

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 933 476**

51 Int. Cl.:

G10L 19/16 (2013.01)

G10L 21/038 (2013.01)

G10L 19/24 (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 21195190 (0)**

97 Fecha y número de publicación de la concesión europea: **19.10.2022 EP 3958259**

54 Título: **Decodificación de secuencias de bits de audio con metadatos de replicación de banda espectral mejorada 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:

09.02.2023

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 2 933 476 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

Decodificación de secuencias de bits de audio con metadatos de replicación de banda espectral mejorada en al menos un elemento de relleno

Referencia cruzada con las solicitudes relacionadas

- 5 Esta solicitud es una solicitud divisional europea de la solicitud de patente europea EP 19213743.8 presentada el 5 de diciembre de 2019.

Campo técnico

10 La invención se refiere al procesamiento de señales de audio. Algunas realizaciones se refieren a la codificación y decodificación de secuencias de bits de audio (por ejemplo, secuencias de bits que tienen un formato MPEG-4 AAC) que incluyen metadatos para controlar la replicación de banda espectral mejorada (eSBR). Otras realizaciones se refieren a la decodificación de dichas secuencias de bits por decodificadores heredados que no están configurados para realizar un procesamiento eSBR y que ignoran dichos metadatos, o a la decodificación de una secuencia de bits de audio que no incluye dichos metadatos, incluso mediante la generación de datos de control de eSBR en respuesta a la secuencia de bits.

15 Antecedentes de la invención

Una secuencia 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 la generación de una secuencia de bits de audio codificada es el formato MPEG-4 *Advanced Audio Coding* (AAC), descrito en el estándar MPEG ISO/IEC 14496-3: 2009. En el estándar MPEG-4, AAC denota "codificación de audio avanzada" (*Advanced Audio Coding*) y HE-AAC denota "codificación de audio avanzada de alta eficiencia" (*High-Efficiency Advanced Audio Coding*).

20 El estándar MPEG-4 AAC define varios perfiles de audio, que determinan qué objetos y herramientas de codificación están presentes en un codificador o decodificador compatible. 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 MPEG-2 AAC de baja complejidad, con algunos ajustes, y no incluye ni el tipo de objeto de replicación de banda espectral ("SBR", *Spectral Band Replication*) ni el tipo de objeto estéreo paramétrico ("PS", *Parametric Stereo*). El perfil HE-AAC es un superconjunto del perfil AAC e incluye además el tipo de objeto SBR. El perfil HE-AAC v2 es un superconjunto del perfil HE-AAC e incluye además el tipo de objeto PS.

30 El tipo de objeto SBR contiene la herramienta de replicación de banda espectral, que es una herramienta de codificación importante que mejora significativamente la eficiencia de compresión de los códecs de audio perceptuales. SBR reconstruye los componentes de alta frecuencia de una señal de audio en el lado del receptor (por ejemplo, en el decodificador). De esta manera, el codificador solo necesita codificar y transmitir componentes de baja frecuencia, permitiendo una calidad de audio mucho mayor a bajas velocidades de datos. SBR se basa en la replicación de las secuencias de armónicos, truncadas previamente para reducir la velocidad de datos, a partir de la señal limitada por ancho de banda disponible y los datos de control obtenidos desde el codificador. La relación entre los componentes tonales y similares al ruido se mantiene mediante un filtrado inverso adaptativo, así como la adición opcional de ruido y sinusoidales. En el estándar MPEG-4 AAC, la herramienta SBR realiza una transposición espectral, en el que una serie de sub-bandas de filtro de espejo de cuadratura (QMF) contiguas se copian desde una parte de banda baja transmitida de una señal de audio a una parte de banda alta de la señal de audio, que es generada en el decodificador.

La transposición espectral puede no ser ideal para ciertos tipos de audio, tal como contenido musical con cruces de frecuencias relativamente bajos. Por lo tanto, se necesitan técnicas para mejorar la replicación de banda espectral.

45 Breve descripción de las realizaciones de la invención

La presente descripción proporciona unidades de procesamiento de audio, procedimientos para decodificar secuencias de bits de audio codificadas y medios o dispositivos de almacenamiento para realizar dichos procedimientos, tal como se reivindica en las reivindicaciones 1, 6 y 9. Las características opcionales se mencionan en las reivindicaciones dependientes.

Breve descripción de los dibujos

La Fig. 1 es un diagrama de bloques de una realización de un sistema que puede estar configurada para realizar una realización del procedimiento de la invención.

La Fig. 2 es un diagrama de bloques de un codificador que es una realización de la unidad de procesamiento de audio de la invención.

- 5 La Fig. 3 es un diagrama de bloques de un sistema que incluye un decodificador que es una realización de la unidad de procesamiento de audio de la invención, y opcionalmente también un post-procesador acoplado a la misma.

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

- 10 La Fig. 5 es un diagrama de bloques de un decodificador que es otra realización de la unidad de procesamiento de audio de la invención.

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

La Fig. 7 es un diagrama de un bloque de una secuencia de bits MPEG-4 AAC, que incluye segmentos en los que se divide.

Notación y nomenclatura

- 15 A lo largo de la presente descripción, incluyendo en las reivindicaciones, la expresión realización de una operación "sobre" 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 una operación directamente sobre la señal o los datos, o sobre una versión procesada de la señal o los datos (por ejemplo, sobre una versión de la señal que ha sido sometida a un filtrado preliminar o pre-procesamiento antes de la realización de la operación sobre la misma).

- 20 A lo largo de la presente descripción, incluyendo en 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, pero no se limitan a, codificadores (por ejemplo, transcodificadores), decodificadores, códecs, sistemas de pre-procesamiento, sistemas de post-procesamiento y sistemas de procesamiento de secuencias de bits (a veces denominados herramientas de procesamiento de secuencias de bits). Prácticamente todos los productos electrónicos de consumo, tales como teléfonos móviles, televisores, ordenadores portátiles y tabletas, contienen una unidad de procesamiento de audio.

- 25 A lo largo de la presente descripción, incluyendo en las reivindicaciones, el término "se acopla" o "acoplado" se usa en un sentido amplio para significar una conexión directa o indirecta. De esta manera, si un primer dispositivo se acopla a un segundo dispositivo, esa conexión puede ser a través de una conexión directa, o a través de una conexión indirecta a través de otros dispositivos y conexiones. Además, los componentes que están integrados en o con otros componentes están también acoplados entre sí.

Descripción detallada de las realizaciones de la invención

- 35 El estándar MPEG-4 AAC contempla que una secuencia de bits codificada con MPEG-4 AAC incluye metadatos indicativos de cada tipo de procesamiento SBR a aplicar (si se aplica alguno) por un decodificador para decodificar el contenido de audio de la secuencia de bits, y/o que controla dicho procesamiento SBR y/o es indicativa de al menos una característica o parámetro de al menos una herramienta SBR a emplear para decodificar el contenido de audio de la secuencia de bits. En la presente memoria, se usa la expresión "metadatos SBR" para denotar metadatos de este tipo que se describen o mencionan en el estándar MPEG-4 AAC.

- 40 El nivel superior de una secuencia 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 (al que se hace referencia en adelante como un "bloque") que contiene datos de audio (típicamente durante un tiempo período de 1024 o 960 muestras) e información relacionada y/u otros datos. En la presente memoria, se usa el término "bloque" para denotar un segmento de una secuencia de bits MPEG-4 AAC que comprende datos de audio (y metadatos correspondientes y opcionalmente también otros datos relacionados) que determina o es indicativa de un (pero no más de uno) elemento "raw_data_block".

- 45 Cada bloque de una secuencia de bits MPEG-4 AAC puede incluir una serie de elementos sintácticos (cada uno de los cuales se materializa también en la secuencia de bits como un segmento de datos). Siete tipos de dichos elementos sintácticos se definen en el estándar MPEG-4 AAC. Cada elemento sintáctico se identifica por un valor diferente del elemento de datos "id_syn_ele". Los ejemplos de elementos sintácticos incluyen un "single_channel_element()", un "channel_pair_element()" y un "fill_element()". Un elemento con un único canal es un

contenedor que incluye datos de audio de un único canal de audio (una señal de audio monofónica). Un elemento con un par de canales incluye datos de audio de dos canales de audio (es decir, una señal de audio estéreo).

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, al que se hace referencia como "datos de relleno".

5 Los elementos de relleno se han usado históricamente para ajustar la velocidad de bits instantánea de las secuencias de bits a transmitir a través de un canal de velocidad constante. Añadiendo la cantidad apropiada de datos de relleno a cada bloque, puede conseguirse una velocidad de datos constante.

