

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 946 760**

51 Int. Cl.:

G10L 19/16 (2013.01)

G10L 19/24 (2013.01)

G10L 21/038 (2013.01)

G10L 19/035 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **10.03.2016** **E 21193211 (6)**

97 Fecha y número de publicación de la concesión europea: **26.04.2023** **EP 3985667**

54 Título: **Descodificar corrientes de bits de audio con metadatos de replicación potenciada de banda espectral en al menos un elemento de relleno**

30 Prioridad:

13.03.2015 EP 15159067

16.03.2015 US 201562133800 P

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

25.07.2023

73 Titular/es:

DOLBY INTERNATIONAL AB (100.0%)
77 Sir John Rogerson's Quay Block C Grand
Canal Docklands
Dublin, D02 VK60, IE

72 Inventor/es:

VILLEMoes, LARS;
PURNHAGEN, HEIKO y
EKSTRAND, PER

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 946 760 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

Descodificar corrientes de bits de audio con metadatos de replicación potenciada de banda espectral en al menos un elemento de relleno

5

Referencia cruzada a aplicación relacionada

Esta solicitud es una solicitud divisional europea de la solicitud de patente Euro-PCT EP16765449.0 (referencia: D15012BEP01), presentada el 10 de marzo de 2016.

10

Esta solicitud reivindica la prioridad de la solicitud de patente europea No. 15159067.6 presentada el 13 de marzo de 2015, y la solicitud provisional de los Estados Unidos No. 62/133,800 presentada el 16 de marzo de 2015.

Campo técnico

15

La invención se refiere al procesamiento de señales de audio. Algunas realizaciones se refieren a la descodificación de corrientes de bits de audio (por ejemplo, corrientes de bits que tienen un formato MPEG-4 AAC) que incluyen metadatos para controlar la replicación de banda espectral potenciada (eSBR). Otras realizaciones se refieren a la descodificación de tales corrientes de bits por descodificadores heredados que no están configurados para realizar el procesamiento eSBR y que ignoran tales metadatos, o a la descodificación de una corriente de bits de audio que no incluye tales metadatos, incluso mediante la generación de datos de control de eSBR en respuesta a la corriente de bits.

20

Antecedentes de la invención

25

Una corriente de bits de audio típica incluye tanto datos de audio (por ejemplo, datos de audio codificados) indicativos de uno o más canales de contenido de audio, como metadatos indicativos de al menos una característica de los datos de audio o del contenido de audio. Un formato bien conocido para generar una corriente de bits de audio codificada es el formato de codificación de audio avanzada (AAC) MPEG-4, descrito en el estándar MPEG ISO/IEC 14496-3: 2009. En el estándar MPEG-4, AAC denota "codificación de audio avanzada" y HE-AAC denota "codificación de audio avanzada de alta eficiencia".

30

El estándar MPEG-4 AAC define varios perfiles de audio, que determinan qué objetos y herramientas de codificación están presentes en un codificador o descodificador de reclamación. Tres de estos perfiles de audio son (1) el perfil AAC, (2) el perfil HE-AAC y (3) el perfil HE-AAC v2. El perfil AAC incluye el tipo de objeto AAC de baja complejidad (o "AAC-LC"). El objeto AAC-LC es la contraparte del perfil de baja complejidad MPEG-2 AAC, con algunos ajustes, y no incluye ni el tipo de objeto de replicación de banda espectral ("SBR") ni el tipo de objeto estéreo paramétrico ("PS"). El perfil HE-AAC es un superconjunto del perfil AAC y además incluye el tipo de objeto SBR. El perfil HE-AAC v2 es un superconjunto del perfil HE-AAC y además incluye el tipo de objeto PS.

35

40

El tipo de objeto SBR contiene la herramienta de replicación de banda espectral, que es una importante herramienta de codificación que mejora significativamente la eficiencia de compresión de los códecs de audio de percepción. SBR reconstruye los componentes de alta frecuencia de una señal de audio en el lado del receptor (por ejemplo, en el descodificador). Por lo tanto, el codificador solo necesita codificar y transmitir componentes de baja frecuencia, lo que permite una calidad de audio mucho mayor a velocidades de datos bajas. SBR se basa en la replicación de las secuencias de armónicos, previamente truncadas para reducir la velocidad de datos, a partir de la señal limitada de ancho de banda disponible y los datos de control obtenidos del codificador. La relación entre los componentes tonales y similares al ruido se mantiene mediante el filtrado inverso adaptativo, así como la adición opcional de ruido y sinusoidales. En el estándar MPEG-4 AAC, la herramienta SBR realiza el parcheo espectral, en el que se copian varias subbandas adyacentes del filtro de espejo en cuadratura (QMF) desde una porción de banda baja transmitida de una señal de audio a una porción de banda alta de la señal de audio, que se genera en el descodificador.

45

50

El parcheo espectral puede no ser ideal para ciertos tipos de audio, como contenido musical con frecuencias de cruce relativamente bajas. Por lo tanto, se necesitan técnicas para mejorar la replicación de banda espectral.

55

Breve descripción de las realizaciones de la invención

La presente divulgación proporciona una unidad de procesamiento de audio, un método para descodificar una corriente de bits de audio codificada y un medio o dispositivo de almacenamiento para realizar dicho método, como se reivindica en las reivindicaciones 1, 5 y 6. Las características opcionales se enumeran en las reivindicaciones dependientes.

60

Breve descripción de los dibujos

La figura 1 es un diagrama de bloques de una realización de un sistema que puede configurarse para realizar una realización del método de la invención.

65

La figura 2 es un diagrama de bloques de un codificador que es un ejemplo de unidad de procesamiento de audio que no está de acuerdo con la invención y está presente solo con fines ilustrativos.

5 La figura 3 es un diagrama de bloques de un sistema que incluye un descodificador que es una realización de la unidad de procesamiento de audio de la invención y, opcionalmente, también un posprocesador acoplado al mismo.

La figura 4 es un diagrama de bloques de un descodificador que es una realización de la unidad de procesamiento de audio de la invención.

10 La figura 5 es un diagrama de bloques de un descodificador que es otra realización de la unidad de procesamiento de audio de la invención.

La figura 6 es un diagrama de bloques de otra realización de la unidad de procesamiento de audio de la invención.

15 La figura 7 es un diagrama de un bloque de una corriente de bits MPEG-4 AAC, incluidos los segmentos en los que se divide.

20 **Nomenclatura y notación**

A lo largo de esta divulgación, incluidas las reivindicaciones, la expresión realizar una operación "en" una señal o datos (por ejemplo, filtrar, escalar, transformar o aplicar ganancia a la señal o los datos) se usa en un sentido amplio para denotar la realización de la operación directamente en la señal o los datos, o en una versión procesada de la señal o los datos (por ejemplo, en una versión de la señal que ha sido sometida a un filtrado preliminar o preprocesamiento antes de la realización de la operación en los mismos).

30 A lo largo de esta divulgación, incluidas las reivindicaciones, la expresión "unidad de procesamiento de audio" se usa en un sentido amplio, para denotar un sistema, dispositivo o aparato, configurado para procesar datos de audio. Los ejemplos de unidades de procesamiento de audio incluyen, entre otros, codificadores (por ejemplo, transcodificadores), descodificadores, códecs, sistemas de preprocesamiento, sistemas de posprocesamiento y sistemas de procesamiento de corriente de bits (a veces denominados herramientas de procesamiento de corriente de bits). Prácticamente todos los productos electrónicos de consumo, como teléfonos móviles, televisores, portátiles y tabletas, contienen una unidad de procesamiento de audio.

35 A lo largo de esta divulgación, incluidas las reivindicaciones, el término "se acopla" o "acoplado" se usa en un sentido amplio para indicar una conexión directa o indirecta. Por lo tanto, si un primer dispositivo se acopla a un segundo dispositivo, esa conexión puede ser a través de una conexión directa o mediante una conexión indirecta a través de otros dispositivos y conexiones. Además, los componentes que están integrados en o con otros componentes también están acoplados entre sí.

40 **Descripción detallada de las realizaciones de la invención**

45 El estándar MPEG-4 AAC contempla que una corriente de bits MPEG-4 AAC codificada incluye metadatos indicativos de cada tipo de procesamiento SBR que se aplicará (si se ha de aplicar alguno) por un descodificador para descodificar el contenido de audio de la corriente de bits, y/o cuál controla tal procesamiento SBR, y/o es indicativo de al menos una característica o parámetro de al menos una herramienta SBR a emplear para descodificar el contenido de audio de la corriente de bits. En el presente documento, usamos la expresión "metadatos SBR" para denotar metadatos de este tipo que se describen o mencionan en el estándar MPEG-4 AAC.

50 El nivel superior de una corriente de bits MPEG-4 AAC es una secuencia de bloques de datos (elementos "raw_data_block"), cada uno de los cuales es un segmento de datos (en el presente documento denominado "bloque") que contiene datos de audio (típicamente durante un tiempo período de 1024 o 960 muestras) e información relacionada y/u otros datos. En el presente documento, usamos el término "bloque" para denotar un segmento de una corriente de bits MPEG-4 AAC que comprende datos de audio (y los metadatos correspondientes y, opcionalmente, también otros datos relacionados) que determina o es indicativo de un elemento (pero no más de uno) "raw_data_block".

60 Cada bloque de una corriente de bits MPEG-4 AAC puede incluir varios elementos sintácticos (cada uno de los cuales también se materializa en la corriente de bits como un segmento de datos). En el estándar MPEG-4 AAC se definen siete tipos de tales elementos sintácticos. Cada elemento sintáctico se identifica por un valor diferente del elemento de datos "id_syn_ele". Ejemplos de elementos sintácticos incluyen un "single_channel_element ()", un "channel_pair_element ()" y un "fill_element ()". Un elemento de un solo canal es un contenedor que incluye datos de audio de un solo canal de audio (una señal de audio monofónica). Un elemento de par de canales incluye datos de audio de dos canales de audio (es decir, una señal de audio estéreo).

65

Un elemento de relleno es un contenedor de información que incluye un identificador (por ejemplo, el valor del elemento "id_syn_ele" indicado anteriormente) seguido de datos, que se denomina "datos de relleno". Los elementos de relleno se han usado históricamente para ajustar la velocidad de bits instantánea de las corrientes de bits que se van a transmitir a través de un canal de velocidad constante. Añadiendo la cantidad apropiada de datos de relleno a cada bloque, se puede lograr una velocidad de datos constante.

