valor que representa una relación entre una energía de la banda de frecuencia completamente cuantificada a cero y una energía del error de cuantificación multi-banda.
- 35 5. Un procedimiento para proporcionar una corriente de audio (126; 212) a partir de una representación en el dominio de la transformada (112; 114; 228a) de una señal de audio de entrada, comprendiendo el procedimiento:  
 40 determinar un error de cuantificación multi-banda sobre una pluralidad de bandas de frecuencia de la señal de audio de entrada, para las que hay disponible una información de ganancia de banda separada; y proporcionar la corriente de audio de modo que la corriente de audio comprenda una información espectral que describe un contenido de audio de las bandas de frecuencia y una información que describe el error de cuantificación multi-banda;  
 45 en el que el error de cuantificación multi-banda (332) se determina sobre una pluralidad de bandas de frecuencia que comprenden cada una al menos una componente espectral cuantificada a un valor diferente de cero mientras se evitan bandas de frecuencia, cuyas componentes espectrales están totalmente cuantificadas a cero.
6. Un programa de ordenador para realizar el procedimiento según la reivindicación 5 cuando el programa de ordenador se ejecuta en un ordenador.

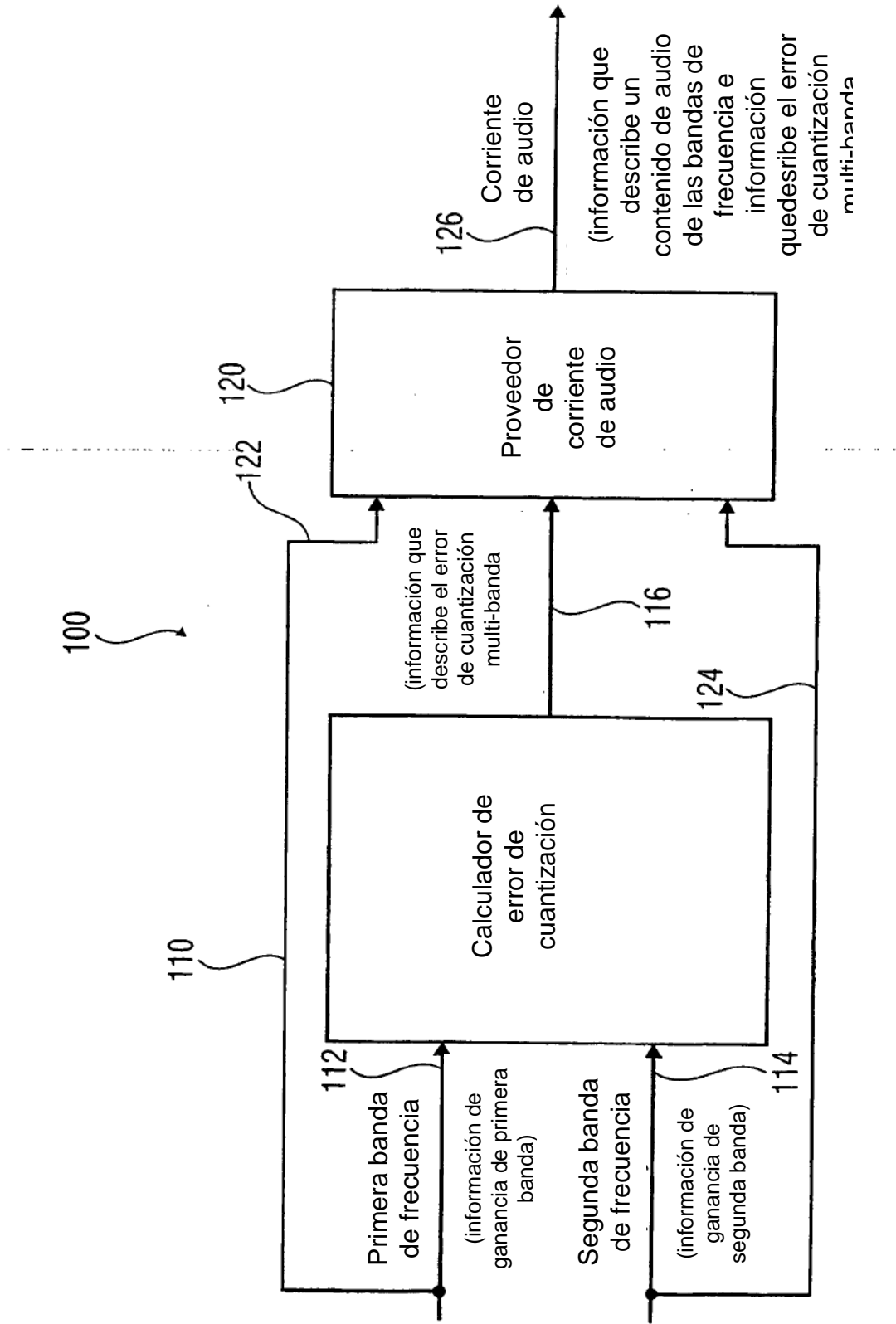
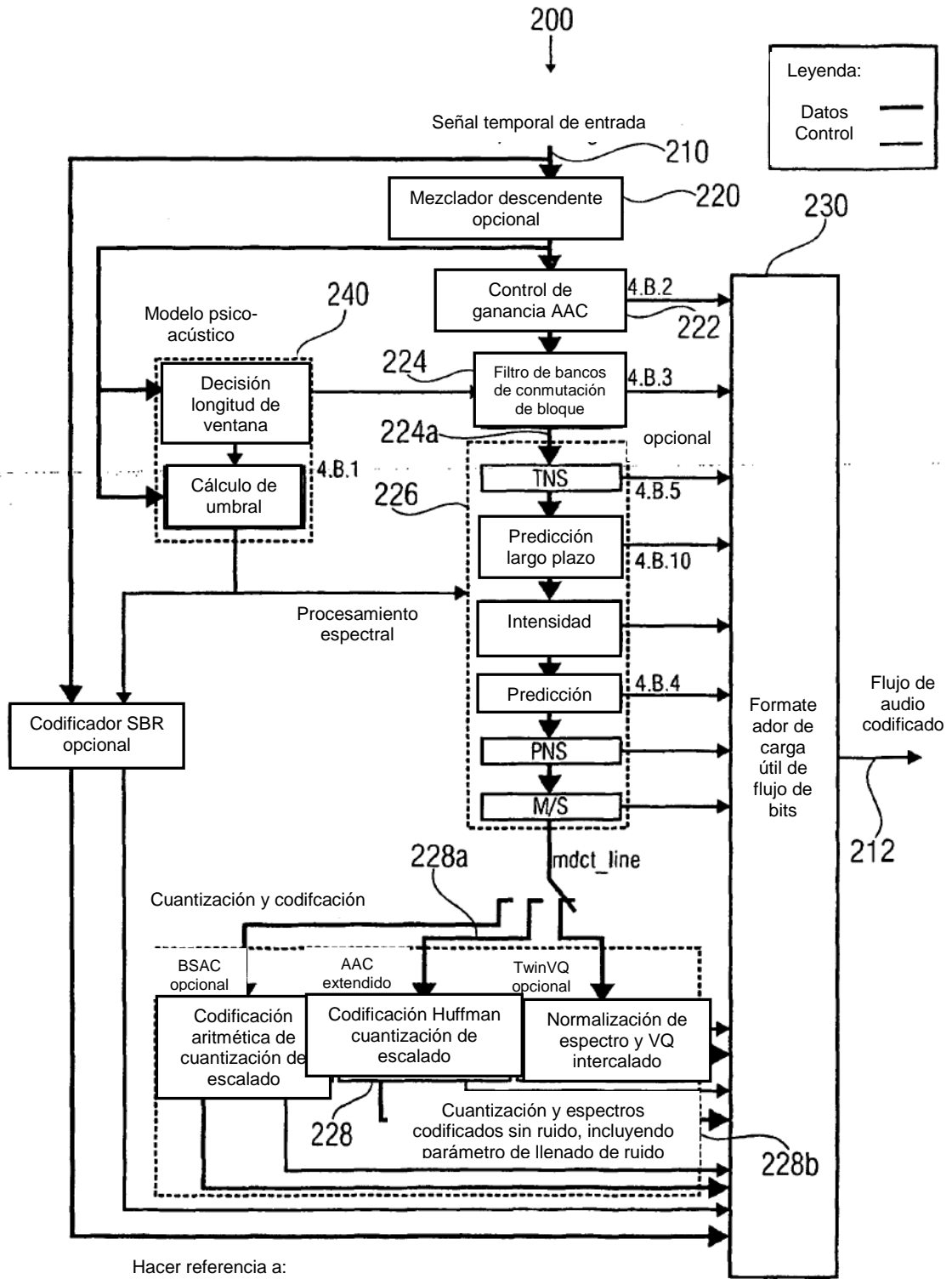
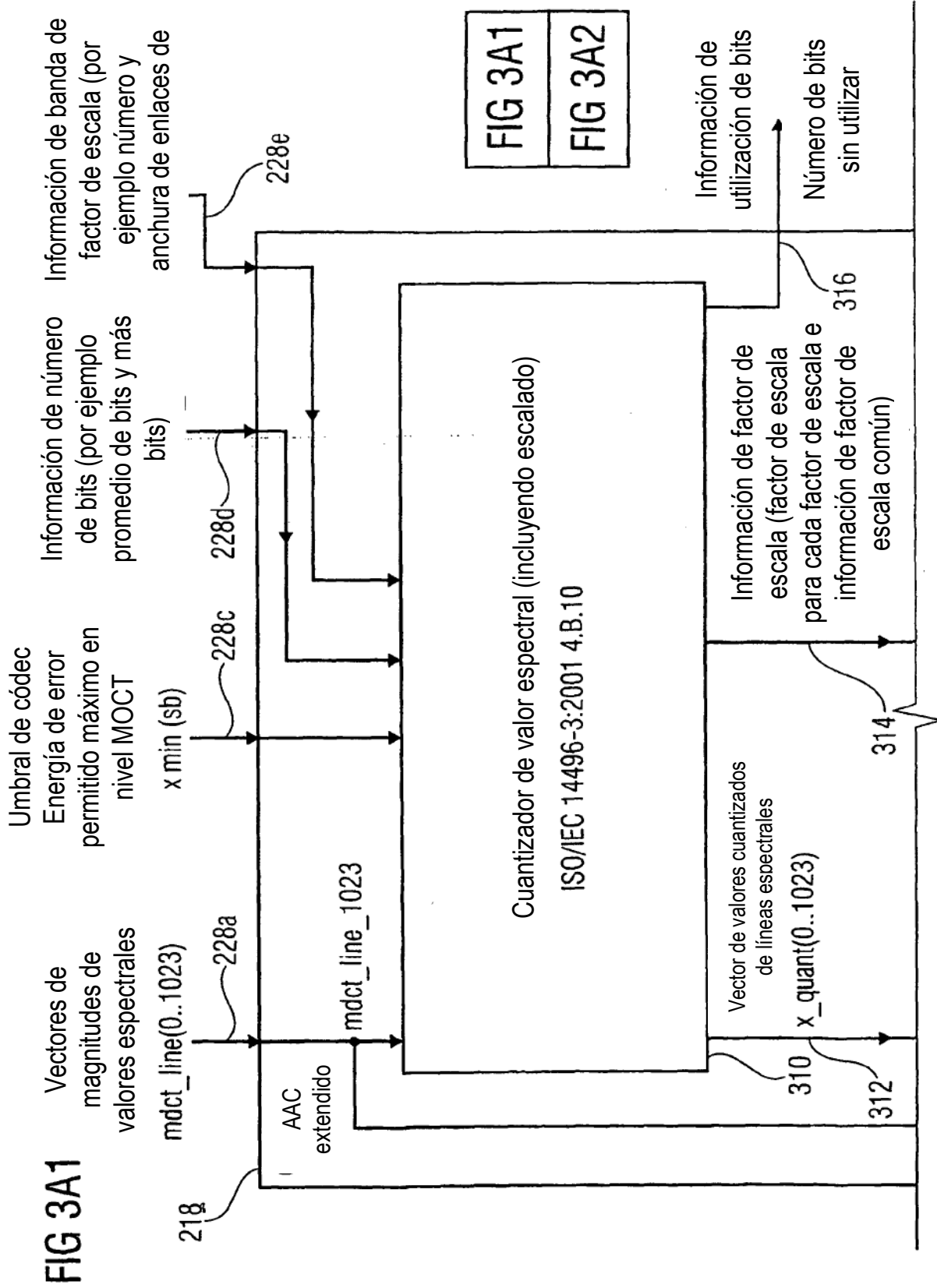