Según las realizaciones de la invención, los datos de relleno pueden incluir una o más cargas útiles de extensión que extienden el tipo de datos (por ejemplo, metadatos) capaces de ser transmitidas en una secuencia de bits. Un decodificador que recibe secuencias de bits con datos de relleno que contienen un nuevo tipo de datos puede ser usado opcionalmente por un dispositivo que recibe la secuencia de bits (por ejemplo, un decodificador) para extender la funcionalidad del dispositivo. De esta manera, tal como puede apreciar una persona experta en la materia, los elementos de relleno son un tipo especial de estructura de datos y son diferentes de las estructuras de datos usadas típicamente para transmitir datos de audio (por ejemplo, cargas útiles de audio que contienen datos de canal).

De acuerdo con la invención, el identificador usado para identificar un elemento de relleno consiste en un entero sin signo de tres bits transmitiéndose primero el bit más significativo ("uimbsb") 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 secuencias de bits de audio es el estándar MPEG de codificación de voz y audio unificado (USAC) (ISO/IEC 23003-3: 2012). El estándar MPEG USAC describe la codificación y la decodificación de contenido de audio usando procesamiento de replicación de banda espectral (incluyendo el procesamiento SBR tal como se describe en el estándar MPEG-4 AAC, e incluye también otras formas mejoradas de procesamiento de replicación de banda espectral). Este procesamiento aplica herramientas de replicación de banda espectral (a las que a veces se hace referencia en la presente memoria como "herramientas SBR mejoradas" o "herramientas eSBR") de una versión expandida y mejorada del conjunto de herramientas SBR descritas en el estándar MPEG-4 AAC. De esta manera, eSBR (tal como se define en el estándar USAC) es una mejora de SBR (tal como se define en el estándar MPEG-4 AAC).

En la presente memoria, se usa la expresión "procesamiento SBR mejorado" (o "procesamiento eSBR") para hacer referencia al procesamiento de replicación de banda espectral usando al menos una herramienta eSBR (por ejemplo, al menos una herramienta eSBR que se describe o se menciona en el estándar MPEG USAC) que no se describe ni se menciona en el estándar MPEG-4 AAC. Los ejemplos de dichas herramientas eSBR son la transposición armónica, el pre-procesamiento adicional con transposición QMF o "pre-aplanamiento" y el conformado de envolvente temporal de muestra entre sub-bandas o "inter-TES".

Una secuencia de bits generada según el estándar MPEG USAC (a la que se hace referencia a veces en la presente memoria como "secuencia de bits USAC") incluye contenido de audio codificado e incluye típicamente metadatos indicativos de cada tipo de procesamiento de replicación de banda espectral a aplicar por un decodificador para decodificar contenido de audio de la secuencia de bits USAC y/o metadatos que controlan dicho procesamiento de replicación de banda espectral y/o son indicativos de al menos una característica o parámetro de al menos una herramienta SBR y/o una herramienta eSBR a ser empleada para decodificar el contenido de audio de la secuencia de bits USAC.

En la presente memoria, se usa la expresión "metadatos SBR mejorados" (o "metadatos eSBR") para hacer referencia a metadatos indicativos de cada tipo de procesamiento de replicación de banda espectral a ser aplicado por un decodificador para decodificar el contenido de audio de una secuencia de bits de audio codificada (por ejemplo, una secuencia 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 una herramienta eSBR para decodificar dicho contenido de audio, pero que no se describe o se 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 se mencionan en el estándar MPEG USAC, pero no en el estándar MPEG-4 AAC. De esta manera, los metadatos eSBR en la presente memoria hacen referencia a metadatos que no son metadatos de SBR, y los metadatos SBR en la presente memoria hacen referencia a metadatos que no son metadatos eSBR.

Una secuencia de bits USAC puede incluir tanto metadatos SBR como metadatos eSBR. Más específicamente, una secuencia 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. Según las realizaciones típicas de la presente invención, los metadatos eSBR (por ejemplo, datos de configuración

específicos de eSBR) se incluyen (según la presente invención) en una secuencia 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 secuencia de bits codificada usando un conjunto de herramientas eSBR (que comprende al menos una herramienta eSBR), por un decodificador regenera la banda de alta frecuencia de la señal de audio, en base a la replicación de secuencias de armónicos que se truncaron durante la codificación. Dicho procesamiento eSBR típicamente ajusta la envolvente espectral de la banda de alta frecuencia generada y aplica filtrado inverso, y añade ruido y componentes sinusoidales con el fin de volver a crear las características espectrales de la señal de audio original.

Según las realizaciones típicas de la invención, se incluyen metadatos eSBR (por ejemplo, se incluyen un pequeño número de bits de control que son metadatos eSBR) en uno o más segmentos de metadatos de una secuencia de bits de audio codificada (por ejemplo, una secuencia de bits MPEG-4 AAC) que incluye también datos de audio codificados en otros segmentos (segmentos de datos de audio). Típicamente, al menos uno de dichos segmentos de metadatos de cada bloque de la secuencia de bits es (o incluye) un elemento de relleno (que incluye un identificador que indica el inicio del elemento de relleno), y los metadatos eSBR se incluyen en el elemento de relleno después del identificador.

La Fig. 1 es un diagrama de bloques de una cadena de procesamiento de audio ejemplar (un sistema de procesamiento de datos de audio), en el que uno o más de los elementos del sistema pueden configurarse según una realización de la presente invención. El sistema incluye los siguientes elementos, acoplados entre sí, tal como se muestra: codificador 1, subsistema 2 de suministro, decodificador 3 y unidad 4 de post-procesamiento. En las variantes del sistema mostrado, se omiten uno o más de los elementos o se incluyen unidades de procesamiento de datos de audio adicionales.

En algunas implementaciones, el codificador 1 (que opcionalmente incluye una unidad de pre-procesamiento) está configurada para aceptar muestras PCM (dominio del tiempo) que comprenden contenido de audio como entrada, y para emitir una secuencia de bits de audio codificadas (que tienen un formato que es compatible con el estándar MPEG-4 AAC) que es indicativa del contenido de audio. Los datos de la secuencia de bits que son indicativos del contenido de audio a veces se denominan en la presente memoria "datos de audio" o "datos de audio codificados". Si el codificador está configurado según una realización típica de la presente invención, la secuencia de bits de audio emitida desde el codificador incluye metadatos eSBR (y típicamente también otros metadatos) así como datos de audio.

Una o más secuencias de bits de audio codificadas emitidas desde el codificador 1 pueden validarse al subsistema 2 de suministro de audio codificado. El subsistema 2 está configurado para almacenar y/o suministrar cada salida de secuencia de bits codificada desde el codificador 1. Una secuencia de bits de audio codificada emitida desde el codificador 1 puede ser almacenada por el subsistema 2 (por ejemplo, en la forma de un disco DVD o Blu ray), o transmitida por el subsistema 2 (que puede implementar un enlace o red de transmisión), o puede ser almacenada y transmitida por el subsistema 2.

El decodificador 3 está configurado para decodificar una secuencia de bits de audio con codificación MPEG-4 AAC (generada por el codificador 1) que recibe a través del subsistema 2. En algunas realizaciones, el decodificador 3 está configurado para extraer metadatos eSBR desde cada bloque de la secuencia de bits y para decodificar la secuencia de bits (incluyendo mediante el procesamiento eSBR usando los metadatos eSBR extraídos) para generar datos de audio decodificados (por ejemplo, secuencias de muestras de audio PCM decodificadas). En algunas realizaciones, el decodificador 3 está configurado para extraer metadatos SBR desde la secuencia de bits (pero para ignorar los metadatos eSBR incluidos en la secuencia de bits), y para decodificar la secuencia de bits (que incluye realizando el procesamiento SBR usando los metadatos SBR extraídos) para generar datos de audio decodificados (por ejemplo, secuencias de muestras de audio PCM decodificadas). Típicamente, el decodificador 3 incluye una memoria intermedia que almacena (por ejemplo, de manera no transitoria) segmentos de la secuencia de bits de audio codificada recibida desde el subsistema 2.

La unidad 4 de post-procesamiento de la Fig. 1 está configurada para aceptar una secuencia de datos de audio decodificados desde el decodificador 3 (por ejemplo, muestras de audio PCM decodificadas), y para realizar un post-procesamiento sobre la misma. La unidad de post-procesamiento puede configurarse también para recuperar el contenido de audio post-procesado (o el audio decodificado recibido desde el decodificador 3) para su reproducción por uno o más altavoces.