De acuerdo con las realizaciones de la invención, los datos de relleno pueden incluir una o más cargas útiles de extensión que amplían el tipo de datos (por ejemplo, metadatos) que pueden transmitirse en una corriente de bits. Un decodificador que recibe corrientes de bits con datos de relleno que contienen un nuevo tipo de datos puede ser usado opcionalmente por un dispositivo que recibe la corriente de bits (por ejemplo, un decodificador) para ampliar la funcionalidad del dispositivo. Por tanto, como puede apreciar un experto en la técnica, los elementos de relleno son un tipo especial de estructura de datos y son diferentes de las estructuras de datos que se usan típicamente para transmitir datos de audio (por ejemplo, cargas útiles de audio que contienen datos de canal).

En algunas realizaciones de la invención, el identificador usado para identificar un elemento de relleno puede consistir en un entero sin signo de tres bits transmitido primero el bit más significativo ("uimsbf") que tiene un valor de 0x6. En un bloque, pueden ocurrir varias instancias del mismo tipo de elemento sintáctico (por ejemplo, varios elementos de relleno).

Otro estándar para codificar corrientes de bits de audio es el estándar de codificación de audio y voz unificado (USAC) MPEG (ISO/IEC 23003-3: 2012). El estándar MPEG USAC describe la codificación y decodificación de contenido de audio usando procesamiento de replicación de banda espectral (incluido el procesamiento SBR como se describe en el estándar MPEG-4 AAC, y también incluye otras formas potenciadas de procesamiento de replicación de banda espectral). Este procesamiento aplica herramientas de replicación de banda espectral (a veces denominadas en el presente documento "herramientas de SBR potenciada" o "herramientas eSBR") de una versión ampliada y potenciada del conjunto de herramientas SBR descritas en el estándar MPEG-4 AAC. Por lo tanto, eSBR (como se define en el estándar USAC) es una mejora de SBR (como se define en el estándar MPEG-4 AAC).

En el presente documento, usamos la expresión "procesamiento de SBR potenciada" (o "procesamiento eSBR") para denotar el procesamiento de replicación de banda espectral usando al menos una herramienta eSBR (por ejemplo, al menos una herramienta eSBR que se describe o menciona en el estándar MPEG USAC) que no se describe ni se menciona en el estándar MPEG-4 AAC. Ejemplos de tales herramientas eSBR son la transposición armónica, el preprocesamiento adicional de parcheo QMF o "preaplanamiento" y la conformación de envolvente temporal de muestra entre subbandas o "inter-TES".

Una corriente de bits generada de acuerdo con el estándar MPEG USAC (a veces denominado en el presente documento como "una corriente de bits USAC") incluye contenido de audio codificado y típicamente incluye metadatos indicativos de cada tipo de procesamiento de replicación de banda espectral para ser aplicado por un decodificador para decodificar el contenido de audio de la corriente de bits USAC y/o metadatos que controlan tal procesamiento de replicación de banda espectral y/o es indicativo de al menos una característica o parámetro de al menos una herramienta SBR y/o herramienta eSBR que se empleará para decodificar el contenido de audio de la corriente de bits USAC.

En el presente documento, usamos la expresión "metadatos de SBR potenciada" (o "metadatos eSBR") para denotar los metadatos indicativos de cada tipo de procesamiento de replicación de banda espectral que aplicará un decodificador para decodificar el contenido de audio de una corriente de bits de audio codificada (por ejemplo, una corriente de bits USAC) y/o que controla dicho procesamiento de replicación de banda espectral, y/o es indicativo de al menos una característica o parámetro de al menos una herramienta SBR y/o herramienta eSBR que se usará para decodificar dicho contenido de audio, pero que no se describe o menciona en el estándar MPEG-4 AAC. Un ejemplo de metadatos eSBR son los metadatos (indicativos de, o para controlar, el procesamiento de replicación de banda espectral) que se describen o mencionan en el estándar MPEG USAC pero no en el estándar MPEG-4 AAC. Por lo tanto, los metadatos eSBR en el presente documento denotan metadatos que no son metadatos SBR, y los metadatos SBR en el presente documento denotan metadatos que no son metadatos eSBR.

Una corriente de bits USAC puede incluir tanto metadatos SBR como metadatos eSBR. Más específicamente, una corriente de bits USAC puede incluir metadatos eSBR que controlan el rendimiento del procesamiento eSBR por un decodificador, y metadatos SBR que controlan el rendimiento del procesamiento SBR por el decodificador. De acuerdo con las realizaciones típicas de la presente invención, los metadatos eSBR (por ejemplo, datos de configuración específicos de eSBR) se incluyen (de acuerdo con la presente invención) en una corriente de bits MPEG-4 AAC (por ejemplo, en el contenedor sbr_extension () al final de una carga útil SBR).

El rendimiento del procesamiento eSBR, durante la decodificación de una corriente de bits codificada usando un conjunto de herramientas eSBR (que comprende al menos una herramienta eSBR), mediante un decodificador regenera la banda de alta frecuencia de la señal de audio, basándose en la replicación de secuencias de armónicos que se truncaron durante la codificación. Tal procesamiento eSBR típicamente ajusta la envolvente espectral de la

banda de alta frecuencia generada y aplica filtrado inverso, y agrega componentes de ruido y sinusoidales para recrear las características espectrales de la señal de audio original.

5 De acuerdo con las realizaciones típicas de la invención, se incluyen metadatos eSBR (por ejemplo, se incluye una pequeña cantidad de bits de control que son metadatos eSBR) en uno o más de los segmentos de metadatos de una corriente de bits de audio codificada (por ejemplo, una corriente de bits MPEG-4 AAC) que también incluye datos de audio codificados en otros segmentos (segmentos de datos de audio). Típicamente, al menos uno de esos segmentos de metadatos de cada bloque de la corriente de bits es (o incluye) un elemento de relleno (incluido un identificador que indica el inicio del elemento de relleno), y los metadatos eSBR se incluyen en el elemento de relleno después del identificador.

15 La figura 1 es un diagrama de bloques de una cadena de procesamiento de audio de ejemplo (un sistema de procesamiento de datos de audio), en el que uno o más de los elementos del sistema pueden configurarse de acuerdo con una realización de la presente invención. El sistema incluye los siguientes elementos, acoplados juntos como se muestra: codificador 1, subsistema 2 de entrega, descodificador 3 y unidad 4 de posprocesamiento. En variaciones del sistema mostrado, se omiten uno o más de los elementos, o se incluyen unidades de procesamiento de datos de audio adicionales.

20 En algunas implementaciones, el codificador 1 (que opcionalmente incluye una unidad de preprocesamiento) está configurado para aceptar muestras PCM (dominio tiempo) que comprenden contenido de audio como entrada, y para emitir una corriente de bits de audio codificada (que tiene un formato compatible con el estándar MPEG-4 AAC) que es indicativo del contenido de audio. Los datos de la corriente de bits que son indicativos del contenido de audio se denominan a veces en el presente documento "datos de audio" o "datos de audio codificados". Si el codificador está configurado de acuerdo con una realización típica de la presente invención, la salida de corriente de bits de audio del codificador incluye metadatos eSBR (y típicamente también otros metadatos) así como datos de audio.

30 Una o más corrientes de bits de audio codificadas emitidas desde el codificador 1 pueden afirmarse en el subsistema 2 de entrega de audio codificado. El subsistema 2 está configurado para almacenar y/o entregar cada salida de corriente de bits codificada desde el codificador 1. Una salida de corriente de bits de audio codificada del codificador 1 puede ser almacenada por el subsistema 2 (por ejemplo, en forma de un disco DVD o Blu ray), o transmitida por el subsistema 2 (que puede implementar un enlace o red de transmisión), o pueden almacenarse y transmitirse ambos por el subsistema 2.

35 El descodificador 3 está configurado para descodificar una corriente de bits de audio codificada MPEG-4 AAC (generada por el codificador 1) que recibe a través del subsistema 2. En algunas realizaciones, el descodificador 3 está configurado para extraer metadatos eSBR de cada bloque de la corriente de bits y descodificar la corriente de bits (incluso realizando el procesamiento eSBR usando los metadatos eSBR extraídos) para generar datos de audio descodificados (por ejemplo, corrientes de muestras de audio PCM descodificadas). En algunas realizaciones, el descodificador 3 está configurado para extraer metadatos SBR de la corriente de bits (pero para ignorar los metadatos eSBR incluidos en la corriente de bits) y para descodificar la corriente de bits (incluso realizando el procesamiento SBR usando los metadatos SBR extraídos) para generar datos de audio descodificados (por ejemplo, corrientes de muestras de audio PCM descodificadas). Típicamente, el descodificador 3 incluye un búfer que almacena (por ejemplo, de manera no transitoria) segmentos de la corriente de bits de audio codificada recibida del subsistema 2.

50 La unidad 4 de posprocesamiento de la figura 1 está configurada para aceptar una corriente de datos de audio descodificados del descodificador 3 (por ejemplo, muestras de audio PCM descodificadas) y para realizar el posprocesamiento en el mismo. La unidad 4 de posprocesamiento también puede configurarse para reproducir el contenido de audio posprocesado (o el audio descodificado recibido del descodificador 3) para su reproducción por uno o más altavoces.

La figura 2 es un diagrama de bloques de un codificador 100. Cualquiera de los componentes o elementos del codificador 100 puede implementarse como uno o más procesos y/o uno o más circuitos (por ejemplo, ASIC, FPGA u otros circuitos integrados), en hardware, software o una combinación de hardware y software. El codificador 100 incluye el codificador 105, la etapa 107 de relleno/formateador, la etapa 106 de generación de metadatos y la memoria intermedia 109, conectados como se muestra. Típicamente también, el codificador 100 incluye otros elementos de procesamiento (no mostrados). El codificador 100 está configurado para convertir una corriente de bits de audio de entrada en una corriente de bits de salida codificada MPEG-4 AAC.

60 El generador 106 de metadatos está acoplado y configurado para generar (y/o pasar a la etapa 107) metadatos (incluidos metadatos eSBR y metadatos SBR) que se incluirán en la etapa 107 en la corriente de bits codificada que se emitirá desde el codificador 100.

El codificador 105 está acoplado y configurado para codificar (por ejemplo, realizando compresión sobre el mismo) los datos de audio de entrada, y para afirmar el audio codificado resultante en la etapa 107 para su inclusión en la corriente de bits codificada que se emitirá desde la etapa 107.