FIG 1



Hacer referencia a:  
ISO/IEC 14496-3: 2005 Subpart 4, Fig. 4.1

FIG 2



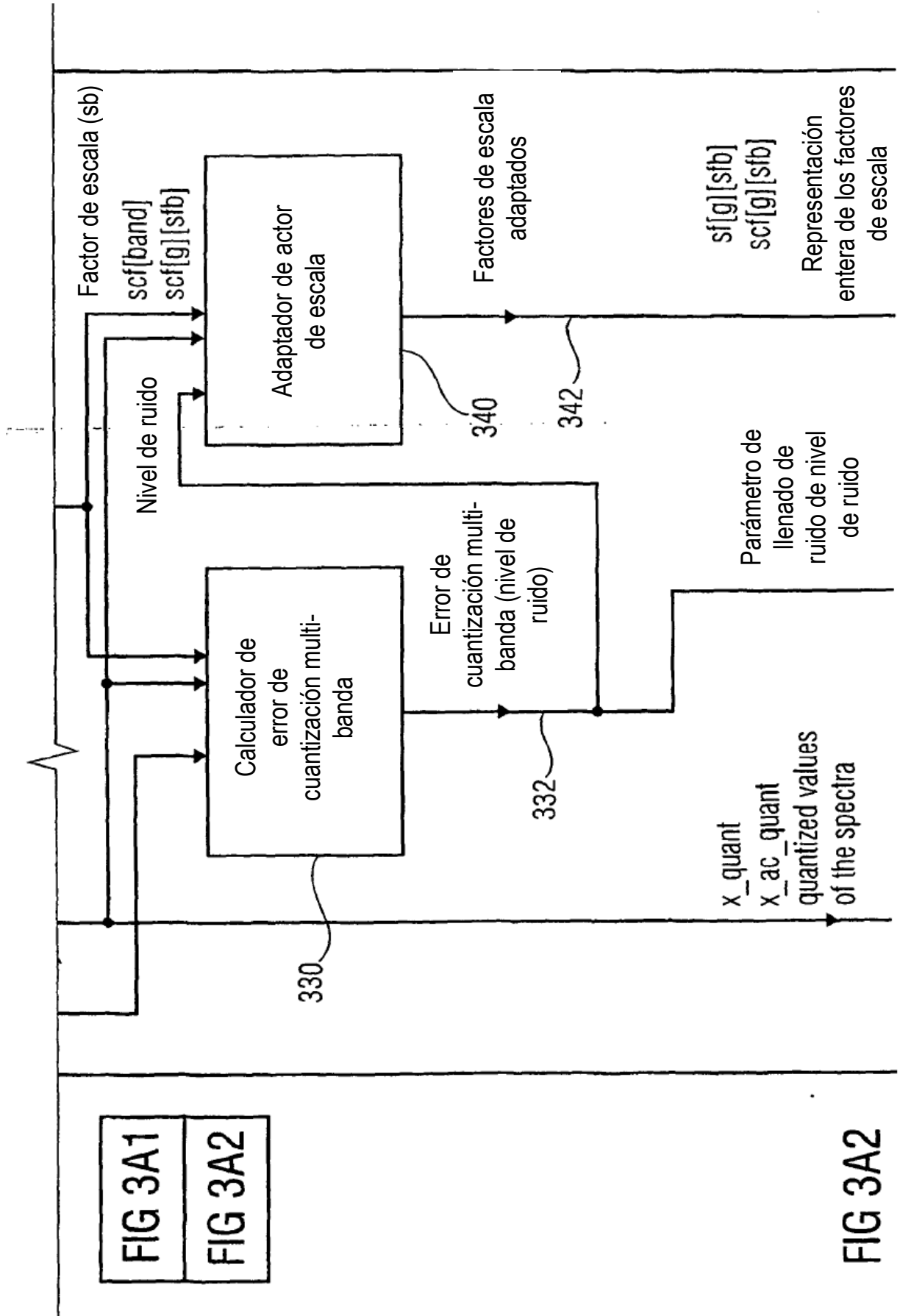
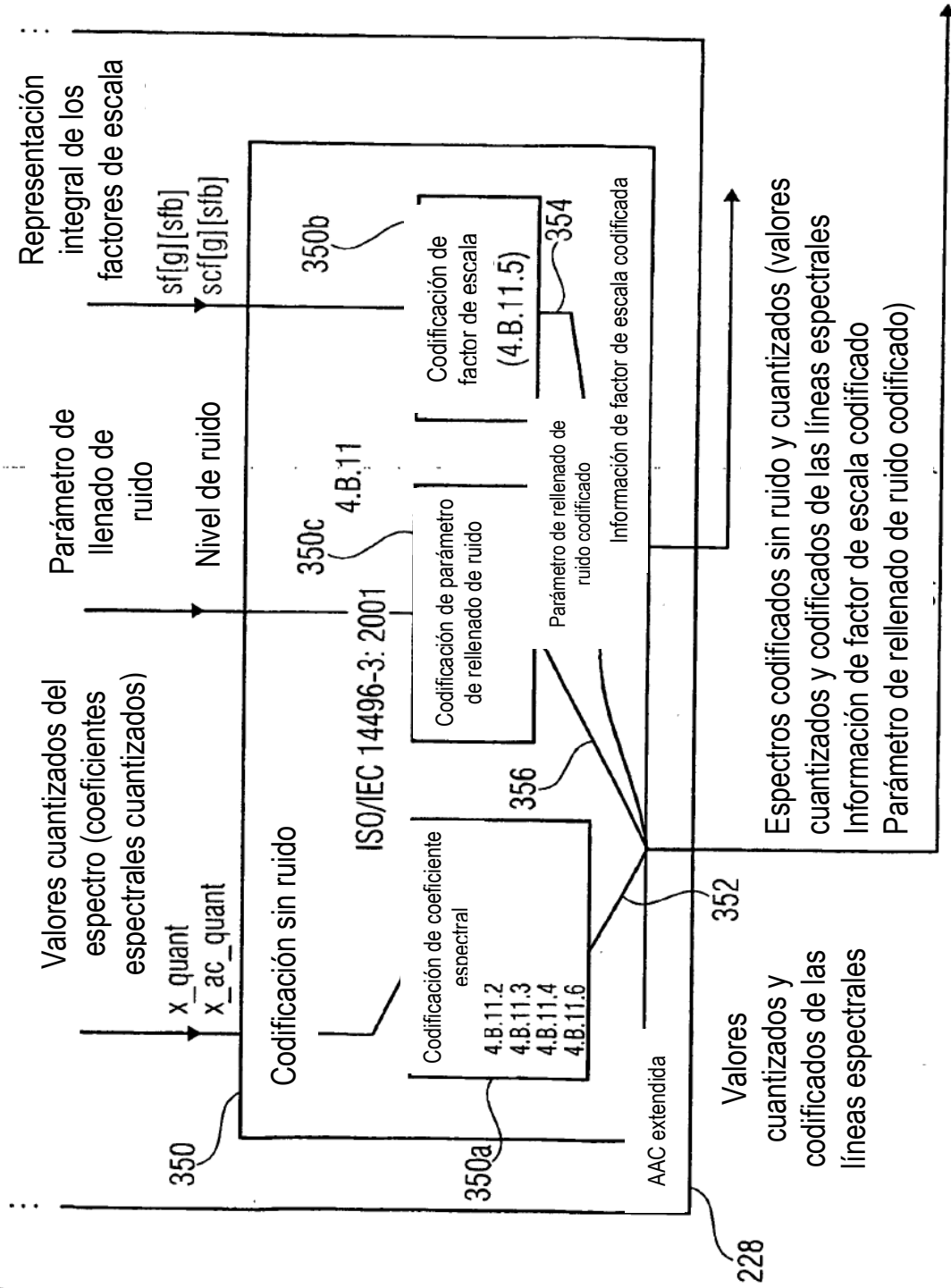