La Fig. 2 es un diagrama de bloques de un codificador (100) que es una realización de la unidad de procesamiento de audio de la invención. Cualquiera de los componentes o elementos del codificador 100 puede implementarse como uno o más procesos y/o uno o más circuitos (por ejemplo, ASICs, FPGAs 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/de formateo, la etapa 106 de generación de metadatos y la memoria 109 intermedia,

conectados tal como se muestra. Típicamente también, el codificador 100 incluye otros elementos de procesamiento (no mostrados). El codificador 100 está configurado para convertir una secuencia de bits de audio de entrada en una secuencia de bits MPEG-4 AAC de salida codificada.

5 El generador 106 de metadatos está acoplado y configurado para generar (y/o pasar a la etapa 107) metadatos (incluyendo metadatos eSBR y metadatos SBR) a ser incluidos por la etapa 107 en la secuencia de bits codificada para ser emitida desde el codificador 100.

El codificador 105 está acoplado y configurado para codificar los datos de audio de entrada (por ejemplo, realizando una compresión sobre los mismos), y para validar el audio codificado resultante a la etapa 107 para su inclusión en la secuencia de bits codificada a ser emitida desde la etapa 107.

10 La etapa 107 está configurada para multiplexar el audio codificado desde el codificador 105 y los metadatos (incluyendo metadatos eSBR y metadatos SBR) desde el generador 106 para generar la secuencia de bits codificada a ser emitida desde la etapa 107, preferiblemente de manera que la secuencia de bits codificada tenga el formato especificado por una de las realizaciones de la presente invención.

15 La memoria 109 intermedia está configurada para almacenar (por ejemplo, de manera no transitoria) al menos un bloque de la secuencia de bits de audio codificada emitida desde la etapa 107, y una secuencia de los bloques de la secuencia de bits de audio codificada se valida a continuación desde la memoria 109 intermedia tal como es emitida desde el codificador 100 a un sistema de suministro.

20 La Fig. 3 es un diagrama de bloques de un sistema que incluye un decodificador (200) que es una realización de la unidad de procesamiento de audio de la invención, y opcionalmente también un post-procesador (300) acoplado a la misma. Cualquiera de los componentes o elementos del decodificador 200 y el post-procesador 300 puede implementarse como uno o más procesos y/o uno o más circuitos (por ejemplo, ASICs, FPGAs u otros circuitos integrados), en hardware, software o una combinación de hardware y software. El decodificador 200 comprende memoria 201 intermedia, deformateador (analizador sintáctico) 205 de carga útil de secuencia de bits, subsistema 202 de decodificación de audio (al que se hace referencia a veces como etapa de decodificación "central" o subsistema de decodificación "central"), etapa 203 de procesamiento eSBR y etapa 204 de generación de bits de control, conectados tal como se muestra. Típicamente también, el decodificador 200 incluye otros elementos de procesamiento (no mostrados).

30 La memoria 201 (memoria intermedia) almacena (por ejemplo, de manera no transitoria) al menos un bloque de una secuencia de bits de audio codificada MPEG-4 AAC recibida por el decodificador 200. En la operación del decodificador 200, una secuencia de los bloques de la secuencia de bits es validada desde la memoria 201 intermedia al deformateador 205.

35 En las variantes de la realización de la Fig. 3 (o la realización de la Fig. 4 que se describirá a continuación), una APU que no es un decodificador (por ejemplo, APU 500 de la Fig. 6) incluye una memoria intermedia (por ejemplo, una memoria intermedia idéntica a la memoria 201 intermedia) que almacena (por ejemplo, de manera no transitoria) al menos un bloque de una secuencia de bits de audio codificada (por ejemplo, una secuencia de bits de audio MPEG-4 AAC) del mismo tipo recibido por la memoria 201 intermedia de la Fig. 3 o la Fig. 4 (es decir, una secuencia de bits de audio codificada que incluye metadatos eSBR).

40 Con referencia una vez más a la Fig. 3, el deformateador 205 está acoplado y configurado para demultiplexar cada bloque de la secuencia de bits para extraer metadatos SBR (incluyendo datos de envoltorio cuantificados) y metadatos eSBR (y típicamente también otros metadatos) a partir de los mismos, para validar al menos los metadatos eSBR y los metadatos SBR a la etapa 203 de procesamiento eSBR, y típicamente también para validar otros metadatos extraídos al subsistema 202 de decodificación (y opcionalmente también para controlar el generador 204 de bits). El deformateador 205 está acoplado y configurado también para extraer datos de audio desde cada bloque de la secuencia de bits, y para validar los datos de audio extraídos al subsistema 202 de decodificación (etapa de decodificación).

50 El sistema de la Fig. 3 incluye también opcionalmente un post-procesador 300. El post-procesador 300 incluye la memoria 301 intermedia (memoria intermedia) y otros elementos de procesamiento (no mostrados) que incluyen al menos un elemento de procesamiento acoplado a la memoria 301 intermedia. La memoria 301 intermedia almacena (por ejemplo, de manera no transitoria) al menos un bloque (o trama) de los datos de audio decodificados recibidos por el post-procesador 300 desde el decodificador 200. Los elementos de procesamiento del post-procesador 300 están acoplados y configurados para recibir y procesar de manera adaptativa una secuencia de los bloques (o tramas) del audio decodificado emitido desde la memoria 301 intermedia, usando metadatos emitidos desde el subsistema 202 de decodificación (y/o el deformateador 205) y/o los bits de control emitidos desde la etapa 204 del decodificador 200.

El subsistema 202 de decodificación de audio del decodificador 200 está configurado para decodificar los datos de audio extraídos por el analizador 205 sintáctico (puede hacerse referencia dicha decodificación como una operación de decodificación "central") para generar datos de audio decodificados y para validar los datos de audio decodificados a la etapa 203 de procesamiento eSBR. La decodificación se realiza en el dominio de la frecuencia y típicamente incluye cuantización inversa seguida por procesamiento espectral. Típicamente, una etapa de procesamiento final en el subsistema 202 aplica una transformación de dominio de la frecuencia a dominio del tiempo a los datos de audio de dominio de la frecuencia decodificados, de manera que la salida del subsistema sea datos de audio decodificados en el dominio del tiempo. La etapa 203 está configurada para aplicar herramientas SBR y herramientas eSBR indicadas por los metadatos eSBR y los eSBR (extraídos por el analizador 205 sintáctico) a los datos de audio decodificados (es decir, para realizar el procesamiento SBR y eSBR sobre la salida del subsistema 202 de decodificación usando los metadatos SBR y eSBR) para generar los datos de audio totalmente decodificados emitidos (por ejemplo, al post-procesador 300) desde el decodificador 200. Típicamente, el decodificador 200 incluye una memoria (accesible por el subsistema 202 y la etapa 203) que almacena los datos de audio deformateados y los metadatos emitidos desde el deformateador 205, y la etapa 203 está configurada para acceder a los datos de audio y los metadatos (incluyendo los metadatos SBR y eSBR) según sea necesario durante el procesamiento SBR y eSBR. El procesamiento SBR y el procesamiento eSBR en la etapa 203 puede considerarse que es el post-procesamiento sobre la salida del subsistema 202 de decodificación central. Opcionalmente, el decodificador 200 incluye también un subsistema de mezcla final (que puede aplicar herramientas estéreo paramétricas ("PS") definidas en el estándar MPEG-4 AAC, usando los metadatos PS extraídos por el deformateador 205 y/o bits de control generados en el subsistema 204) que está acoplado y configurado para realizar una mezcla en la salida de la etapa 203 para generar audio mezclado totalmente descodificado, que se emite desde el decodificador 200. De manera alternativa, el post-procesador 300 está configurado para realizar una mezcla sobre la salida del decodificador 200 (por ejemplo, usando los metadatos PS extraídos por el deformateador 205 y/o los bits de control generados en el subsistema 204).

En respuesta a los metadatos extraídos por el deformateador 205, el generador 204 de bits de control puede generar datos de control, y los datos de control pueden usarse dentro del decodificador 200 (por ejemplo, en un subsistema de mezcla final) y/o pueden validarse como salida del decodificador 200 (por ejemplo, al post-procesador 300 para su uso en el post-procesamiento). En respuesta a los metadatos extraídos desde la secuencia de bits de entrada (y opcionalmente también en respuesta a los datos de control), la etapa 204 puede generar (y validar al post-procesador 300) bits de control que indican que los datos de audio decodificados emitidos desde la etapa 203 de procesamiento eSBR debería someterse a tipo de post-procesamiento específico. En algunas implementaciones, el decodificador 200 está configurado para validar los metadatos extraídos por el deformateador 205 desde la secuencia de bits de entrada al post-procesador 300, y el post-procesador 300 está configurado para realizar el post-procesamiento sobre los datos de audio decodificados emitidos desde el decodificador 200 usando los metadatos.