- 5 La etapa 107 está configurada para multiplexar el audio codificado del codificador 105 y los metadatos (incluidos los metadatos eSBR y los metadatos SBR) del generador 106 para generar la corriente de bits codificada que se emitirá desde la etapa 107, preferiblemente para que la corriente de bits codificada tenga el formato especificado por una de las realizaciones de la presente invención.
- 10 La memoria intermedia 109 está configurada para almacenar (por ejemplo, de manera no transitoria) al menos un bloque de la salida de la corriente de bits de audio codificada de la etapa 107, y una secuencia de los bloques de la corriente de bits de audio codificada se afirma desde la memoria intermedia 109 como salida del codificador 100 a un sistema de entrega.
- 15 La figura 3 es un diagrama de bloques de un sistema que incluye un descodificador 200 que es una realización de la unidad de procesamiento de audio de la invención y, opcionalmente, también un posprocesador 300 acoplado a la misma. Cualquiera de los componentes o elementos del descodificador 200 y posprocesador 300 puede implementarse como uno o más procesos y/o uno o más circuitos (por ejemplo, ASIC, FPGA u otros circuitos integrados), en hardware, software o una combinación de hardware y software. El descodificador 200 comprende
- 20 memoria intermedia 201, desformateador (analizador) 205 de carga útil de corriente de bits, subsistema 202 de descodificación de audio (a veces denominado etapa de descodificación "central" o subsistema de descodificación "central"), etapa 203 de procesamiento eSBR y etapa 204 de generación de bits de control, conectado como se muestra. Típicamente también, el descodificador 200 incluye otros elementos de procesamiento (no mostrados).
- 25 La memoria intermedia (búfer) 201 almacena (por ejemplo, de manera no transitoria) al menos un bloque de una corriente de bits de audio MPEG-4 AAC codificada recibido por el descodificador 200. En el funcionamiento del descodificador 200, se afirma una secuencia de los bloques de la corriente de bits desde el búfer 201 al desformateador 205.
- 30 En variaciones de la realización de la figura 3 (o la realización de la figura 4 que se describirá), una APU que no es un descodificador (por ejemplo, APU 500 de la figura 6) incluye una memoria intermedia (por ejemplo, una memoria intermedia idéntica a un búfer 201) que almacena (por ejemplo, de manera no transitoria) al menos un bloque de una corriente de bits de audio codificada (por ejemplo, una corriente de bits de audio MPEG-4 AAC) del mismo tipo recibido por el búfer 201 de la figura 3 o la figura 4 (es decir, una corriente de bits de audio codificada que incluye
- 35 metadatos eSBR).

Con referencia nuevamente a la figura 3, el desformateador 205 está acoplado y configurado para demultiplexar cada bloque de la corriente de bits para extraer metadatos SBR (incluidos datos de envolvente cuantificados) y metadatos eSBR (y típicamente también otros metadatos) de los mismos, para afirmar al menos los metadatos eSBR. y los metadatos SBR a la etapa 203 de procesamiento eSBR, y típicamente también para afirmar otros metadatos extraídos al subsistema 202 de descodificación (y opcionalmente también al generador 204 de bits de control). El desformateador 205 también está acoplado y configurado para extraer datos de audio de cada bloque de la corriente de bits y para afirmar los datos de audio extraídos en el subsistema 202 de descodificación (etapa de descodificación).

- 45 El sistema de la figura 3 opcionalmente también incluye posprocesador 300. El posprocesador 300 incluye una memoria intermedia (búfer) 301 y otros elementos de procesamiento (no mostrados) que incluyen al menos un elemento de procesamiento acoplado al búfer 301. El búfer 301 almacena (por ejemplo, de manera no transitoria) al menos un bloque (o trama) de los datos de audio descodificados recibidos por el posprocesador 300 desde el
- 50 descodificador 200. Los elementos de procesamiento del posprocesador 300 están acoplados y configurados para recibir y procesar de forma adaptativa una secuencia de los bloques (o tramas) de la salida de audio descodificada del búfer 301, usando la salida de metadatos del subsistema 202 de descodificación (y/o desformateador 205) y/o bits de control emitidos desde la etapa 204 del descodificador 200.

- 55 El subsistema 202 de descodificación de audio del descodificador 200 está configurado para descodificar los datos de audio extraídos por el analizador 205 (tal descodificación puede denominarse operación de descodificación "central") para generar datos de audio descodificados y para afirmar los datos de audio descodificados en la etapa 203 de procesamiento eSBR.. La descodificación se realiza en el dominio frecuencia y típicamente incluye cuantificación inversa seguida de procesamiento espectral. Típicamente, una etapa final de procesamiento en el
- 60 subsistema 202 aplica una transformación de dominio frecuencia a dominio tiempo a los datos de audio de dominio frecuencia descodificados, de modo que la salida del subsistema son datos de audio descodificados de dominio tiempo. La etapa 203 está configurada para aplicar herramientas SBR y herramientas eSBR indicadas por los metadatos eSBR y la eSBR (extraída por el analizador 205) a los datos de audio descodificados (es decir, para realizar el procesamiento SBR y eSBR en la salida del subsistema 202 de descodificación usando los metadatos
- 65 SBR y eSBR) para generar los datos de audio completamente descodificados que se envían (por ejemplo, al posprocesador 300) desde el descodificador 200. Típicamente, el descodificador 200 incluye una memoria

(accesible por el subsistema 202 y la etapa 203) que almacena los datos de audio desformateados y la salida de metadatos del desformateador 205, y la etapa 203 está configurada para acceder a los datos de audio y metadatos (incluidos los metadatos SBR y los metadatos eSBR) como necesario durante el procesamiento SBR y eSBR. Se puede considerar que el procesamiento SBR y el procesamiento eSBR en la etapa 203 son posprocesamiento en la salida del subsistema 202 de descodificación central. Opcionalmente, el descodificador 200 también incluye un subsistema de mezcla ascendente final (que puede aplicar herramientas estéreo paramétricas ("PS") definidas en el estándar MPEG-4 AAC, usando metadatos PS extraídos por desformateador 205 y/o bits de control generados en el subsistema 204) que es acoplado y configurado para realizar una mezcla ascendente en la salida de la etapa 203 para generar audio de mezcla ascendente totalmente descodificado que se emite desde el descodificador 200. Alternativamente, el posprocesador 300 está configurado para realizar una mezcla ascendente en la salida del descodificador 200 (por ejemplo, usando metadatos PS extraídos por el desformateador 205 y/o bits de control generados en el subsistema 204).

En respuesta a los metadatos extraídos por el desformateador 205, el generador 204 de bits de control puede generar datos de control, y los datos de control pueden usarse dentro del descodificador 200 (por ejemplo, en un subsistema de mezcla final) y/o afirmarse como salida del descodificador 200 (por ejemplo, para el posprocesador 300 para su uso en posprocesamiento). En respuesta a los metadatos extraídos de la corriente de bits de entrada (y opcionalmente también en respuesta a los datos de control), la etapa 204 puede generar (y afirmar al posprocesador 300) bits de control que indican que la salida de datos de audio descodificados de la etapa 203 de procesamiento eSBR debe someterse a un tipo específico de posprocesamiento. En algunas implementaciones, el descodificador 200 está configurado para afirmar los metadatos extraídos por el desformateador 205 de la corriente de bits de entrada al posprocesador 300, y el posprocesador 300 está configurado para realizar el posprocesamiento en la salida de datos de audio descodificados desde el descodificador 200 usando los metadatos.

La figura 4 es un diagrama de bloques de una unidad 210 de procesamiento de audio ("APU") que es otra realización de la unidad de procesamiento de audio de la invención. La APU 210 es un descodificador heredado que no está configurado para realizar procesamiento eSBR. Cualquiera de los componentes o elementos de APU 210 puede implementarse como uno o más procesos y/o uno o más circuitos (por ejemplo, ASIC, FPGA u otros circuitos integrados), en hardware, software o una combinación de hardware y software. La APU 210 comprende memoria intermedia 201, desformateador (analizador) 215 de carga útil de corriente de bits, subsistema 202 de descodificación de audio (a veces denominado etapa de descodificación "central" o subsistema de descodificación "central") y etapa 213 de procesamiento SBR, conectada como se muestra. Típicamente también, la APU 210 incluye otros elementos de procesamiento (no mostrados).

Los elementos 201 y 202 de la APU 210 son idénticos a los elementos numerados idénticamente del descodificador 200 (de la figura 3) y la descripción anterior de ellos no se repetirá. En el funcionamiento de la APU 210, una secuencia de bloques de una corriente de bits de audio codificada (una corriente de bits MPEG-4 AAC) recibida por la APU 210 se afirma desde el búfer 201 al desformateador 215.

El desformateador 215 está acoplado y configurado para demultiplexar cada bloque de la corriente de bits para extraer metadatos SBR (incluidos los datos de envoltorio cuantificados) y típicamente también otros metadatos de los mismos, pero para ignorar los metadatos eSBR que pueden incluirse en la corriente de bits de acuerdo con cualquier realización de la presente invención. El desformateador 215 está configurado para afirmar al menos los metadatos SBR en la etapa 213 de procesamiento SBR. El desformateador 215 también está acoplado y configurado para extraer datos de audio de cada bloque de la corriente de bits y para afirmar los datos de audio extraídos en el subsistema 202 de descodificación (etapa de descodificación).

El subsistema 202 de descodificación de audio del descodificador 200 está configurado para descodificar los datos de audio extraídos por el desformateador 215 (tal descodificación puede denominarse operación de descodificación "central") para generar datos de audio descodificados y para afirmar los datos de audio descodificados en la etapa 213 de procesamiento SBR. La descodificación se realiza en el dominio frecuencia. Típicamente, una etapa final de procesamiento en el subsistema 202 aplica una transformación de dominio frecuencia a dominio tiempo a los datos de audio de dominio frecuencia descodificados, de modo que la salida del subsistema son datos de audio descodificados de dominio tiempo. La etapa 213 está configurada para aplicar herramientas SBR (pero no herramientas eSBR) indicadas por los metadatos SBR (extraídos por el desformateador 215) a los datos de audio descodificados (es decir, para realizar el procesamiento SBR en la salida del subsistema 202 de descodificación usando los metadatos SBR) para generar los datos de audio totalmente descodificados que se envían (por ejemplo, al posprocesador 300) desde la APU 210. Típicamente, APU 210 incluye una memoria (accesible por el subsistema 202 y la etapa 213) que almacena los datos de audio desformateados y la salida de metadatos del desformateador 215, y la etapa 213 está configurada para acceder a los datos de audio y metadatos (incluidos los metadatos SBR) según sea necesario durante el procesamiento SBR. Se puede considerar que el procesamiento SBR en la etapa 213 es un posprocesamiento en la salida del subsistema 202 de descodificación central. Opcionalmente, la APU 210 también incluye un subsistema de mezcla ascendente final (que puede aplicar herramientas estéreo paramétricas ("PS") definidas en el estándar MPEG-4 AAC, usando metadatos PS extraídos por el desformateador 215) que está acoplado y configurado para realizar mezcla ascendente en la salida de la etapa 213 para generar audio mezclado y totalmente descodificado que se emite desde la APU 210. Alternativamente, se configura un posprocesador para

realizar una mezcla ascendente en la salida de la APU 210 (por ejemplo, usando metadatos PS extraídos por el desformateador 215 y/o bits de control generados en la APU 210).