FIG 3B



**FIG 4A** Número de línea

Codificador:

```

1  Calcular error de cuantización promedio:
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 Calcular factor de escala cero:
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( ((float) sbWidth * noiseVal * noiseVal) / log(2.f) ));
21 }
22 }
23 else {
24 scf = don't care
25 }

```

Valor de magnitud de línea espectral escalada no cuantizado

Valor de magnitud de línea espectral escalada cuantizado

} Opcional: cuantización de nivel de ruido

} Opcional: cuantización inversa de nivel de ruido



Cuantización de nivel de ruido

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

**FIG 4B**

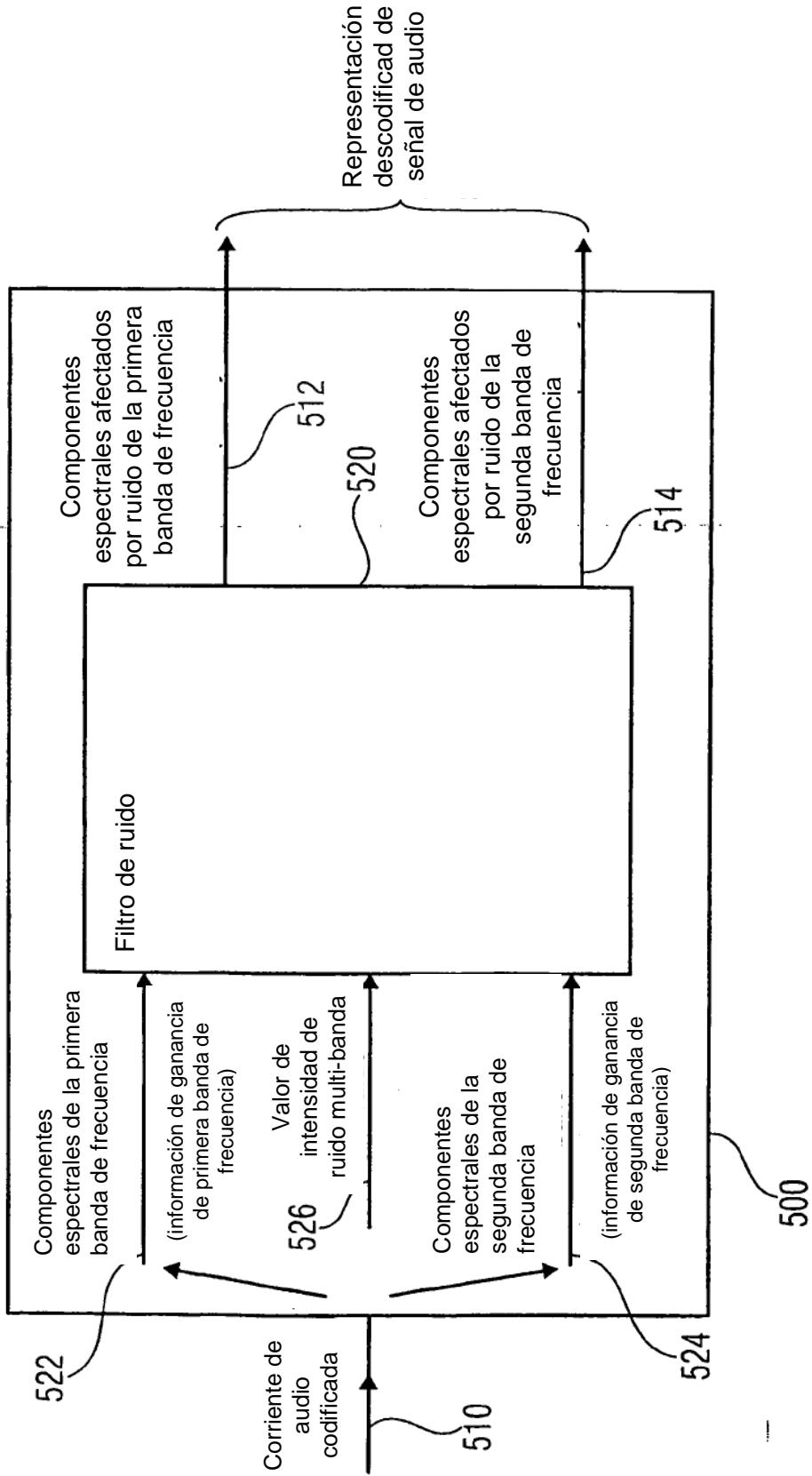


FIG 5

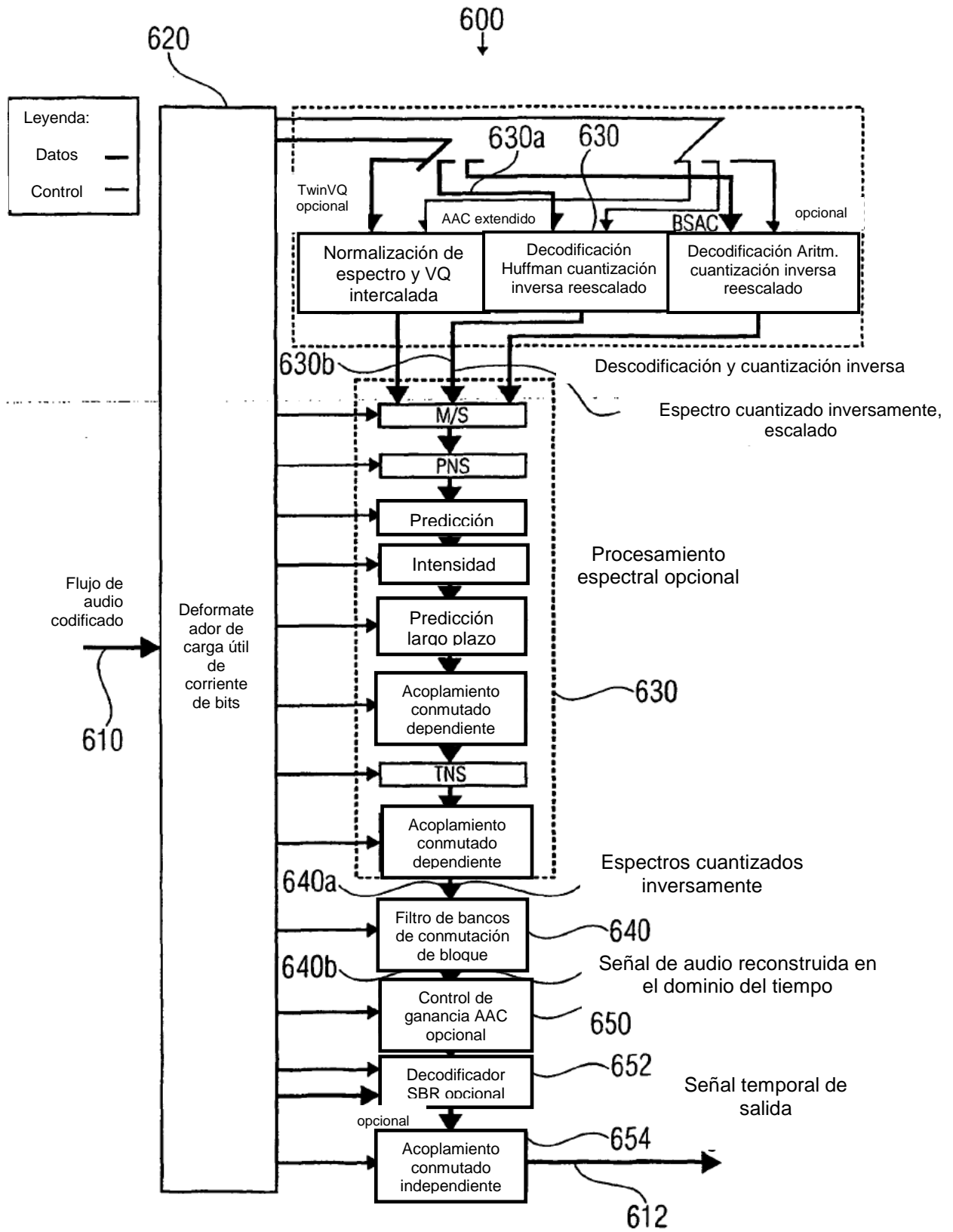


FIG 6

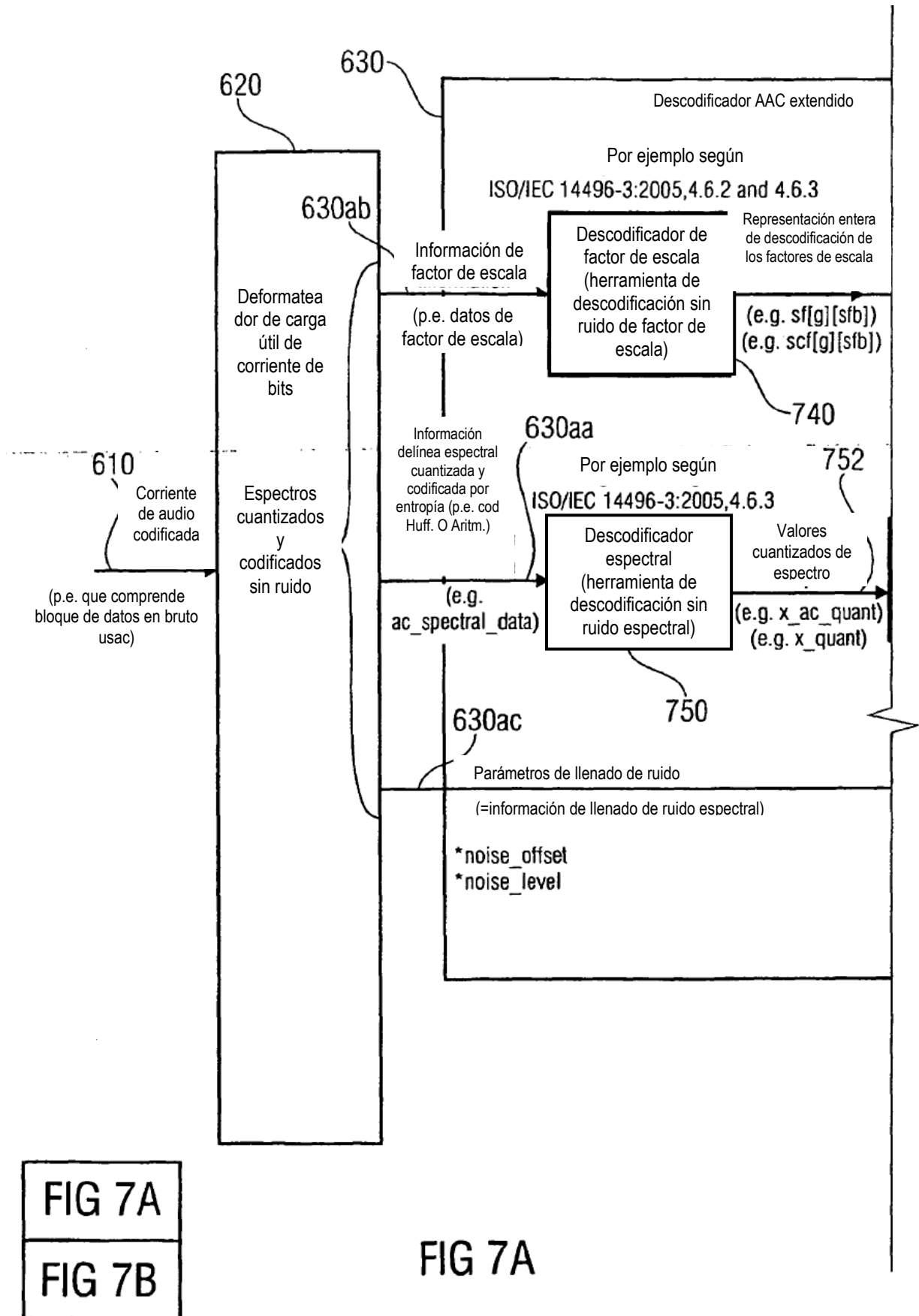
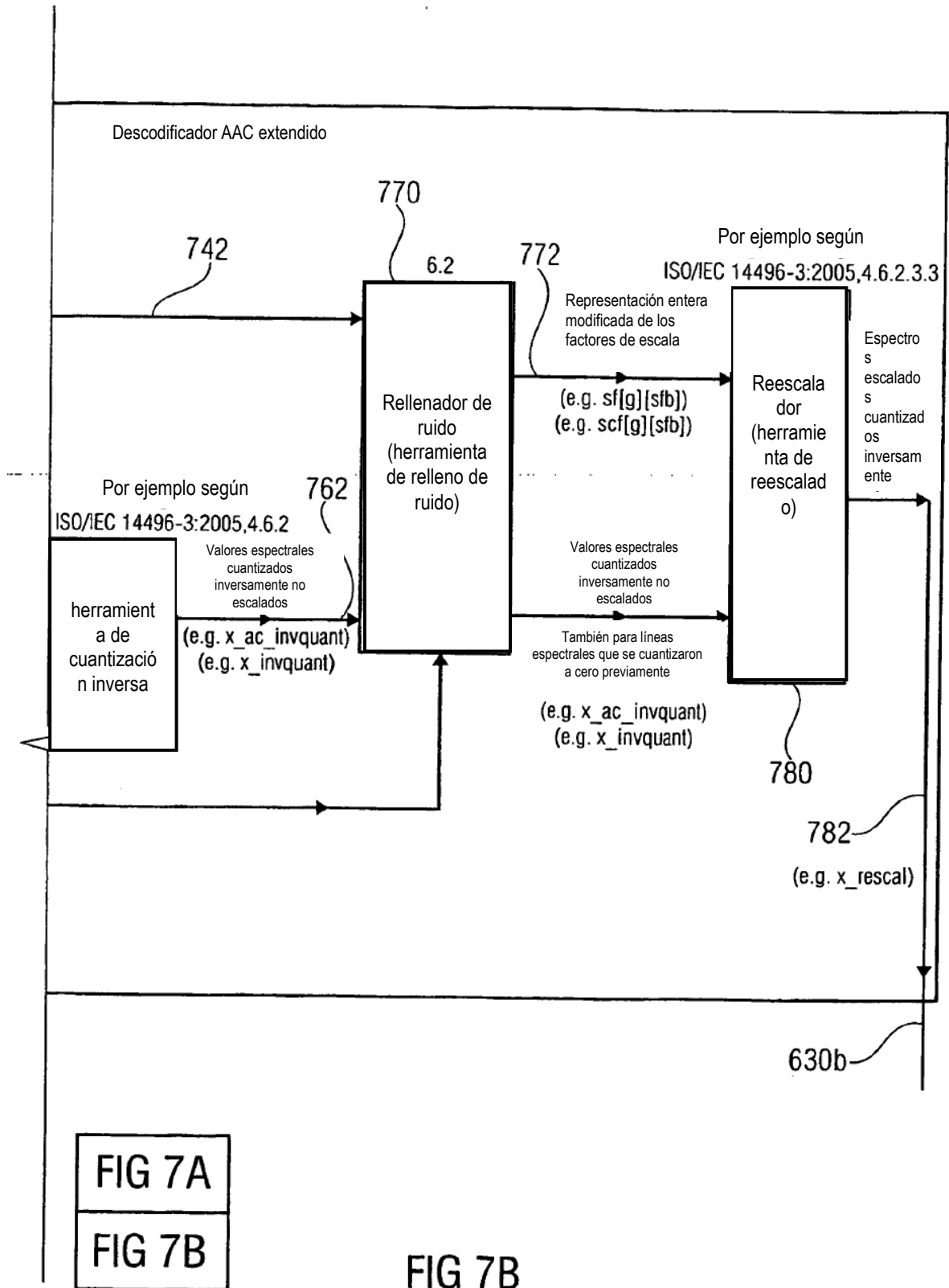