La Fig. 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 decodificador heredado que no está configurado para realizar el procesamiento eSBR. Cualquiera de los componentes o elementos de la APU 210 puede implementarse como uno o más procesos y/o uno o más circuitos (por ejemplo, ASICs, FPGAs u otros circuitos integrados), en hardware, software o una combinación de hardware y software. La APU 210 comprende una memoria 201 intermedia, un deformateador (analizador sintáctico) 215 de carga útil de secuencia de bits, un subsistema 202 de decodificación de audio (al que se hace referencia a veces como etapa de decodificación "central" o subsistema de decodificación "central") y una etapa 213 de procesamiento SBR, conectados tal 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 decodificador 200 (de la Fig. 3) y no se repetirá la descripción anterior de los mismos. Durante la operación de la APU 210, una secuencia de bloques de una secuencia de bits de audio codificada (una secuencia de bits MPEG-4 AAC) recibida por el APU 210 se valida desde la memoria 201 intermedia al deformateador 215.

El deformateador 215 está acoplado y configurado para demultiplexar cada bloque de la secuencia de bits para extraer los metadatos SBR (incluyendo los datos de envolvente cuantificados) y, típicamente, también otros metadatos a partir de los mismos, pero para ignorar los metadatos eSBR que puedan estar incluidos en la secuencia de bits según cualquier realización de la presente invención. El deformateador 215 está configurado para validar al menos los metadatos SBR a la etapa 213 de procesamiento SBR. El deformateador 215 está también acoplado y configurado para extraer los datos de audio de cada bloque de la secuencia de bits, y para validar los datos de audio extraídos al subsistema 202 de decodificación (etapa de decodificación).

El subsistema 202 de decodificación de audio del decodificador 200 está configurado para decodificar los datos de audio extraídos por el deformateador 215 (puede hacerse referencia a dicha decodificación como una operación de decodificación "central") para generar los datos de audio decodificados y para validar los datos de audio decodificados a la etapa 213 de procesamiento SBR. La decodificación se realiza en el dominio de la frecuencia.

Típicamente, una etapa de procesamiento final en el subsistema 202 aplica una transformación de dominio de la frecuencia a dominio del tiempo a los datos de audio de dominio de la frecuencia decodificados, de manera que la salida del subsistema sean datos de audio decodificados en el dominio del tiempo. La etapa 213 está configurada para aplicar las herramientas SBR (pero no herramientas eSBR) indicadas por los metadatos SBR (extraídos por el deformateador 215) a los datos de audio decodificados (es decir, para realizar un procesamiento SBR sobre la salida del subsistema 202 de decodificación usando los metadatos SBR) para generar los datos de audio totalmente decodificados que se emiten (por ejemplo, al post-procesador 300) desde la APU 210. Típicamente, la APU 210 incluye una memoria (accesible por el subsistema 202 y la etapa 213) que almacena los datos de audio deformateados y los metadatos emitidos desde el deformateador 215, y la etapa 213 está configurada para acceder a los datos de audio y los metadatos (incluyendo los metadatos SBR) según sea necesario durante el procesamiento SBR. El procesamiento SBR en la etapa 213 puede considerarse como un post-procesamiento sobre la salida del subsistema 202 de decodificación central. Opcionalmente, la APU 210 incluye también un subsistema de mezcla final (que puede aplicar las herramientas estéreo paramétricas ("PS") definidas en el Estándar MPEG-4 AAC, usando los metadatos PS extraídos por el deformateador 215) que está acoplado y configurado para realizar una mezcla en la salida de la etapa 213 para generar audio mezclado totalmente descodificado que se emite desde la APU 210. De manera alternativa, un post-procesador está configurado para realizar una mezcla en la salida de la APU 210 (por ejemplo, usando los metadatos PS extraídos por el deformateador 215 y/o los bits de control generados en la APU 210).

Diversas implementaciones del codificador 100, el decodificador 200 y la APU 210 están configuradas para realizar diferentes realizaciones del procedimiento de la invención.

Según algunas realizaciones, los metadatos eSBR se incluyen (por ejemplo, se incluyen un pequeño número de bits de control que son metadatos eSBR) en una secuencia de bits de audio codificada (por ejemplo, una secuencia de bits MPEG-4 AAC), de manera que los decodificadores heredados (que no están configurados para analizar sintácticamente los metadatos eSBR, o para usar cualquier herramienta eSBR a la que pertenecen los metadatos eSBR) puedan ignorar los metadatos eSBR pero, sin embargo, decodifiquen la secuencia de bits en la medida de lo posible sin el uso de los metadatos eSBR ni ninguna herramienta eSBR a la que pertenecen los metadatos eSBR, típicamente sin ninguna penalización significativa en la calidad de audio decodificada. Sin embargo, los decodificadores eSBR configurados para analizar sintácticamente la secuencia 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 del uso de al menos una de dichas herramientas eSBR. Por lo tanto, las realizaciones de la invención proporcionan un medio para transmitir eficientemente datos o metadatos de control de replicación de banda espectral mejorada (eSBR) de una manera compatible con versiones anteriores.

Típicamente, los metadatos eSBR en la secuencia 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 haber sido aplicadas o no por un codificador durante la generación de la secuencia de bits):

- Transposición armónica;
- Pre-procesamiento de transposición QMF adicional (pre-aplanamiento); y
- Conformado de envolvente temporal de muestra entre sub-bandas o "inter-TES".

Por ejemplo, los metadatos eSBR incluidos en la secuencia de bits pueden ser indicativos de los valores de los parámetros (descritos en el estándar MPEG USAC y en la presente descripción): `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`.

En la presente memoria, la notación `X[ch]`, donde `X` es algún parámetro, indica que el parámetro pertenece al canal ("`ch`") de contenido de audio de una secuencia de bits codificada a ser decodificada. En aras de la simplicidad, a veces se omite la expresión `[ch]` y se supone que el parámetro relevante pertenece a un canal de contenido de audio.

En la presente memoria, la notación `X[ch][env]`, donde `X` es algún parámetro, indica que el parámetro pertenece a la envolvente SBR ("`env`") del canal ("`ch`") del contenido de audio de una secuencia de bits codificada a ser decodificada. En aras de la simplicidad, a veces se omiten las expresiones `[env]` y `[ch]`, y se supone que el parámetro relevante pertenece a una envolvente SBR de un canal de contenido de audio.

Tal como se ha indicado, el estándar MPEG USAC contempla que una secuencia de bits USAC incluya metadatos eSBR que controlan el rendimiento del procesamiento eSBR por un decodificador. Los metadatos eSBR incluyen los siguientes parámetros de metadatos de un bit: `harmonicSBR`; `bs_interTES`; y `bs_pvc`.