5 Varias implementaciones del descodificador 200 y la APU 210 están configuradas para realizar diferentes realizaciones del método de la invención.

10 De acuerdo con algunas realizaciones, se incluyen metadatos eSBR (por ejemplo, se incluye una pequeña cantidad de bits de control que son metadatos eSBR) en una corriente de bits de audio codificada (por ejemplo, una corriente de bits MPEG-4 AAC), de modo que los descodificadores heredados (que no están configurados para analizar los metadatos eSBR, o para usar cualquier herramienta eSBR a la que pertenezcan los metadatos eSBR) puede ignorar los metadatos eSBR pero, sin embargo, descodificar la corriente de bits en la medida de lo posible sin usar los metadatos eSBR o cualquier herramienta eSBR a la que pertenezcan los metadatos eSBR, típicamente sin ninguna penalización significativa en la calidad del audio descodificado. Sin embargo, los descodificadores eSBR configurados para analizar la corriente de bits para identificar los metadatos eSBR y para usar al menos una herramienta eSBR en respuesta a los metadatos eSBR, disfrutarán de los beneficios de usar al menos una de tales herramientas eSBR. Por lo tanto, las realizaciones de la invención proporcionan un medio para transmitir eficazmente datos o metadatos de control de replicación de banda espectral potenciada (eSBR) de una manera compatible con versiones anteriores.

20 Típicamente, los metadatos eSBR en la corriente de bits son indicativos de (por ejemplo, son indicativos de al menos una característica o parámetro de) una o más de las siguientes herramientas eSBR (que se describen en el estándar MPEG USAC, y que pueden tener o no aplicado por un codificador durante la generación de la corriente de bits):

- transposición armónica;
- 25 • procesamiento adicional de parcheo QMF (preaplanamiento); y
- conformación de envolvente temporal de muestra entre subbandas o "inter-TES".

30 Por ejemplo, los metadatos eSBR incluidos en la corriente de bits pueden ser indicativos de los valores de los parámetros (descritos en el estándar MPEG USAC y en la presente divulgación): harmonicSBR[ch], sbrPatchingMode[ch], sbrOversamplingFlag[ch], sbrPitchInBins[ch], sbrPitchInBins[ch], bs_interTes, bs_temp_shape[ch][env], bs_inter_temp_shape_mode[ch][env] y bs_sbr_preprocessing.

35 En el presente documento, la notación X[ch], donde X es algún parámetro, denota que el parámetro pertenece al canal ("ch") del contenido de audio de una corriente de bits codificada a descodificar. Para simplificar, a veces omitimos la expresión [ch] y asumimos que el parámetro relevante pertenece a un canal de contenido de audio.

40 En el presente documento, la notación X[ch][env], donde X es algún parámetro, denota que el parámetro pertenece a la envolvente SBR ("env") del canal ("ch") del contenido de audio de una corriente de bits codificada a descodificar. Para simplificar, a veces omitimos las expresiones [env] y [ch], y asumimos que el parámetro relevante pertenece a una envolvente SBR de un canal de contenido de audio.

45 Como se señaló, el estándar MPEG USAC contempla que una corriente de bits USAC incluye metadatos eSBR que controlan el rendimiento del procesamiento eSBR mediante un descodificador. Los metadatos eSBR incluyen los siguientes parámetros de metadatos de un bit: harmonicSBR; bs_interTES; y bs_pvc.

50 El parámetro "harmonicSBR" indica el uso de parcheo armónico (transposición armónica) para SBR. Específicamente, harmonicSBR = 0 indica parcheo espectral no armónico como se describe en la sección 4.6.18.6.3 del estándar MPEG-4 AAC; y harmonicSBR = 1 indica parcheo armónico de SBR (del tipo usado en eSBR, como se describe en la sección 7.5.3 o 7.5.4 del estándar MPEG USAC). El parcheo armónico de SBR no se usa de acuerdo con la replicación de banda espectral que no es eSBR (es decir, SBR que no es eSBR). A lo largo de esta divulgación, el parcheo espectral se denomina una forma base de replicación de banda espectral, mientras que la transposición armónica se denomina forma potenciada de replicación de banda espectral.

55 El valor del parámetro "bs_interTES" indica el uso de la herramienta inter-TES de eSBR.

El valor del parámetro "bs_pvc" indica el uso de la herramienta PVC de eSBR.

60 Durante la descodificación de una corriente de bits codificada, el rendimiento de la transposición armónica durante una etapa de procesamiento eSBR de la descodificación (para cada canal, "ch", del contenido de audio indicado por la corriente de bits) se controla mediante los siguientes parámetros de metadatos eSBR: sbrPatchingMode[ch]; sbrOversamplingFlag[ch]; sbrPitchInBinsFlag[ch]; y sbrPitchInBins[ch].

El valor "sbrPatchingMode[ch]" indica el tipo de transpositor usado en eSBR: sbrPatchingMode[ch] = 1 indica parcheo no armónico como se describe en la sección 4.6.18.6.3 del estándar MPEG-4 AAC; sbrPatchingMode[ch] = 0 indica parcheo armónico de SBR como se describe en la sección 7.5.3 o 7.5.4 del estándar MPEG USAC.

- 5 El valor "sbrOversamplingFlag[ch]" indica el uso de sobremuestreo de dominio frecuencia adaptativo de señal en eSBR en combinación con el parcheo SBR armónico basado en DFT como se describe en la sección 7.5.3 del estándar MPEG USAC. Este indicador controla el tamaño de las DFT que se usan en el transpositor: 1 indica sobremuestreo de dominio frecuencia adaptativo de señal habilitado como se describe en la sección 7.5.3.1 del estándar MPEG USAC; 0 indica que el sobremuestreo del dominio frecuencia adaptativo de la señal está
- 10 inhabilitado, como se describe en la sección 7.5.3.1 del estándar MPEG USAC.

El valor "sbrPitchInBinsFlag [ch]" controla la interpretación del parámetro sbrPitchInBins[ch]: 1 indica que el valor en sbrPitchInBins[ch] es válido y mayor que cero; 0 indica que el valor de sbrPitchInBins[ch] se establece en cero.

- 15 El valor "sbrPitchInBins[ch]" controla la adición de términos de productos cruzados en el transpositor armónico de SBR. El valor sbrPitchInBins[ch] es un valor entero en el rango [0,127] y representa la distancia medida en intervalos de frecuencia para una DFT de 1536 líneas que actúa sobre la frecuencia de muestreo del codificador central.

- 20 En el caso de que una corriente de bits MPEG-4 AAC sea indicativa de un par de canales SBR cuyos canales no están acoplados (en lugar de un solo canal SBR), la corriente de bits es indicativa de dos instancias de la sintaxis anterior (para transposición armónica o no armónica), una para cada canal del sbr_channel_pair_element ().

- 25 La transposición armónica de la herramienta eSBR típicamente mejora la calidad de las señales musicales descodificadas a frecuencias de cruce relativamente bajas. La transposición de armónicos debe implementarse en el descodificador mediante una transposición de armónicos basada en DFT o basada en QMF. La transposición no armónica (es decir, la copia o el parcheo espectral heredado) generalmente mejora las señales de voz. Por lo tanto, un punto de partida en la decisión sobre qué tipo de transposición es preferible para codificar contenido de audio específico es seleccionar el método de transposición dependiendo de la detección de voz/música con transposición armónica empleada en el contenido musical y parcheo espectral en el contenido de voz.
- 30

- El rendimiento del preaplanamiento durante el procesamiento eSBR se controla mediante el valor de un parámetro de metadatos eSBR de un bit conocido como "bs_sbr_preprocessing", en el sentido de que el preaplanamiento se realiza o no dependiendo del valor de este bit único. Cuando se usa el algoritmo de parcheo SBR QMF, como se describe en la sección 4.6.18.6.3 del estándar MPEG-4 AAC, se puede realizar el paso de preaplanamiento (cuando lo indique el parámetro "bs_sbr_preprocessing") en un esfuerzo para evitar discontinuidades en la forma de la envolvente espectral de una señal de alta frecuencia que se introduce en un ajustador de envolvente posterior (el ajustador de envolvente realiza otra etapa del procesamiento eSBR). El preaplanamiento típicamente mejora el funcionamiento de la etapa de ajuste de envolvente subsiguiente, dando como resultado una señal de banda alta que se percibe como más estable.
- 35
- 40

- El rendimiento de la conformación de envolvente temporal de muestra entre subbandas (la herramienta "inter-TES"), durante el procesamiento eSBR en un descodificador, se controla mediante los siguientes parámetros de metadatos eSBR para cada envolvente de SBR ("env") de cada canal ("ch") del contenido de audio de una corriente de bits USAC que se está descodificando: bs_temp_shape [ch] [env]; y bs_inter_temp_shape_mode [ch] [env].
- 45

- La herramienta inter-TES procesa las muestras de subbanda QMF después del ajustador de envolvente. Este paso de procesamiento conforma la envolvente temporal de la banda de frecuencia más alta con una granularidad temporal más fina que la del ajustador de envolvente. Al aplicar un factor de ganancia a cada muestra de subbanda QMF en una envolvente SBR, inter-TES conforma la envolvente temporal entre las muestras de subbanda QMF.
- 50

- El parámetro "bs_temp_shape [ch] [env]" es un indicador que señala el uso de inter-TES. El parámetro "bs_inter_temp_shape_mode [ch] [env]" indica (como se define en el estándar MPEG USAC) los valores del parámetro γ en inter-TES.