FIG 7A  
FIG 7B

FIG 7A



$$x\_invquant = \text{Sign}(x\_quant) \cdot |x\_quant|^{\frac{4}{3}}$$

o:

$$x\_ac\_invquant = \text{Sign}(x\_ac\_quant) \cdot |x\_ac\_quant|^{\frac{4}{3}}$$

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

↓ x\_ac\_quant

Para bandas de factor de escala por debajo de una banda de factor de escala máxima predeterminada, obtener valores espectrales cuantizados inversamente no escalados como una función de valores espectrales cuantizados no escalados

↓ x\_ac\_invquant

FIG 8C

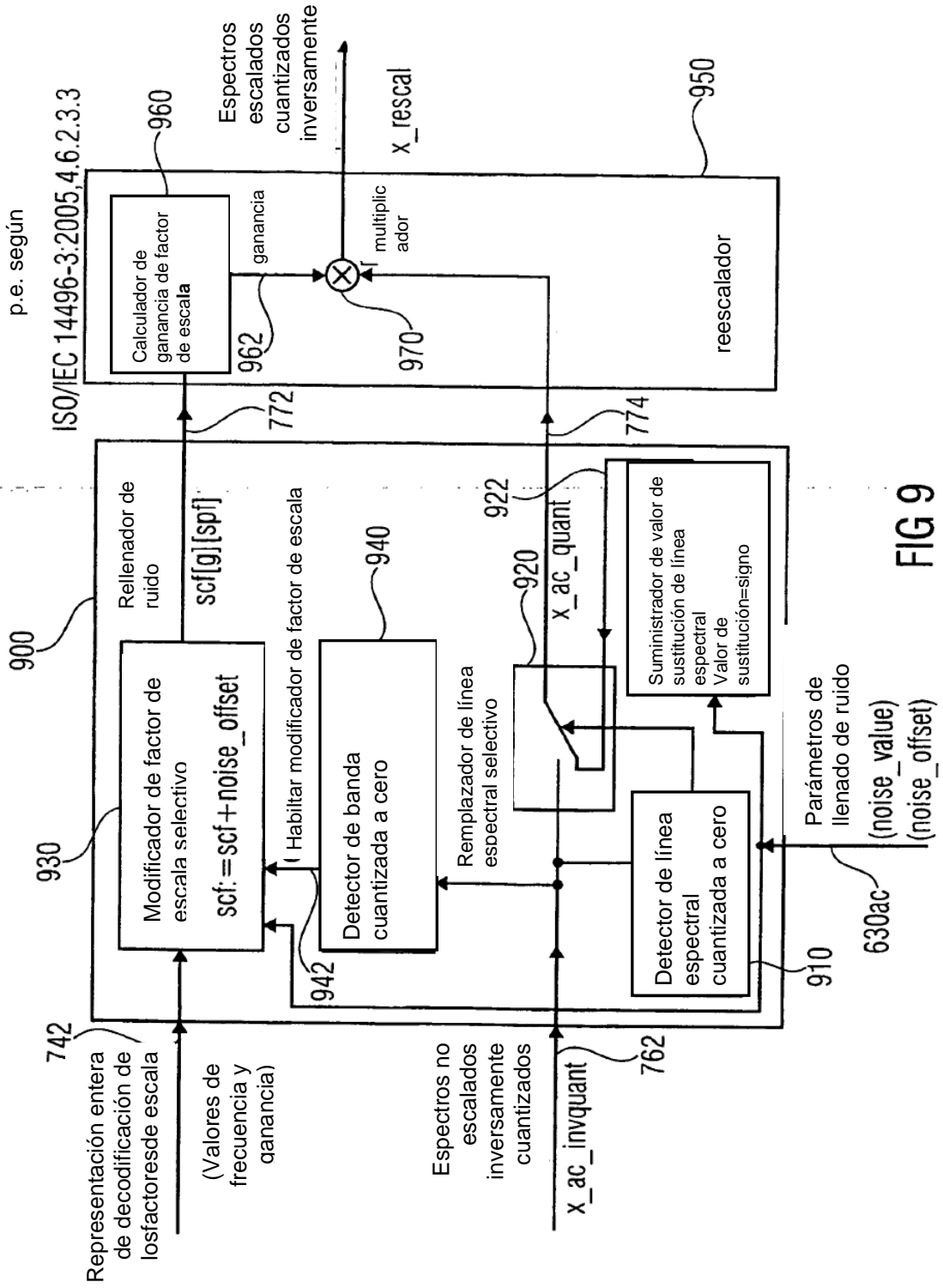


FIG 9

## Proceso de descodificación: Proceso de rellenado 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_lo_zero=1;
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_inquant[g][win][sfb][bin] == 0) {
17            x_ac_inquant[g][win][sfb][bin] = randomSign() * noiseVal;
18          }
19        }
20      }
21    }
22  }
23 }
24 }
24a else {
24b   band_quantized_lo_zero=0;
24c }
25 if(band_quantized_lo_zero) {
26   scf[g][sfb] = scf[g][sfb] + noise_offset;
27 }
28 }
29 }

```

--Suponer que la banda está cuantizada a cero

--para bandas de factor de escala, empezar por encima de noiseFillingStartOffset  
 --añadir ruido de amplitud noiseVal a las líneas espectrales cuantizadas a cero;

--para bandas de factor de escala, empezar por encima de oiseFillingStartOffset, si un único  
 --contenedor de factor de escala es diferente 0, entonces la banda no se cuantiza a cero;