- El parámetro "harmonicSBR" indica el uso de transposición armónica (transposición armónica) para SBR. Específicamente, harmonicSBR = 0 indica ausencia de transposición espectral armónica tal como se describe en la Sección 4.6.18.6.3 del estándar MPEG-4 AAC; y harmonicSBR = 1 indica presencia de transposición armónica SBR (del tipo usado en eSBR, tal como se describe en la Sección 7.5.3 o 7.5.4 del estándar MPEG USAC). La transposición armónica SBR no se usa según la replicación de banda espectral no eSBR (es decir, SBR que no es eSBR). A lo largo de la presente descripción, se hace referencia a la transposición espectral como una forma básica de replicación de banda espectral, mientras que se hace referencia a la transposición armónica como una forma mejorada de replicación de banda espectral.
- 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.
- Durante la decodificación de una secuencia de bits codificada, el rendimiento de la transposición armónica durante una etapa de procesamiento eSBR de la decodificación (para cada canal, "ch", del contenido de audio indicado por la secuencia de bits) es controlado mediante los siguientes parámetros de metadatos eSBR: sbrPatchingMode[ch]; sbrOversamplingFlag[ch]; sbrPitchInBinsFlag[ch]; y sbrPitchInBins[ch].
- El valor "sbrPatchingMode[ch]" indica el tipo de dispositivo de transposición usado en eSBR: sbrPatchingMode[ch] = 1 indica una transposición no armónica tal como se describe en la Sección 4.6.18.6.3 del estándar MPEG-4 AAC; sbrPatchingMode[ch] = 0 indica una transposición armónica SBR tal como se describe en la Sección 7.5.3 o 7.5.4 del estándar MPEG USAC.
- El valor "sbrOversamplingFlag[ch]" indica el uso de sobremuestreo adaptativo de señal en el dominio de la frecuencia en eSBR en combinación con la transposición SBR armónica basada en DFT tal como se describe en la Sección 7.5.3 del estándar MPEG USAC. Este indicador controla el tamaño de las DFTs que se utilizan en el dispositivo de transposición: 1 indica sobremuestreo adaptativo de señal en el dominio de la frecuencia habilitado tal como se describe en la Sección 7.5.3.1 del estándar MPEG USAC; 0 indica sobremuestreo adaptativo de señal en el dominio de la frecuencia deshabilitado tal 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] está establecido a cero.
- El valor "sbrPitchInBins[ch]" controla la adición de términos de productos cruzados en el dispositivo de transposición armónico SBR. El valor sbrPitchInBins[ch] es un valor entero en el intervalo [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.
- En el caso en el que una secuencia de bits MPEG-4 AAC es indicativa de un par de canales SBR, cuyos canales no estén acoplados (en lugar de un único canal SBR), la secuencia de bits es indicativa de dos instancias de la sintaxis anterior (para transposición armónica o no armónica), una para cada canal de sbr_channel_pair_element().
- La transposición armónica de la herramienta eSBR típicamente mejora la calidad de las señales musicales decodificadas a frecuencias de cruce relativamente bajas. La transposición no armónica (es decir, la transposición espectral heredada) típicamente mejora las señales de voz. Por lo tanto, un punto de partida en la decisión sobre qué tipo de transposición es preferible para codificar un contenido de audio específico es seleccionar el procedimiento de transposición dependiendo de la detección de voz/música empleándose la transposición armónica en el contenido musical y la transposición espectral en el contenido de velocidad.
- El rendimiento del pre-aplanamiento durante el procesamiento eSBR es controlado por el valor de un parámetro de metadatos eSBR de un bit conocido como "bs_sbr_preprocessing", en el sentido de que el pre-aplanamiento se realiza o no dependiendo del valor de este único bit. Cuando se usa el algoritmo de transposición SBR QMF, tal como se describe en la Sección 4.6.18.6.3 del estándar MPEG-4 AAC, puede realizarse la etapa de pre-aplanamiento (cuando lo indica el parámetro "bs_sbr_preprocessing") en un intento de evitar discontinuidades en la forma de la envolvente espectral de una señal de alta frecuencia introducida a un ajustador de envolvente posterior (el ajustador de envolvente realiza otra etapa del procesamiento eSBR). Típicamente, el pre-aplanamiento mejora el funcionamiento de la etapa de ajuste de envolvente posterior, resultando en una señal de banda alta que se percibe como más estable.
- El rendimiento de la conformación de envolvente temporal de muestra entre sub-bandas (la herramienta "inter-TES") durante el procesamiento eSBR en un decodificador es controlado mediante los siguientes parámetros de metadatos eSBR para cada envolvente de SBR ("env") de cada canal ("ch") del contenido de audio de una secuencia de bits USAC que está siendo decodificada: bs_temp_shape[ch][env] y bs_inter_temp_shape_mode[ch][env].

La herramienta inter-TES procesa las muestras de sub-banda QMF después del ajustador de envolvente. Esta etapa 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 sub-banda QMF en una envolvente SBR, inter-TES conforma la envolvente temporal entre las muestras de sub-banda QMF.

- 5 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 (tal como se define en el estándar MPEG USAC) los valores del parámetro y en inter-TES.

Se espera que el requisito de velocidad de bits global para incluir en una secuencia de bits MPEG-4 AAC metadatos eSBR indicativos de las herramientas eSBR indicadas anteriormente (transposición armónica, pre-aplanamiento e inter_TES) sea del orden de unos pocos cientos de bits por segundo, ya que solo se transmiten los datos de control diferencial necesarios para realizar el procesamiento eSBR según algunas realizaciones de la invención. Los decodificadores heredados pueden ignorar esta información, ya que se incluye de manera compatible con versiones anteriores (tal como se explicará más adelante). Por lo tanto, el efecto perjudicial sobre la velocidad de bits asociado con la inclusión de los metadatos eSBR es insignificante, por una serie de razones, incluyendo las siguientes:

- 10
- 15 • La penalización en la velocidad de bits (debida a la inclusión de los metadatos eSBR) es una fracción muy pequeña de la velocidad de bits total, ya que solo se transmiten los datos de control diferencial necesarios para realizar el procesamiento eSBR (y no una multidifusión de los datos de control SBR);
 - El ajuste de la información de control relacionada con SBR no depende típicamente de los detalles de la transposición; y
 - 20 • La herramienta inter-TES (empleada durante el procesamiento eSBR) realiza un post-procesamiento de un solo extremo de la señal transpuesta.

De esta manera, las realizaciones de la invención proporcionan un medio para transmitir eficientemente datos o metadatos de control de replicación de banda espectral mejorada (eSBR) de una manera compatible con versiones anteriores. Esta transmisión eficiente de los datos de control eSBR reduce los requisitos de memoria en los decodificadores, codificadores y transcodificadores que emplean los aspectos de la invención, sin que tengan efectos adversos tangibles en la velocidad de bits. Además, los requisitos de complejidad y de procesamiento asociados con la realización eSBR según las realizaciones de la invención se reducen también, ya que los datos SBR solo deben procesarse una vez y no deben transmitirse mediante multidifusión, siendo este el caso si eSBR se tratara como un tipo de objeto completamente separado en MPEG-4 AAC en lugar de estar integrado en el códec MPEG-4 AAC de una manera compatible con versiones anteriores.

A continuación, con referencia a la Fig. 7, se describen elementos de un bloque ("raw_data_block") de una secuencia de bits MPEG-4 AAC en la que se incluyen metadatos eSBR según algunas realizaciones de la presente invención. La Fig. 7 es un diagrama de un bloque (un "raw_data_block") de la secuencia de bits MPEG-4 AAC, que muestra algunos de los segmentos del mismo.

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

- 45 Un fill_element (al que se hace referencia en la presente memoria como un elemento de relleno) de una secuencia de bits MPEG-4 AAC incluye un identificador ("ID2" de la Fig. 7) que indica el inicio de un elemento de relleno y datos de relleno después del identificador. El identificador ID2 consiste en un entero sin signo de tres bits transmitiéndose primero el bit más significativo ("uimbsb") que tiene un valor de 0x6. Los datos de relleno pueden incluir un elemento extension_payload() (al que se hace referencia a veces como 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
- 50 se identifican por el parámetro "extension_type", que es un entero sin signo de cuatro bits transmitiéndose primero el bit más significativo ("uimbsb").

Los datos de relleno (por ejemplo, una carga útil de extensión de los mismos) pueden incluir una cabecera o un identificador (por ejemplo, "cabecera1" de la Fig. 7) que indica un segmento de datos de relleno que es indicativo de un objeto SBR (es decir, la cabecera inicializa un tipo de "objeto SBR", al que se hace referencia como

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 '1101' o '1110' para el campo extension_type en la cabecera, con el identificador '1101' que identifica una carga útil de extensión con datos SBR e identificando '1110' una extensión de carga útil con datos SBR con una verificación de redundancia cíclica (Cyclic Redundancy Check, CRC) para verificar la exactitud de los datos SBR.

Cuando la cabecera (por ejemplo, el campo extension_type) inicializa un tipo de objeto SBR, los metadatos SBR (a los que se hace referencia a veces en la presente memoria como "datos de replicación de banda espectral" y a los que se hace referencia como sbr_data() en el estándar MPEG-4 AAC) siguen a la cabecera, y al menos un elemento de extensión de replicación de banda espectral (por ejemplo, el "elemento de extensión SBR" del elemento de relleno 1 de la Fig. 7) puede seguir a los metadatos de SBR. A dicho elemento de extensión de replicación de banda espectral (un segmento de la secuencia de bits) se hace referencia como un contenedor "sbr_extension()" en el estándar MPEG-4 AAC. Un elemento de extensión de replicación de banda espectral incluye opcionalmente una cabecera (por ejemplo, "cabecera de extensión SBR" del elemento de relleno 1 de la Fig. 7).

El estándar MPEG-4 AAC contempla que un elemento de extensión de replicación de banda espectral puede incluir datos PS (estéreo paramétricos) para los datos de audio de un programa. El estándar MPEG-4 AAC contempla que cuando la cabecera de un elemento de relleno (por ejemplo, de una carga útil de extensión del mismo) inicializa un tipo de objeto SBR (tal como lo hace la "cabecera1" de la Fig. 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.