- 55 Se espera que el requisito general de velocidad de bits para incluir en una corriente de bits MPEG-4 AAC metadatos eSBR indicativos de las herramientas eSBR mencionadas anteriormente (transposición armónica, preaplanamiento e inter_TES) sea del orden de unos pocos cientos de bits por segundo porque sólo los datos de control diferencial necesarios para realizar el procesamiento eSBR se transmiten de acuerdo con algunas realizaciones de la invención. Los descodificadores heredados pueden ignorar esta información porque se incluye de una manera compatible con versiones anteriores (como se explicará más adelante). Por lo tanto, el efecto perjudicial sobre la velocidad de bits asociada con la inclusión de metadatos eSBR es insignificante, por varias razones, incluidas las siguientes:
- 60

- La penalización de la velocidad de bits (debido a que se incluyen los metadatos eSBR) es una fracción muy pequeña de la velocidad de bits total porque solo se transmiten los datos de control diferencial necesarios para realizar el procesamiento eSBR (y no una transmisión simultánea de los datos de control de SBR);
- 65

El ajuste de la información de control relacionada con SBR típicamente no depende de los detalles de la transposición; y

- 5 La herramienta inter-TES (empleada durante el procesamiento eSBR) realiza un posprocesamiento de un solo extremo de la señal transpuesta.

Por tanto, las realizaciones de la invención proporcionan un medio para transmitir eficazmente datos o metadatos de control de replicación de banda espectral potenciada (eSBR) de una manera compatible con versiones anteriores. Esta transmisión eficiente de los datos de control de eSBR reduce los requisitos de memoria en los descodificadores y transcodificadores que emplean aspectos de la invención, mientras que no tiene ningún efecto adverso tangible sobre la velocidad de bits. Además, la complejidad y los requisitos de procesamiento asociados con la realización de eSBR de acuerdo con las realizaciones de la invención también se reducen porque los datos de SBR deben procesarse solo una vez y no transmitirse simultáneamente, lo que sería el caso si eSBR se tratara como un tipo de objeto completamente separado en MPEG-4 AAC en lugar de integrarse en el códec MPEG-4 AAC de forma compatible con versiones anteriores.

10 A continuación, con referencia a la figura 7, describimos elementos de un bloque ("raw_data_block") de una corriente de bits MPEG-4 AAC en el que se incluyen metadatos eSBR de acuerdo con algunas realizaciones de la presente invención. La figura 7 es un diagrama de un bloque (un "raw_data_block") de la corriente de bits MPEG-4 AAC, que muestra algunos de sus segmentos.

Un bloque de una corriente de bits MPEG-4 AAC puede incluir al menos un "single_channel_element ()" (por ejemplo, el elemento de un solo canal que se muestra en la figura 7), y/o al menos un "channel_pair_element ()" (no se muestra específicamente en la figura 7 aunque puede estar presente), incluidos los datos de audio de un programa de audio. El bloque también puede incluir una serie de elementos de relleno (por ejemplo, elemento 1 de relleno y/o elemento 2 de relleno de la figura 7) que incluyen datos (por ejemplo, metadatos) relacionados con el programa. Cada "single_channel_element ()" incluye un identificador (por ejemplo, "ID1" de la figura 7) que indica el inicio de un elemento de canal único, y puede incluir datos de audio indicativos de un canal diferente de un programa de audio multicanal. Cada elemento de par de canales incluye un identificador (no mostrado en la figura 7) que indica el inicio de un elemento de par de canales, y puede incluir datos de audio indicativos de dos canales del programa.

Un fill_element (denominado en el presente documento elemento de relleno) de una corriente de bits MPEG-4 AAC incluye un identificador ("ID2" de la figura 7) que indica el inicio de un elemento de relleno y datos de relleno después del identificador. El identificador ID2 puede consistir en un entero sin signo de tres bits transmitido primero el bit más significativo ("uimbsbf") que tiene un valor de 0x6. Los datos de relleno pueden incluir un elemento extension_payload () (a veces denominado en el presente documento como una carga útil de extensión) cuya sintaxis se muestra en la tabla 4.57 del estándar MPEG-4 AAC. Existen varios tipos de cargas útiles de extensión y se identifican mediante el parámetro "extension_type", que es un entero sin signo de cuatro bits transmitido primero el bit más significativo ("uimbsbf").

Los datos de relleno (por ejemplo, una carga útil de extensión de los mismos) pueden incluir un encabezado o identificador (por ejemplo, "header1" de la figura 7) que indica un segmento de datos de relleno que es indicativo de un objeto SBR (es decir, el encabezado inicializa un de "tipo de objeto SBR", denominado sbr_extension_data () en el estándar MPEG-4 AAC). Por ejemplo, una carga útil de extensión de replicación de banda espectral (SBR) se identifica con el valor de '1101' o '1110' para el campo extension_type en el encabezado, con el identificador '1101' que identifica una carga útil de extensión con datos SBR y '1110' que identifica una extensión de la carga útil con datos SBR con una verificación de redundancia cíclica (CRC) para verificar la exactitud de los datos SBR.

50 Cuando el encabezado (por ejemplo, el campo extension_type) inicializa un tipo de objeto SBR, los metadatos SBR (a veces denominados en el presente documento "datos de replicación de banda espectral" y denominados sbr_data () en el estándar MPEG-4 AAC) siguen al encabezado, y al menos un elemento de extensión de replicación de banda espectral (por ejemplo, el "elemento de extensión de SBR" del elemento 1 de relleno de la figura 7) puede seguir los metadatos SBR. Tal elemento de extensión de replicación de banda espectral (un segmento de la corriente de bits) se denomina contenedor "sbr_extension ()" en el estándar MPEG-4 AAC. Un elemento de extensión de replicación de banda espectral incluye opcionalmente un encabezado (por ejemplo, "encabezado de extensión SBR" del elemento 1 de relleno de la figura 7).

60 El estándar MPEG-4 AAC contempla que un elemento de extensión de replicación de banda espectral puede incluir datos PS (estéreo paramétrico) para datos de audio de un programa. El estándar MPEG-4 AAC contempla que cuando el encabezado de un elemento de relleno (por ejemplo, de una carga útil de extensión del mismo) inicializa un tipo de objeto SBR (al igual que "header1" de la figura 7) y un elemento de extensión de replicación de banda espectral del elemento de relleno incluye datos PS, el elemento de relleno (por ejemplo, la carga útil de extensión del mismo) incluye datos de replicación de banda espectral y un parámetro "bs_extension_id" cuyo valor (es decir,

bs_extension_id = 2) indica que los datos PS están incluidos en un elemento de extensión de replicación de banda espectral del elemento de relleno.

5 De acuerdo con algunas realizaciones de la presente invención, los metadatos eSBR (por ejemplo, un indicador indicativo de si el procesamiento de replicación de banda espectral potenciada (eSBR) se realizará en el contenido de audio del bloque) se incluyen en un elemento de extensión de replicación de banda espectral de un elemento de relleno. Por ejemplo, tal indicador se indica en el elemento 1 de relleno de la figura 7, donde el indicador aparece después del encabezado (el "encabezado de extensión SBR" del elemento 1 de relleno) del "elemento de extensión SBR" del elemento 1 de relleno. Opcionalmente, tal indicador y los metadatos eSBR adicionales se incluyen en un elemento de extensión de replicación de banda espectral después del encabezado del elemento de extensión de replicación de banda espectral (por ejemplo, en el elemento de extensión SBR del elemento 1 de relleno en la figura 7, después del encabezado de extensión SBR). De acuerdo con algunas realizaciones de la presente invención, un elemento de relleno que incluye metadatos eSBR también incluye un parámetro "bs_extension_id" cuyo valor (por ejemplo, bs_extension_id = 3) indica que los metadatos eSBR están incluidos en el elemento de relleno y que el procesamiento eSBR ha de realizarse en el contenido de audio del bloque relevante.

20 De acuerdo con algunas realizaciones de la invención, los metadatos eSBR se incluyen en un elemento de relleno (por ejemplo, el elemento 2 de relleno de la figura 7) de una corriente de bits MPEG-4 AAC que no sea un elemento de extensión de replicación de banda espectral (elemento de extensión de SBR) del elemento de relleno. Esto se debe a que los elementos de relleno que contienen un extension_payload () con datos SBR o datos SBR con una CRC no contienen ninguna otra carga útil de extensión de ningún otro tipo de extensión. Por lo tanto, en las realizaciones en las que los metadatos eSBR se almacenan en su propia carga útil de extensión, se usa un elemento de relleno independiente para almacenar los metadatos eSBR. Tal elemento de relleno incluye un identificador (por ejemplo, "ID2" de la figura 7) que indica el inicio de un elemento de relleno y datos de relleno después del identificador. Los datos de relleno pueden incluir un elemento extension_payload () (a veces denominado en el presente documento como una carga útil de extensión) cuya sintaxis se muestra en la tabla 4.57 del estándar MPEG-4 AAC. Los datos de relleno (por ejemplo, una carga útil de extensión de los mismos) incluyen un encabezado (por ejemplo, "header2" del elemento 2 de relleno de la figura 7) que es indicativo de un objeto eSBR (es decir, el encabezado inicializa un tipo de objeto de replicación de banda espectral potenciada (eSBR)), y los datos de relleno (por ejemplo, una carga útil de extensión de los mismos) incluyen metadatos eSBR después del encabezado. Por ejemplo, el elemento 2 de relleno de la figura 7 incluye dicho encabezado ("header2") y también incluye, después del encabezado, metadatos eSBR (es decir, el "indicador" en el elemento 2 de relleno, que indica si la replicación de banda espectral potenciada (eSBR) se debe realizar en el contenido de audio del bloque). Opcionalmente, también se incluyen metadatos eSBR adicionales en los datos de relleno del elemento 2 de relleno de la figura 7, después del header2. En las realizaciones que se describen en el presente párrafo, el encabezado (por ejemplo, header2 de la figura 7) tiene un valor de identificación que no es uno de los valores convencionales especificados en la tabla 4.57 del estándar MPEG-4 AAC, y en cambio es indicativo de una carga útil de extensión eSBR (de modo que el campo extension_type del encabezado indica que los datos de relleno incluyen metadatos eSBR).