--para bandas de factor de escala, que empiezan por debajo de noiseFillingStartOffset, si  
 --siempre se asume que la banda no está cuantizada a cero;

--para bandas de factor de escala cuantizadas a cero, modificar el factor de escala de banda  
 --en función del valor de desplazamiento de ruido



**Elementos de datos**

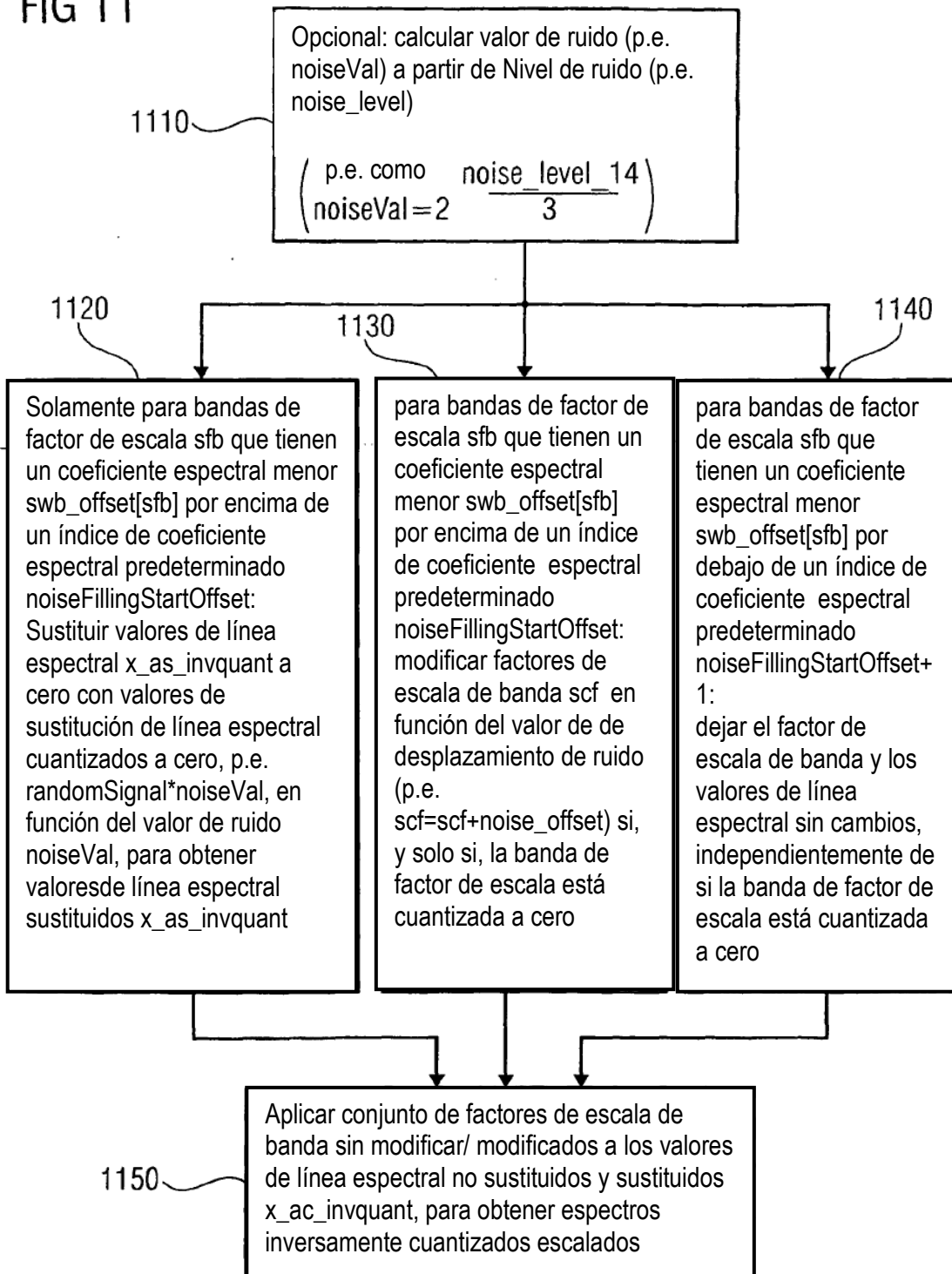
<b>noise_offset</b>	Desplazamiento adicional para modificar los factores de escala de bandas cuantizadas a cero
<b>noise_level</b>	Entero que representa el ruido de cuantización a añadir a cada línea espectral cuantizada a a cero

**Elementos de ayuda**

<b>x_ac_invquant[g][win][sfb][bin]</b>	Coefficiente espectral AAC para grupo g, ventana win, banda de factor de escala sfb, contenedor de coeficientes tras cuantización inversa
<b>noiseFillingStartOffset[win]</b>	Un desplazamiento general o frecuencia de inicio de llenado de ruido. El desplazamiento se define para secuencias cortas (window_sequence==EIGHT_SHORT_SEQUENCE) y 160 en otros casos
<b>noiseVal</b>	El valor de ruido absoluto que sustituye cada contenedor cuantizado a cero.
<b>randomSign</b>	Signo aleatorio (-1,1) multiplicado por noiseVal
<b>band_quantized_to_zero</b>	Etiqueta para indicar si un sfb está completamente cuantizado a cero
<b>swb_offset[sfb]</b>	Índice del coeficiente espectral menor de banda de factor de escala sfb
<b>num_window_groups</b>	Número de grupos de ventanas que comparten un conjunto de factores de escala