Según algunas realizaciones de la presente invención, los metadatos eSBR (por ejemplo, un indicador que indica si debe realizarse o no un procesamiento de replicación de banda espectral mejorada (eSBR) sobre 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, dicho indicador se indica en el elemento de relleno 1 de la Fig. 7, donde el indicador se encuentra después de la cabecera (la "cabecera de extensión SBR" del elemento de relleno 1) del "elemento de extensión SBR" del elemento de relleno 1. Opcionalmente, dicho indicador y los metadatos eSBR adicionales se incluyen en un elemento de extensión de replicación de banda espectral después de la cabecera del elemento de extensión de replicación de banda espectral (por ejemplo, en el elemento de extensión SBR del elemento de relleno 1 en la Fig. 7, después de la cabecera de extensión SBR). Según algunas realizaciones de la presente invención, un elemento de relleno que incluye metadatos eSBR incluye también 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 debe realizarse un procesamiento eSBR sobre el contenido de audio del bloque relevante.

Según algunas realizaciones de la invención, los metadatos eSBR se incluyen en un elemento de relleno (por ejemplo, el elemento de relleno 2 de la Fig. 7) de una secuencia de bits MPEG-4 AAC distinto de un elemento de extensión de replicación de banda espectral (elemento de extensión SBR) del elemento de relleno. Esto es debido a que los elementos de relleno que contienen una extension_payload() con datos SBR o datos SBR con un 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 su propia carga útil de extensión, se usa un elemento de relleno separado para almacenar los metadatos eSBR. Dicho elemento de relleno incluye un identificador (por ejemplo, "ID2" de la Fig. 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() (al que se hace referencia a veces en la presente memoria como 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 del mismo) incluyen una cabecera (por ejemplo, "cabecera2" del elemento de relleno 2 de la Fig. 7) que es indicativa de un objeto eSBR (es decir, la cabecera inicializa un tipo de objeto de replicación de banda espectral mejorada (eSBR)), y los datos de relleno (por ejemplo, una carga útil de extensión del mismo) incluyen metadatos eSBR después de la cabecera. Por ejemplo, el elemento de relleno 2 de la Fig. 7 incluye dicha cabecera ("cabecera2") e incluye también, después de la cabecera, metadatos eSBR (es decir, el "indicador" en el elemento de relleno 2, que es indicativo de si debe realizarse o no un procesamiento de replicación de banda espectral mejorada (eSBR) sobre el contenido de audio del bloque). Opcionalmente, se incluyen también metadatos eSBR adicionales en los datos de relleno del elemento de relleno 2 de la Fig. 7, después de la cabecera2. En las realizaciones que se describen en el presente párrafo, la cabecera (por ejemplo, la cabecera 2 de la Fig. 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 manera que el campo extension_type de la cabecera indique que los datos de relleno incluyen metadatos eSBR).

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

una memoria (por ejemplo, la memoria 201 intermedia de la Fig. 3 o 4) configurada para almacenar al menos un bloque de una secuencia de bits de audio codificada (por ejemplo, al menos un bloque de una secuencia de bits MPEG-4 AAC);

5 un deformateador de carga útil de secuencia de bits (por ejemplo, el elemento 205 de la Fig. 3 o el elemento 215 de la Fig. 4) acoplado a la memoria y configurado para demultiplexar al menos una parte de dicho bloque de la secuencia de bits; y

un subsistema de decodificación (por ejemplo, los elementos 202 y 203 de la Fig. 3, o los elementos 202 y 213 de la Fig. 4), acoplado y configurado para decodificar al menos una parte del contenido de audio de dicho bloque de la secuencia de bits, en el que el bloque incluye:

10 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 un valor de 0x6, de la Tabla 4.85 del estándar MPEG-4 AAC), y datos de relleno después del identificador, en el que los datos de relleno incluyen:

15 al menos un indicador que identifica si debe realizarse o no el procesamiento de replicación de banda espectral mejorada (eSBR) sobre el contenido de audio del bloque (por ejemplo, usando los datos de replicación de banda espectral y los metadatos eSBR incluidos en el bloque).

20 El indicador es metadatos eSBR, y un ejemplo del indicador es el indicador sbrPatchingMode. Otro ejemplo del indicador es el indicador harmonicSBR. Ambos indicadores indican si debe realizarse o no una forma básica de replicación de banda espectral o una forma mejorada de replicación espectral sobre los datos de audio del bloque. La forma básica de la replicación espectral es la transposición espectral, y la forma mejorada de la replicación de banda espectral es la transposición armónica.

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

25 La memoria puede ser una memoria intermedia (por ejemplo, una implementación de la memoria 201 intermedia de la Fig. 4) que almacena (por ejemplo, de manera no transitoria) el al menos un bloque de la secuencia de bits de audio codificada.

Se estima que la complejidad del rendimiento del procesamiento eSBR (usando las herramientas de transposición armónica eSBR, pre-aplanamiento e inter_TES) por un decodificador eSBR durante la decodificación de una secuencia de bits MPEG-4 AAC que incluye metadatos eSBR (indicativos de estas herramientas eSBR) sería como sigue (para la decodificación típica con los parámetros indicados):

- 30
- Transposición armónica (16 kbps, 14400/28800 Hz)
 - Basada en DFT: 3,68 WMOPS (millones de operaciones ponderadas por segundo);
 - Basada en QMF: 0,98 WMOPS;
 - Pre-procesamiento de transposición QMF (pre-aplanamiento): 0,1WMOPS; y
 - Conformado de envolvente temporal de muestra entre sub-bandas (inter-TES): como máximo 0,16
- 35 WMOPS.

Se conoce que típicamente el rendimiento de la transposición basada en DFT es mejor que el de la transposición basada en QMF para transitorios.

Según algunas realizaciones de la presente invención, un elemento de relleno (de una secuencia de bits de audio codificada) que incluye metadatos eSBR incluye también un parámetro (por ejemplo, un parámetro "bs_extension_id") cuyo valor (por ejemplo, bs_extension_id = 3) señala que los metadatos eSBR están incluidos en el elemento de relleno y que debe realizarse un procesamiento eSBR sobre 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) señala que un contenedor sbr_extension() del elemento de relleno incluye datos PS. Por ejemplo, tal como se indica en la Tabla 1 a continuación, dicho parámetro que tiene el valor bs_extension_id = 2 puede señalar que un contenedor sbr_extension() del elemento de relleno incluye datos PS, y dicho parámetro que tiene el valor bs_extension_id = 3 puede señalar que un contenedor sbr_extension() del elemento de relleno incluye metadatos eSBR:

40

45

Tabla 1

bs_extension_id	Significado	
0	Reservado	
1	Reservado	
2	EXTENSION_ID_PS	
3	EXTENSION_ID_ESBR	

Según 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 tal como se indica en la Tabla 2 a continuación (en la que "sbr_extension()" indica un contenedor que es el elemento de extensión de replicación de banda espectral, "bs_extension_id" es tal como se describe en la Tabla 1 anterior, "ps_data" indica datos PS y "esbr_data" indica metadatos eSBR):

5

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 ejemplar, el esbr_data() al que hace referencia en la Tabla 2 anterior es indicativo de los valores de los siguientes parámetros de metadatos:

10

1. cada uno de los parámetros de metadatos de un bit descritos anteriormente "harmonicSBR"; "bs_interTES"; y "bs_sbr_preprocessing";
2. para cada canal ("ch") de contenido de audio de la secuencia de bits codificada a ser decodificada, cada uno de los parámetros descritos anteriormente: "sbrPatchingMode[ch]"; "sbrOversamplingFlag[ch]"; "sbrPitchInBinsFlag[ch]"; y "sbrPitchInBins[ch]"; y

3. para cada envolvente SBR ("env") de cada canal ("ch") de contenido de audio de la secuencia de bits codificada a ser decodificada, 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:

5

Tabla 3

esbr_data()		
{		
harmonicSBR;	1	
bs_interTes;	1	
bs_sbr_preprocessing;	1	
if (harmonicSBR) {		
if (sbrPatchingMode[0] == 0) {		
sbrOversamplingFlag[0];	1	
if (sbrPitchInBinsFlag[0])	1	
sbrPitchInBins[0];	7	
Else		
sbrPitchInBins[0] = 0;		
} else {		
sbrOversamplingFlag[0] = 0;		
sbrPitchInBins[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	
}		
}		
}		

La sintaxis anterior permite una implementación eficiente de una forma mejorada de replicación de banda espectral, tal como transposición armónica, como una extensión a un decodificador heredado. Específicamente, los datos de eSBR de la Tabla 3 incluyen solo aquellos parámetros necesarios para realizar la forma mejorada de replicación de banda espectral que no están soportados ya en la secuencia de bits o no son directamente derivables a partir de los parámetros ya soportados en la secuencia de bits. Todos los demás parámetros y datos de procesamiento necesarios para realizar la forma mejorada de replicación de banda espectral se extraen desde los parámetros preexistentes en ubicaciones ya definidas en la secuencia de bits.