40 En una primera clase de realizaciones, la invención es una unidad de procesamiento de audio (por ejemplo, un descodificador), que comprende:

45 una memoria (por ejemplo, el búfer 201 de la figura 3 o 4) configurada para almacenar al menos un bloque de una corriente de bits de audio codificada (por ejemplo, al menos un bloque de una corriente de bits MPEG-4 AAC);

50 un desformateador de carga útil de corriente de bits (por ejemplo, el elemento 205 de la figura 3 o el elemento 215 de la figura 4) acoplado a la memoria y configurado para demultiplexar al menos una porción de dicho bloque de la corriente de bits; y

un subsistema de descodificación (por ejemplo, elementos 202 y 203 de la figura 3, o elementos 202 y 213 de la figura 4), acoplado y configurado para descodificar al menos una porción del contenido de audio de dicho bloque de la corriente de bits, en el que el bloque incluye:

55 un elemento de relleno, que incluye un identificador que indica un inicio del elemento de relleno (por ejemplo, el identificador "id_syn_ele" que tiene el valor 0x6, de la tabla 4.85 del estándar MPEG-4 AAC), y datos de relleno después del identificador, donde los datos de relleno incluyen:

60 al menos un indicador que identifica si el procesamiento de replicación de banda espectral potenciada (eSBR) debe realizarse en el contenido de audio del bloque (por ejemplo, usando datos de replicación de banda espectral y metadatos eSBR incluidos en el bloque).

65 El indicador son los metadatos eSBR y un ejemplo del indicador es el indicador sbrPatchingMode. Otro ejemplo del indicador es el indicador harmonicSBR. Ambos indicadores indican si se va a realizar una forma base de replicación de banda espectral o una forma potenciada de replicación espectral en los datos de audio del bloque. La forma base

de la replicación espectral es el parcheo espectral, y la forma potenciada de replicación de banda espectral es la transposición armónica.

5 En algunas realizaciones, los datos de relleno también incluyen metadatos eSBR adicionales (es decir, metadatos eSBR distintos del indicador).

La memoria puede ser una memoria intermedia (por ejemplo, una implementación del búfer 201 de la figura 4) que almacena (por ejemplo, de manera no transitoria) al menos dicho bloque de la corriente de bits de audio codificada.

10 Se estima que la complejidad del rendimiento del procesamiento eSBR (usando las herramientas de transposición armónica, preaplanamiento e inter_TES de eSBR) por un descodificador eSBR durante la descodificación de una corriente de bits MPEG-4 AAC que incluye metadatos eSBR (indicativo de estas herramientas eSBR) sería el siguiente (para descodificación típica con los parámetros indicados):

- 15 • transposición armónica (16 kbps, 14400/28800 Hz)
 - basado en DFT: 3,68 WMOPS (millones de operaciones ponderadas por segundo);
 - basado en QMF: 0,98 WMOPS;
- 20 • preprocesamiento de parcheo QMF (preaplanamiento): 0.1WMOPS; y
- conformación de envolvente temporal de muestra entre subbandas (inter-TES): como máximo 0,16 WMOPS.

25 Se sabe que la transposición basada en DFT funciona típicamente mejor que la transposición basada en QMF para transitorios.

De acuerdo con algunas realizaciones de la presente invención, un elemento de relleno (de una corriente de bits de audio codificada) que incluye metadatos eSBR también incluye un parámetro (por ejemplo, un parámetro "bs_extension_id") cuyo valor (por ejemplo, bs_extension_id = 3) indica que los metadatos eSBR están incluidos en el elemento de relleno y que el procesamiento eSBR ha de realizarse en el contenido de audio del bloque relevante, y/o un parámetro (por ejemplo, el mismo parámetro "bs_extension_id") cuyo valor (por ejemplo, bs_extension_id = 2) indica que un contenedor sbr_extension () del elemento de relleno incluye datos PS. Por ejemplo, como se indica en la tabla 1 a continuación, tal parámetro que tiene el valor bs_extension_id = 2 puede indicar que un contenedor sbr_extension () del elemento de relleno incluye datos PS, y dicho parámetro que tiene el valor bs_extension_id = 3 puede indicar que un contenedor sbr_extension () del elemento de relleno incluye metadatos eSBR:

Tabla 1

| bs_extension_id | Significado | |
|-----------------|-------------------|--|
| 0 | Reservado | |
| 1 | Reservado | |
| 2 | EXTENSION_ID_PS | |
| 3 | EXTENSION_ID_ESBR | |

40 De acuerdo con algunas realizaciones de la invención, la sintaxis de cada elemento de extensión de replicación de banda espectral que incluye metadatos eSBR y/o datos PS es como se indica en la tabla 2 a continuación (en la que "sbr_extension ()" denota un contenedor que es el elemento de extensión de replicación de banda espectral, "bs_extension_id" es como se describe en la tabla 1 anterior, "ps_data" denota datos PS y "esbr_data" denota metadatos eSBR):

45 Tabla 2

| | | |
|---|--|--------|
| sbr_extension(bs_extension_id, num_bits_left) | | |
| { | | |
| switch (bs_extension_id){ | | |
| case EXTENSION_ID_PS: | | |
| num_bits_left-=ps_data(); | | Nota 1 |
| break; | | |
| case EXTENSION_ID_ESBR: | | |

| | | |
|--|--|--------|
| num_bits_left -= esbr_data(); | | Nota 2 |
| break; | | |
| default: | | |
| bs_fill_bits; | | Nota 3 |
| num_bits_left = 0; | | |
| break; | | |
| } | | |
| } | | |
| Nota 1: ps_data() devuelve el número de bits leídos. | | |
| Nota 2: esbr_data() devuelve el número de bits leídos. | | |
| Nota 3: el parámetro bs_fill_bits comprende N bits, donde N = num_bits_left. | | |

En una realización de ejemplo, el esbr_data () al que se hace referencia en la tabla 2 anterior es indicativo de los valores de los siguientes parámetros de metadatos:

- 5 1. cada uno de los parámetros de metadatos de un bit descritos anteriormente "harmonicSBR"; "bs_interTES"; y "bs_sbr_preprocessing";
- 10 2. para cada canal ("ch") de contenido de audio de la corriente de bits codificada a descodificar, cada uno de los parámetros descritos anteriormente: "sbrPatchingMode[ch]"; "sbrOversamplingFlag[ch]"; "sbrPitchInBinsFlag [ch]"; y "sbrPitchInBins[ch]"; y
- 15 3. para cada envolvente SBR ("env") de cada canal ("ch") de contenido de audio de la corriente de bits codificada a descodificar, cada uno de los parámetros descritos anteriormente: "bs_temp_shape [ch] [env]"; y "bs_inter_temp_shape_mode [ch] [env]".

Por ejemplo, en algunas realizaciones, esbr_data () puede tener la sintaxis indicada en la tabla 3, para indicar estos parámetros de metadatos:

Tabla 3

| | | |
|----------------------------------|---|--|
| esbr_data() | | |
| { | | |
| harmonicSBR; | 1 | |
| bs_interTes; | 1 | |
| bs_sbr_preprocessing; | 1 | |
| if (harmonicSBR) { | | |
| if (sbrPatchingMode[0] == 0){ | | |
| sbrOversamplingFlag[0]; | 1 | |
| if (sbrPichinBinsFlag[0]) | 1 | |
| sbrPichinBins[0]; | 7 | |
| Else | | |
| sbrPichinBins[0@] = 0; | | |
| } else { | | |
| sbrOversamplingFlag[0] = 0; | | |
| sbrPithinBins[0] = 0; | | |
| } | | |

| | | |
|---|---|--|
| } | | |
| | | |
| if (bs_interTes) { | | |
| /* se implementa un bucle sobre ch y env */ | | |
| bs_temp_shape[ch][env]; | 1 | |
| if (bs_temp_shape[ch][env]) { | | |
| bs_inter_temp_shape_mode[ch][env]; | 2 | |
| } | | |
| } | | |
| } | | |

En la tabla 3, el número en la columna central indica el número de bits del parámetro correspondiente en la columna de la izquierda.

- 5 La sintaxis anterior permite una implementación eficiente de una forma potenciada de replicación de banda espectral, como la transposición armónica, como una extensión de un descodificador heredado. Específicamente, los datos eSBR de la tabla 3 incluyen solo aquellos parámetros necesarios para realizar la forma potenciada de replicación de banda espectral que no están ya soportados en la corriente de bits o directamente derivables de los parámetros ya soportados en la corriente de bits. Todos los demás parámetros y datos de procesamiento necesarios para realizar la forma potenciada de replicación de banda espectral se extraen de parámetros preexistentes en ubicaciones ya definidas en la corriente de bits. Esto contrasta con una implementación alternativa (y menos eficiente) que simplemente transmite todos los metadatos de procesamiento usados para la replicación de banda espectral potenciada.
- 10
- 15 Por ejemplo, un descodificador compatible con MPEG-4 HE-AAC o HE-AAC v2 puede ampliarse para incluir una forma potenciada de replicación de banda espectral, como la transposición armónica. Esta forma potenciada de replicación de banda espectral se suma a la forma base de replicación de banda espectral que ya es compatible con el descodificador. En el contexto de un descodificador compatible con MPEG-4 HE-AAC o HE-AAC v2, esta forma base de replicación de banda espectral es la herramienta SBR de parcheo espectral QMF como se define en la sección 4.6.18 del estándar MPEG-4 AAC.
- 20

Al realizar la forma potenciada de replicación de banda espectral, un descodificador HE-AAC extendido puede reutilizar muchos de los parámetros de la corriente de bits ya incluidos en la carga útil de extensión SBR de la corriente de bits. Los parámetros específicos que pueden reutilizarse incluyen, por ejemplo, los diversos parámetros que determinan la tabla de bandas de frecuencia maestra. Estos parámetros incluyen bs_start_freq (parámetro que determina el inicio de la tabla de frecuencia maestra), bs_stop_freq (parámetro que determina la parada de la tabla de frecuencia maestra), bs_freq_scale (parámetro que determina el número de bandas de frecuencia por octava) y bs_alter_scale (parámetro que modifica la escala de las bandas de frecuencia). Los parámetros que pueden reutilizarse también incluyen parámetros que determinan la tabla de bandas de ruido (bs_noise_bands) y los parámetros de la tabla de bandas de limitador (bs_limiter_bands).

25

30

Además de los numerosos parámetros, un descodificador HE-AAC extendido también puede reutilizar otros elementos de datos cuando se realiza una forma potenciada de replicación de banda espectral de acuerdo con realizaciones de la invención. Por ejemplo, los datos de la envolvente y los datos del piso de ruido también pueden extraerse de los datos bs_data_env y bs_noise_env y usarse durante la forma potenciada de replicación de banda espectral.