<b>mux_sfb</b>	Número de Bandas de factor de escala por grupo
<b>window_group_len</b>	Número de ventanas en cada grupo
<b>g</b>	Índice de grupo
<b>win</b>	Índice de ventana dentro del grupo
<b>sfb</b>	Índice de banda de factor de escala dentro del grupo
<b>swb</b>	Índice de banda de ventana de factor de escala dentro de la ventana
<b>bin</b>	Índice de coeficiente
<b>num_windows</b>	Número de ventanas de la secuencia de ventanas actual

**FIG 10B**

FIG 11



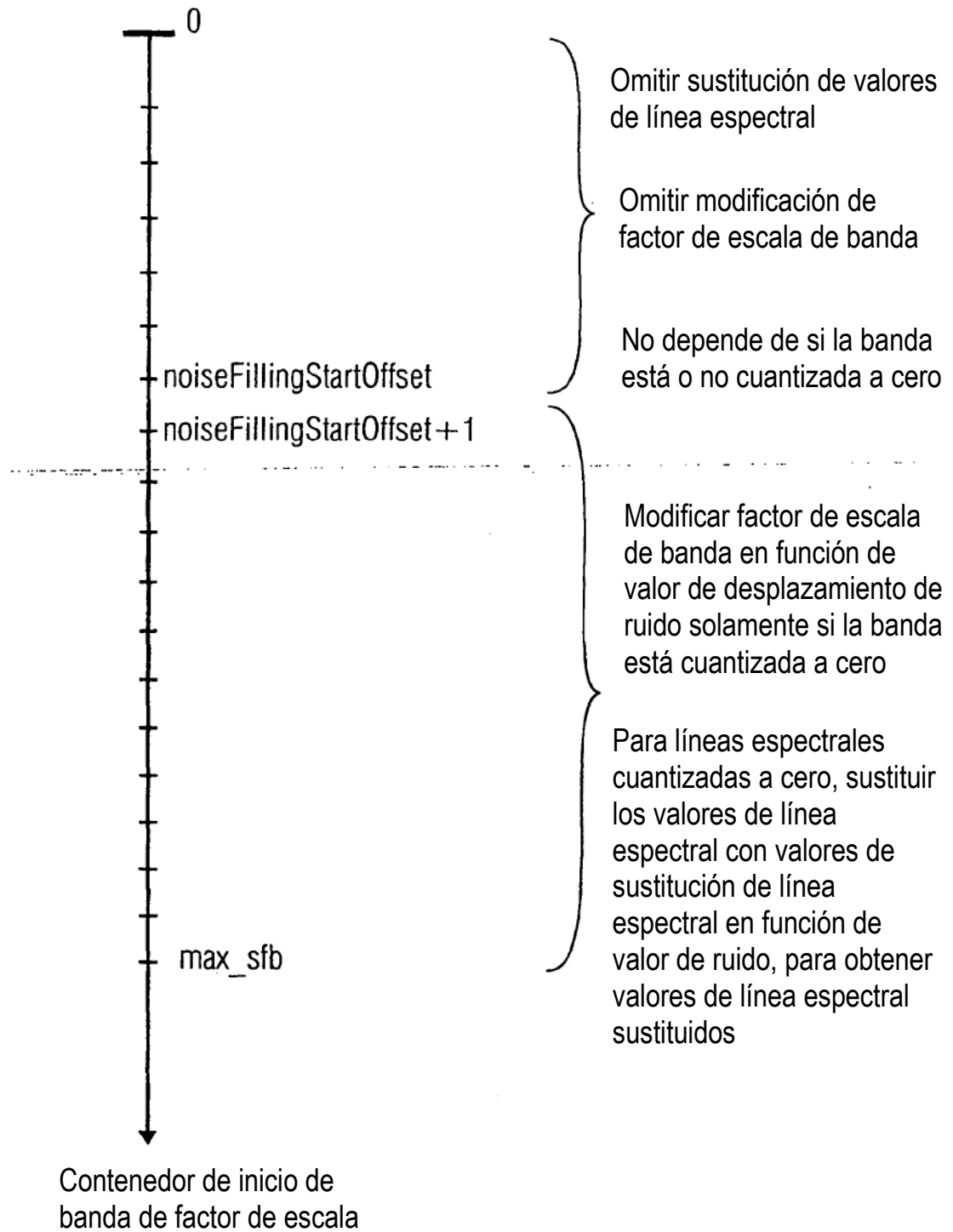


FIG 12

Número de línea	Valor de ruido en descodificador
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	descodificador: Cálculo del "valor de sustitución 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

Carga útil de corriente de bits USAC

(Tab. 4.3)

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

### FIG 14A

single\_channel\_element ()

(Tab. 4.4)

```
{
  fd_channel_stream (*, *, noise Filling)
}
```

### FIG 14B

channel\_pair\_element

(Tab. 4.5)

```
{
  fd_channel_stream (*, *, noise Filling)(1st channel); and/or
  fd_channel_stream (*, *, noise Filling)(2nd channel)
}
```

### FIG 14C

fd\_channel\_stream ()

(Tab. 4.8)

```
{
  global_gain;           e.g. 8 bit
  noise_offset;         e.g. 3 bit
  noise_level;          e.g. 5 bit
```

scale\_factor\_data ();

tus\_data (); Opcional

```
ac_spectral_data ()
}
```

### FIG 14D

Sintaxis de individual\_channel\_stream()

Sintaxis	núm. de bits	código
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