Por ejemplo, un decodificador compatible con MPEG-4 HE-AAC o HE-AAC v2 puede extenderse para incluir una forma mejorada de replicación de banda espectral, tal como transposición armónica. Esta forma mejorada de replicación de banda espectral es adicional a la forma básica de replicación de banda espectral ya soportada por el decodificador. En el contexto de un decodificador compatible con MPEG-4 HE-AAC o HE-AAC v2, esta forma básica de replicación de banda espectral es la herramienta de transposición espectral QMF SBR tal como se define en la Sección 4.6.18 del estándar MPEG-4 AAC.

Cuando se realiza la forma mejorada de replicación de banda espectral, un decodificador HE-AAC extendido puede reutilizar muchos de los parámetros de secuencia de bits ya incluidos en la carga útil de extensión SBR de la secuencia 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 del parámetro 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 altera la escala de las bandas de frecuencia). Los parámetros que pueden reutilizarse incluyen también parámetros que determinan la tabla de banda de ruido (`bs_noise_bands`) y los parámetros de la tabla de banda de limitador (`bs_limiter_bands`). Por consiguiente, en diversas realizaciones, al menos algunos de los parámetros equivalentes especificados en el estándar USAC se omiten de la secuencia de bits, reduciendo de esta manera la sobrecarga de control en la secuencia de bits. Típicamente, cuando un parámetro especificado en el estándar AAC tiene un parámetro equivalente especificado en el estándar USAC, el parámetro equivalente especificado en el estándar USAC tiene el mismo nombre que el parámetro especificado en el estándar AAC, por ejemplo, el factor de escala de envolvente `EOrigMapped`. Sin embargo, el parámetro equivalente especificado en el estándar USAC típicamente tiene un valor diferente, que se "ajusta" para el procesamiento SBR mejorado definido en el estándar USAC en lugar del procesamiento SBR definido en el estándar AAC.

Además de los numerosos parámetros, otros elementos de datos pueden ser reutilizados también por un decodificador HE-AAC extendido cuando se realiza una forma mejorada de replicación de banda espectral según las realizaciones de la invención. Por ejemplo, los datos de envolvente y de ruido de fondo pueden extraerse también a partir de los datos `bs_data_env` y `bs_noise_env` y pueden usarse durante la forma mejorada de replicación de banda espectral.

En esencia, estas realizaciones explotan los parámetros de configuración y los datos de envolvente ya soportados por un decodificador HE-AAC o HE-AAC v2 heredado en la carga útil de extensión SBR para permitir una forma mejorada de replicación de banda espectral que requiere la menor cantidad posible de datos de transmisión adicionales. Por consiguiente, los decodificadores extendidos que soportan una forma mejorada de replicación de banda espectral pueden crearse de manera muy eficiente basándose en elementos de secuencia de bits ya definidos (por ejemplo, aquellos en la carga útil de extensión SBR) y añadiendo solo aquellos parámetros necesarios para soportar la forma mejorada de replicación de banda espectral (en una carga útil de extensión de elemento de relleno). Esta característica de reducción de datos combinada con la colocación de los parámetros recientemente añadidos en un campo de datos reservado, tal como un contenedor de extensión, reduce sustancialmente las barreras para crear un decodificador que soporte una mejora para la replicación de banda espectral al garantizar que la secuencia de bits sea compatible con el decodificador heredado que no soporta la forma mejorada de replicación de banda espectral.

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

En algunas realizaciones, la invención es un procedimiento que incluye una etapa de codificación de datos de audio para generar una secuencia de bits codificada (por ejemplo, una secuencia de bits MPEG-4 AAC), que incluye metadatos eSBR en al menos un segmento de al menos un bloque de la secuencia de bits codificada y los datos de audio en al menos otro segmento del bloque. En realizaciones típicas, el procedimiento incluye una etapa de multiplexación de los datos de audio con los metadatos eSBR en cada bloque de la secuencia de bits codificada. En la decodificación típica de la secuencia de bits codificada en un decodificador eSBR, el decodificador extrae los metadatos eSBR desde la secuencia de bits (incluyendo mediante el análisis y la demultiplexación de los metadatos

eSBR y los datos de audio) y usa los metadatos eSBR para procesar los datos de audio para generar una secuencia de datos de audio decodificados.

Otro aspecto de la invención es un decodificador eSBR configurado para realizar un procesamiento eSBR (por ejemplo, usando al menos una de las herramientas eSBR conocidas como transposición armónica, pre-aplanamiento o inter_TES) durante la decodificación de una secuencia de bits de audio codificada (por ejemplo, una secuencia de bits de MPEG-4 AAC) que no incluye metadatos eSBR. Un ejemplo de dicho decodificador se describirá con referencia a la Fig. 5.

El decodificador (400) eSBR de la Fig. 5 incluye una memoria 201 intermedia (que es idéntica a la memoria 201 de las Fig. 3 y 4), un deformateador 215 de carga útil de secuencia de bits (que es idéntico al deformateador 215 de la Fig. 4), un subsistema 202 de decodificación de audio (al que se hace referencia a veces como etapa de decodificación "central" o subsistema de decodificación "central", y que es idéntico al subsistema 202 de decodificación central de la Fig. 3), un subsistema 401 de generación de datos de control eSBR y una etapa 203 de procesamiento eSBR (que es idéntica a la etapa 203 de la Fig. 3), conectados como se muestra. Típicamente también, el decodificador 400 incluye otros elementos de procesamiento (no mostrados).

Durante el funcionamiento del decodificador 400, una secuencia de bloques de una secuencia de bits de audio codificada (una secuencia de bits MPEG-4 AAC) recibida por el decodificador 400 se valida desde la memoria 201 intermedia al deformateador 215.

El deformateador 215 está acoplado y configurado para demultiplexar cada bloque de la secuencia de bits para extraer los metadatos SBR (incluyendo los datos de envoltorio cuantificados) y típicamente también otros metadatos a partir de los mismos. El deformateador 215 está configurado para validar al menos los metadatos SBR a la etapa 203 de procesamiento eSBR. El deformateador 215 está acoplado y configurado también para extraer los datos de audio desde cada bloque de la secuencia de bits, y para validar los datos de audio extraídos al subsistema de decodificación 202 (etapa de decodificación).

El subsistema 202 de decodificación de audio del decodificador 400 está configurado para decodificar los datos de audio extraídos por el deformateador 215 (puede hacerse referencia a dicha decodificación como una operación de decodificación "central") para generar los datos de audio decodificados y para validar los datos de audio decodificados a la etapa 203 de procesamiento eSBR. La decodificación se realiza en el dominio de la frecuencia. Típicamente, una etapa final de procesamiento en el subsistema 202 aplica una transformación de dominio de la frecuencia a dominio del tiempo a los datos de audio decodificados en el dominio de la frecuencia, de manera que la salida del subsistema sean datos de audio decodificados en el dominio del tiempo. La etapa 203 está configurada para aplicar herramientas SBR (y herramientas eSBR) indicadas por los metadatos SBR (extraídos por el deformateador 215) y por los metadatos eSBR generados en el subsistema 401, a los datos de audio decodificados (es decir, para realizar un procesamiento SBR y eSBR sobre la salida del subsistema 202 de decodificación usando los metadatos SBR y eSBR) para generar los datos de audio completamente decodificados 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 deformateados y los metadatos emitidos desde el deformateador 215 (y opcionalmente también el subsistema 401), y la etapa 203 está configurada para acceder a los datos de audio y los metadatos según sea necesario durante el procesamiento SBR y eSBR. El procesamiento SBR en la etapa 203 puede considerarse que es un post-procesamiento sobre la salida del subsistema 202 de decodificación central. Opcionalmente, el decodificador 400 incluye también un subsistema de mezcla final (que puede aplicar herramientas estéreo paramétricas ("PS") definidas en el estándar MPEG-4 AAC, usando los metadatos PS extraídos por el deformateador 215) que está acoplado y configurado para realizar una mezcla en la salida de la etapa 203 para generar audio mezclado totalmente descodificado que se emite desde la APU 210.