35

En esencia, estas realizaciones explotan los parámetros de configuración y los datos de envolvente ya soportados por un descodificador HE-AAC o HE-AAC v2 heredado en la carga útil de extensión SBR para permitir una forma potenciada de replicación de banda espectral que requiere la menor cantidad posible de datos transmitidos adicionales. En consecuencia, los descodificadores extendidos que soportan una forma potenciada de replicación de banda espectral pueden crearse de una manera muy eficiente confiando en elementos de corriente de bits ya definidos (por ejemplo, aquellos en la carga útil de extensión SBR) y agregando solo los parámetros necesarios para soportar la forma potenciada de replicación de banda espectral (en una carga útil de extensión de elemento de relleno). Esta característica de reducción de datos combinada con la ubicación de los parámetros recién agregados en un campo de datos reservado, como un contenedor de extensión, reduce sustancialmente las barreras para crear un descodificador que soporte una replicación de banda espectral potenciada al garantizar que la corriente de bits

40

45

sea compatible con versiones anteriores con el decodificador heredado que no soporte la forma potenciada de replicación de banda espectral.

5 En algunos ejemplos que no están de acuerdo con la invención y están presentes solo con fines ilustrativos, un método que incluye un paso de codificación de datos de audio para generar una corriente de bits codificada (por ejemplo, una corriente de bits MPEG-4 AAC), incluyendo la inclusión de metadatos eSBR en al menos un segmento de al menos un bloque de la corriente de bits codificada y datos de audio en al menos otro segmento del bloque. En ejemplos típicos, el método incluye un paso de multiplexar los datos de audio con los metadatos eSBR en cada bloque de la corriente de bits codificada. En la decodificación típica de la corriente de bits codificada en un decodificador eSBR, el decodificador extrae los metadatos eSBR de la corriente de bits (incluso analizando y demultiplexando los metadatos eSBR y los datos de audio) y usa los metadatos eSBR para procesar los datos de audio para generar una corriente de datos de audio descodificados.

15 Otro aspecto de la invención es un decodificador eSBR configurado para realizar procesamiento eSBR (por ejemplo, usando al menos una de las herramientas eSBR conocidas como transposición armónica, preaplamiento o inter_TES) durante la decodificación de una corriente de bits de audio codificada (por ejemplo, una corriente de bits MPEG-4 AAC) que no incluye metadatos eSBR. Se describirá un ejemplo de tal decodificador con referencia a la figura 5.

20 El decodificador eSBR 400 de la figura 5 incluye memoria intermedia 201 (que es idéntica a la memoria 201 de las figuras 3 y 4), desformateador 215 de carga útil de corriente de bits (que es idéntico al desformateador 215 de la figura 4), subsistema 202 de descodificación de audio (a veces denominado etapa de descodificación "central" o subsistema de descodificación "central", y que es idéntico al subsistema 202 de descodificación de la figura 3), el subsistema 401 de generación de datos de control eSBR y la etapa 203 de procesamiento eSBR (que es idéntica a etapa 203 de la figura 3), conectado como se muestra. También típicamente, el decodificador 400 incluye otros elementos de procesamiento (no mostrados).

25 En el funcionamiento del decodificador 400, una secuencia de bloques de una corriente de bits de audio codificada (una corriente bits MPEG-4 AAC) recibida por el decodificador 400 se afirma desde el búfer 201 al desformateador 215.

30 El desformateador 215 está acoplado y configurado para demultiplexar cada bloque de la corriente de bits para extraer metadatos SBR (incluidos datos de envolvente cuantificados) y típicamente también otros metadatos de los mismos. El desformateador 215 está configurado para afirmar al menos los metadatos SBR en la etapa 203 de procesamiento eSBR. El desformateador 215 también está acoplado y configurado para extraer datos de audio de cada bloque de la corriente de bits y para afirmar los datos de audio extraídos en el subsistema 202 de descodificación (etapa de descodificación).

35 El subsistema 202 de descodificación de audio del decodificador 400 está configurado para descodificar los datos de audio extraídos por el desformateador 215 (tal descodificación puede denominarse operación de descodificación "central") para generar datos de audio descodificados y para afirmar los datos de audio descodificados en la etapa 203 de procesamiento eSBR. La descodificación se realiza en el dominio frecuencia. Típicamente, una etapa final de procesamiento en el subsistema 202 aplica una transformación de dominio frecuencia a dominio tiempo a los datos de audio de dominio frecuencia descodificados, de modo que la salida del subsistema son datos de audio descodificados de dominio tiempo. La etapa 203 está configurada para aplicar herramientas SBR (y herramientas eSBR) indicadas por los metadatos SBR (extraídos por el desformateador 215) y por los metadatos eSBR generados en el subsistema 401, a los datos de audio descodificados (es decir, para realizar el procesamiento SBR y eSBR en la salida del subsistema 202 de descodificación usando los metadatos SBR y eSBR) para generar los datos de audio completamente descodificados que se emiten desde el decodificador 400. Típicamente, el decodificador 400 incluye una memoria (accesible por el subsistema 202 y la etapa 203) que almacena los datos de audio desformateados y la salida de metadatos del desformateador 215 (y opcionalmente también el subsistema 401), y la etapa 203 está configurada para acceder a los datos de audio y metadatos según sea necesario durante el procesamiento SBR y eSBR. Se puede considerar que el procesamiento SBR en la etapa 203 es un posprocesamiento en la salida del subsistema 202 de descodificación central. Opcionalmente, el decodificador 400 también incluye un subsistema de mezcla ascendente final (que puede aplicar herramientas estéreo paramétricas ("PS") definidas en el estándar MPEG-4 AAC, usando metadatos PS extraídos por el desformateador 215) que está acoplado y configurado para realizar mezcla ascendente en la salida de la etapa 203 para generar audio mezclado y totalmente descodificado que se emite desde la APU 210.

40 El subsistema 401 de generación de datos de control de la figura 5 está acoplado y configurado para detectar al menos una propiedad de la corriente de bits de audio codificada que se va a descodificar, y para generar datos de control eSBR (que pueden ser o incluir metadatos eSBR de cualquiera de los tipos incluidos en corrientes de bits de audio codificadas de acuerdo con otras realizaciones de la invención) en respuesta a al menos un resultado del paso de detección. Los datos de control de eSBR se afirman en la etapa 203 para activar la aplicación de herramientas eSBR individuales o combinaciones de herramientas eSBR al detectar una propiedad específica (o combinación de propiedades) de la corriente de bits, y/o para controlar la aplicación de tales herramientas eSBR. Por ejemplo, para

controlar el rendimiento del procesamiento eSBR mediante la transposición armónica, algunas realizaciones del subsistema 401 de generación de datos de control incluirían: un detector de música (por ejemplo, una versión simplificada de un detector de música convencional) para configurar el parámetro `sbrPatchingMode[ch]` (y afirmar el parámetro establecido en la etapa 203) en respuesta a la detección de que la corriente de bits es o no indicativa de música; un detector de transitorios para establecer el parámetro `sbrOversamplingFlag[ch]` (y afirmar el parámetro establecido en la etapa 203) en respuesta a la detección de la presencia o ausencia de transitorios en el contenido de audio indicado por la corriente de bits; y/o un detector de tono para establecer los parámetros `sbrPitchInBinsFlag[ch]` y `sbrPitchInBins[ch]` (y afirmar los parámetros establecidos en la etapa 203) en respuesta a la detección del tono del contenido de audio indicado por la corriente de bits. Otros aspectos de la invención son los métodos de descodificación de corriente de bits de audio realizados por cualquier realización del descodificador de la invención descrito en este párrafo y en el párrafo anterior.

Los aspectos de la invención incluyen un método de descodificación del tipo para el que está configurada (por ejemplo, programada) cualquier realización de la APU, sistema o dispositivo de la invención. Otros aspectos de la invención incluyen un sistema o dispositivo configurado (por ejemplo, programado) para realizar cualquier realización del método de la invención, y un medio legible por computadora (por ejemplo, un disco) que almacena código (por ejemplo, de manera no transitoria) para implementar cualquier realización del método de la invención o pasos del mismo. Por ejemplo, el sistema de la invención puede ser o incluir un procesador de propósito general programable, procesador de señal digital o microprocesador, programado con software o firmware y/o configurado de otra manera para realizar cualquiera de una variedad de operaciones sobre datos, incluida una realización del método de la invención o pasos del mismo. Tal procesador de propósito general puede ser o incluir un sistema informático que incluye un dispositivo de entrada, una memoria y un circuito de procesamiento programado (y/o configurado de otra manera) para realizar una realización del método de la invención (o pasos del mismo) en respuesta a los datos afirmados en el mismo.

Las realizaciones de la presente invención pueden implementarse en hardware, firmware o software, o una combinación de ambos (por ejemplo, como una matriz lógica programable). A menos que se especifique lo contrario, los algoritmos o procesos incluidos como parte de la invención no están intrínsecamente relacionados con ninguna computadora u otro aparato en particular. En particular, se pueden usar varias máquinas de uso general con programas escritos de acuerdo con las enseñanzas del presente documento, o puede ser más conveniente construir aparatos más especializados (por ejemplo, circuitos integrados) para realizar los pasos requeridos del método. Por lo tanto, la invención puede implementarse en uno o más programas informáticos que se ejecutan en uno o más sistemas informáticos programables (por ejemplo, el descodificador 200 de la figura 3 (o un elemento del mismo) o el descodificador 210 de la figura 4 (o un elemento del mismo), o descodificador 400 de la figura 5 (o un elemento del mismo)) cada uno de los cuales comprende al menos un procesador, al menos un sistema de almacenamiento de datos (que incluye memoria volátil y no volátil y/o elementos de almacenamiento), al menos un dispositivo o puerto de entrada, y al menos un dispositivo o puerto de salida. El código de programa se aplica a los datos de entrada para realizar las funciones descritas en el presente y generar información de salida. La información de salida se aplica a uno o más dispositivos de salida, de manera conocida.

Cada uno de estos programas puede implementarse en cualquier lenguaje informático deseado (incluyendo máquina, ensamblaje o lenguajes de programación de procedimiento, lógico u orientado a objetos de alto nivel) para comunicarse con un sistema informático. En cualquier caso, el lenguaje puede ser un lenguaje compilado o interpretado.

Por ejemplo, cuando se implementan mediante secuencias de instrucciones de software informático, varias funciones y pasos de realizaciones de la invención pueden implementarse mediante secuencias de instrucciones de software multiproceso que se ejecutan en hardware de procesamiento de señales digitales adecuado, en cuyo caso los diversos dispositivos, pasos y funciones de las realizaciones pueden corresponder a porciones de las instrucciones del software.

Cada uno de dichos programas informáticos se almacena o se descarga preferiblemente en un medio o dispositivo de almacenamiento (por ejemplo, memoria o medios de estado sólido, o medios magnéticos u ópticos) legibles por una computadora programable de propósito general o especial, para configurar y operar la computadora cuando el sistema informático lee el medio o dispositivo de almacenamiento para realizar los procedimientos descritos en el presente documento. El sistema de la invención también puede implementarse como un medio de almacenamiento legible por computadora, configurado con (es decir, almacenando) un programa informático, donde el medio de almacenamiento así configurado hace que un sistema informático opere de una manera específica y predefinida para realizar las funciones descritas en el presente documento.

Se han descrito varias realizaciones de la invención. No obstante, se entenderá que se pueden realizar diversas modificaciones sin apartarse del alcance de la invención tal como se define en las reivindicaciones adjuntas. Son posibles numerosas modificaciones y variaciones de la presente invención a la luz de las enseñanzas anteriores. Debe entenderse que, dentro del alcance de las reivindicaciones adjuntas, la invención se puede poner en práctica de otra manera que la descrita específicamente en el presente documento. Cualquier número de referencia

contenido en las siguientes reivindicaciones tiene únicamente fines ilustrativos y no debe usarse para interpretar o limitar las reivindicaciones de ninguna manera.

REIVINDICACIONES

1.- Una unidad (210) de procesamiento de audio que comprende:

5 un búfer (201) configurada para almacenar al menos un bloque de una corriente de bits de audio codificada;

un desformateador (215) de carga útil de corriente de bits acoplado al búfer y configurado para demultiplexar al menos una porción de al menos dicho bloque de la corriente de bits de audio codificada; y

10 un subsistema (202) de descodificación acoplado al desformateador (215) de carga útil de corriente de bits y configurado para descodificar al menos una porción de al menos dicho bloque de la corriente de bits de audio codificada, en el que al menos dicho bloque de la corriente de bits de audio codificada incluye:

15 un elemento de relleno con un identificador que indica un inicio del elemento de relleno y datos de relleno después del identificador, en el que los datos de relleno incluyen al menos un indicador que identifica si se va a realizar una forma base de replicación de banda espectral o una forma potenciada de replicación de banda espectral en el contenido de audio de al menos un bloque de la corriente de bits de audio codificada, en el que la forma base de la replicación de banda espectral incluye parcheo espectral, la forma potenciada de replicación de banda espectral incluye transposición armónica, un valor del indicador indica que dicha forma potenciada de replicación de banda espectral debe realizarse en el contenido de audio, y otro valor del indicador indica que dicha forma base de la replicación de banda espectral, pero no dicha trasposición armónica, debe realizarse en el contenido de audio; y en el que

25 los datos de relleno incluyen además metadatos de replicación de banda espectral potenciada, y los metadatos de replicación de banda espectral potenciada no incluyen uno o más parámetros usados tanto para el parcheo espectral como para la transposición armónica,

en el que el uno o más parámetros usados tanto para el parcheo espectral como la transposición armónica están contenidos en una carga útil de extensión del elemento de relleno.

30 2.- La unidad de procesamiento de audio de la reivindicación 1, en la que uno o más parámetros usados tanto para el parcheo espectral como para la transposición armónica incluyen uno o más parámetros que definen una tabla de bandas de frecuencia maestra.

35 3.- La unidad de procesamiento de audio de la reivindicación 1, en la que uno o más parámetros usados tanto para el parcheo espectral como para la transposición armónica incluyen factores de escala de envolvente o factores de escala de piso de ruido.

40 4.- La unidad de procesamiento de audio de cualquiera de las reivindicaciones anteriores, en la que si al menos dicho indicador identifica la forma potenciada de procesamiento de replicación de banda espectral, un segundo indicador identifica si el sobremuestreo del dominio frecuencia adaptativo de la señal está habilitado o deshabilitado.

5.- Un método para descodificar una corriente de bits de audio codificada, comprendiendo el método:

45 recibir al menos un bloque de una corriente de bits de audio codificada;

demultiplexar al menos una porción de al menos dicho bloque de la corriente de bits de audio codificada; y

50 descodificar al menos una porción de al menos dicho bloque de la corriente de bits de audio codificada,

en el que al menos dicho bloque de la corriente de bits de audio codificada incluye:

55 un elemento de relleno con un identificador que indica un inicio del elemento de relleno y datos de relleno después del identificador, en el que

60 los datos de relleno incluyen al menos un indicador que identifica si se va a realizar una forma base de replicación de banda espectral o una forma potenciada de replicación de banda espectral en el contenido de audio de al menos un bloque de la corriente de bits de audio codificada, en el que la forma base de la replicación de banda espectral incluye parcheo espectral, la forma potenciada de replicación de banda espectral incluye transposición armónica, un valor del indicador indica que dicha forma potenciada de replicación de banda espectral debe realizarse en el contenido de audio, y otro valor del indicador indica que dicha forma base de la replicación de banda espectral, pero no dicha transposición armónica, debe realizarse en el contenido de audio; y en el que

65 los datos de relleno incluyen además metadatos de replicación de banda espectral potenciada, y los metadatos de replicación de banda espectral potenciada no incluyen uno o más parámetros usados tanto para el parcheo espectral como para la transposición armónica,

en el que el uno o más parámetros usados tanto para el parcheo espectral como la transposición armónica están contenidos en una carga útil de extensión del elemento de relleno.

- 5 6.- Un medio o dispositivo de almacenamiento en el que se almacena un programa informático legible por una computadora programable de propósito general o especial, para configurar y operar la computadora cuando la computadora lee el medio o dispositivo de almacenamiento para realizar el método de la reivindicación 5.

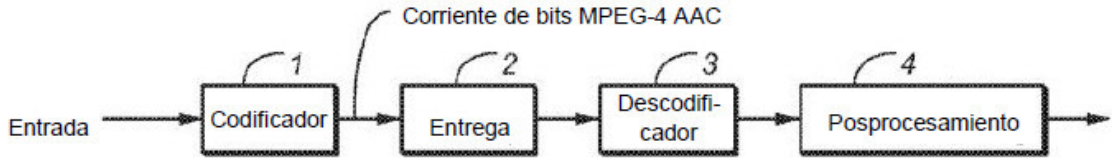


FIG. 1

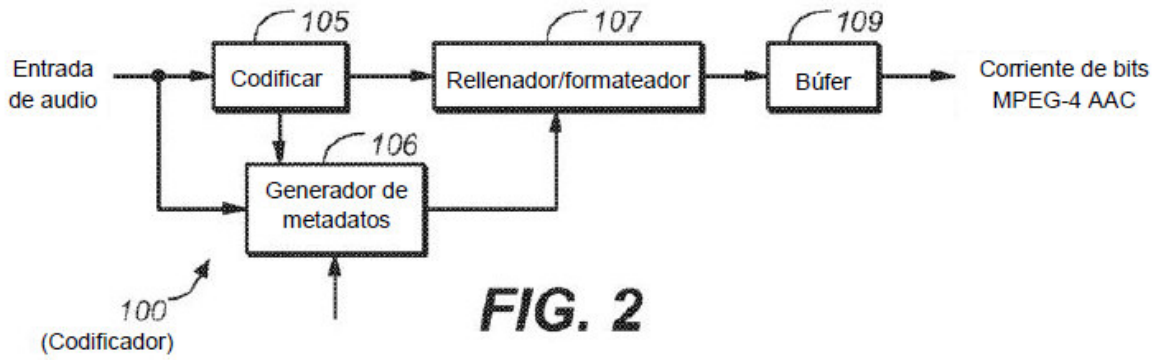


FIG. 2

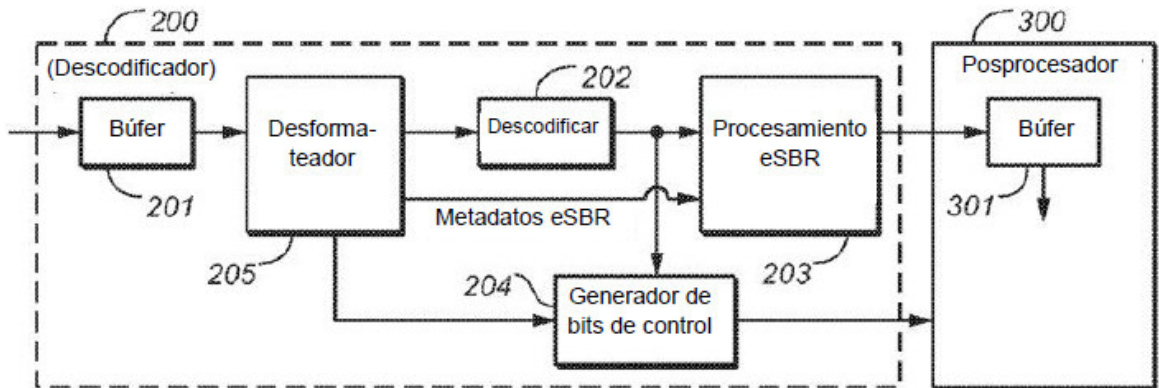


FIG. 3

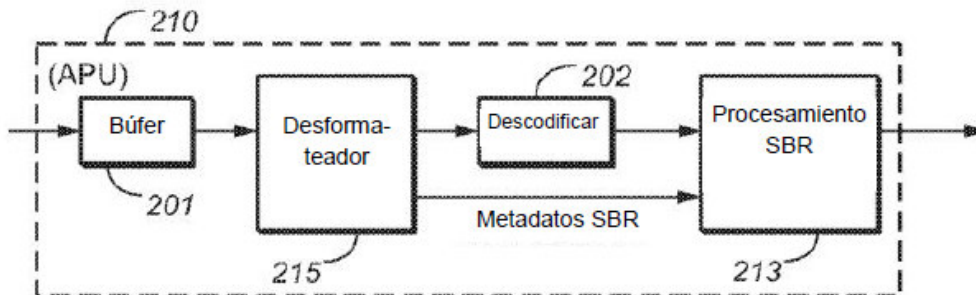


FIG. 4