El subsistema 401 de generación de datos de control de la Fig. 5 está acoplado y configurado para detectar al menos una propiedad de la secuencia de bits de audio codificada a ser decodificada, y para generar datos de control eSBR (que pueden ser o pueden incluir metadatos eSBR de cualquiera de los tipos incluidos en las secuencias de bits de audio codificadas según otras realizaciones de la invención) en respuesta a al menos un resultado de la etapa de detección. Los datos de control eSBR se validan a la etapa 203 para activar la aplicación de herramientas eSBR individuales o combinaciones de herramientas eSBR tras la detección de una propiedad específica (o combinación de propiedades) de la secuencia de bits, y/o para controlar la aplicación de dichas herramientas eSBR. Por ejemplo, con el fin de controlar el rendimiento del procesamiento eSBR usando transposición armónica, algunas realizaciones del subsistema 401 de control de generación de datos incluirían: un detector de música (por ejemplo, una versión simplificada de un detector de música convencional) para configurar el parámetro sbrPatchingMode[ch] (validar afirmar el parámetro establecido a la etapa 203) en respuesta a la detección de que la secuencia de bits es indicativa o no de música; un detector de transitorios para establecer el parámetro sbrOversamplingFlag[ch] (y validar el parámetro establecido a la etapa 203) en respuesta a la detección de la presencia o ausencia de transitorios en el contenido de audio indicado por la secuencia de bits; y/o un detector de tono para establecer los

parámetros `sbrPitchInBinsFlag[ch]` y `sbrPitchInBins[ch]` (y validar los parámetros establecidos a la etapa 203) en respuesta a la detección del tono del contenido de audio indicado por la secuencia de bits. Otros aspectos de la invención son los procedimientos de decodificación de secuencia de bits de audio realizados por cualquier realización del decodificador de la invención descrito en este párrafo y en el párrafo anterior.

Los aspectos de la invención incluyen un procedimiento de codificación o de decodificación del tipo que cualquier realización de la APU, del sistema o del dispositivo de la invención está configurado (por ejemplo, programado) para realizar. Otros aspectos de la invención incluyen un sistema o dispositivo configurado (por ejemplo, programado) para realizar cualquier realización del procedimiento de la invención, y un medio legible por ordenador (por ejemplo, un disco) que almacena código (por ejemplo, de manera no transitoria) para implementar cualquier realización del procedimiento de la invención o etapas del mismo. Por ejemplo, el sistema de la invención puede ser o puede incluir un procesador programable de propósito general, un procesador de señal digital o un microprocesador, programado con software o firmware y/o si no configurado para realizar cualquiera de entre una diversidad de operaciones sobre los datos, incluyendo una realización del procedimiento de la invención o etapas del mismo. Dicho procesador de propósito general puede ser o puede incluir un sistema informático que incluye un dispositivo de entrada, una memoria y un circuito de procesamiento programado (y/o si no configurado) para realizar una realización del procedimiento de la invención (o etapas del mismo) en respuesta a los datos validados al mismos.

Las realizaciones de la presente invención pueden implementarse en hardware, firmware o software, o en 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 inherentemente relacionados con ningún ordenador u otro aparato particular. En particular, pueden usarse diversas máquinas de propósito general con programas escritos según las enseñanzas de la presente memoria, o puede ser más conveniente construir aparatos más especializados (por ejemplo, circuitos integrados) para realizar las etapas de procedimiento requeridas. De esta manera, la invención puede implementarse en uno o más programas de ordenador que se ejecutan en uno o más sistemas informáticos programables (por ejemplo, una implementación de cualquiera de los elementos de la Fig. 1, o el codificador 100 de la Fig. 2 (o un elemento del mismo), o el decodificador 200 de la Fig. 3 (o un elemento del mismo), o el decodificador 210 de la Fig. 4 (o un elemento del mismo), o el decodificador 400 de la Fig. 5 (o un elemento del mismo) comprendiendo cada uno de los mismos al menos un procesador, en al menos un sistema de almacenamiento de datos (incluyendo 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 del programa se aplica a los datos de entrada para realizar las funciones descritas en la presente memoria y para generar la información de salida. La información de salida se aplica a uno o más dispositivos de salida, de manera conocida.

Cada uno de dichos programas puede implementarse en cualquier lenguaje informático deseado (incluyendo los lenguajes de programación máquina, ensamblador, con procedimientos de alto nivel, lógicos u orientados a objetos) para comunicarse con un sistema informático. En cualquier caso, el idioma puede ser un lenguaje compilado o interpretado.

Por ejemplo, cuando se implementan mediante secuencias de instrucciones de software de ordenador, diversas funciones y etapas de las realizaciones de la invención pueden implementarse mediante secuencias de instrucciones de software multiproceso que se ejecutan en hardware de procesamiento de señal digital adecuado, en cuyo caso los diversos dispositivos, etapas y funciones de las realizaciones pueden corresponder a partes de las instrucciones de software.

Preferiblemente, cada uno de dichos programas de ordenador se almacena en, o se descarga a, un dispositivo o medio de almacenamiento (por ejemplo, una memoria o un medio de estado sólido, o un medio magnético u óptico) legible por un ordenador programable de propósito general o especial, para configurar y operar el ordenador cuando el almacenamiento el sistema informático lee los medios o dispositivos para realizar los procedimientos descritos en la presente memoria. El sistema de la invención puede implementarse también como un medio de almacenamiento legible por ordenador, configurado con (es decir, que almacena) un programa de ordenador, en el que el medio de almacenamiento configurado de esta manera causa que un sistema informático opere de una manera específica y predefinida para realizar las funciones descritas en la presente memoria.

Se han descrito una serie de realizaciones de la invención. Sin embargo, se entenderá que pueden realizarse 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 puede llevarse a la práctica de una forma distinta a la descrita específicamente en la presente memoria. Todos los números de referencia contenidos en las siguientes reivindicaciones tienen solo propósitos ilustrativos y no deberían usarse para interpretar o limitar en modo alguno las reivindicaciones.

REIVINDICACIONES

1. Unidad (210) de procesamiento de audio que comprende:

una memoria (201) intermedia configurada para almacenar al menos un bloque de una secuencia de bits de audio codificada;

5 un deformateador (215) de carga útil de secuencia de bits acoplado a la memoria intermedia y configurado para demultiplexar al menos una parte del al menos un bloque de la secuencia de bits de audio codificada; y

un subsistema (202) de decodificación acoplado al deformateador (215) de carga útil de la secuencia de bits y configurado para decodificar al menos una parte del al menos un bloque de la secuencia de bits de audio codificada, en el que el al menos un bloque de la secuencia de bits de audio codificada incluye:

10 un elemento de relleno con un identificador que indica un inicio del elemento de relleno y datos de relleno después del identificador, caracterizado por que los datos de relleno incluyen:

al menos un indicador que identifica si el procesamiento de replicación de banda espectral mejorada, eSBR, debe realizarse o no sobre el contenido de audio del al menos un bloque de la secuencia de bits de audio codificada, y

15 metadatos de replicación de banda espectral mejorada que no incluyen uno o más parámetros usados tanto para la transposición espectral como para la transposición armónica, en el que los metadatos de replicación de banda espectral mejorada son metadatos configurados para habilitar al menos una herramienta eSBR que se describe en el estándar *Moving Picture Experts Group*, MPEG, *Unified Speech and Audio Coding*, USAC, y que no se describe en el estándar MPEG ISO/IEC 14496-3:2009; comprendiendo adicionalmente la unidad de procesamiento de audio:

20 un subsistema (203) de procesamiento de replicación de banda espectral mejorada configurado para realizar procesamiento de replicación de banda espectral mejorada usando o en respuesta al al menos un indicador, en el que la replicación de banda espectral mejorada incluye transposición armónica, en el que la unidad de procesamiento de audio es un decodificador de audio, y el identificador es un entero sin signo de tres bits transmitiéndose primero el bit más significativo y que tiene un valor de 0x6.

2. Unidad de procesamiento de audio según la reivindicación 1, en la que el al menos un bloque de la secuencia de bits de audio codificada incluye metadatos de replicación de banda espectral.

3. Unidad de procesamiento de audio según la reivindicación 2, en la que los metadatos de replicación de banda espectral mejorada no incluyen un parámetro equivalente a un parámetro de los metadatos de replicación de banda espectral.

4. Unidad de procesamiento de audio según la reivindicación 2 o la reivindicación 3, en la que los metadatos de replicación de banda espectral:

son metadatos configurados para habilitar al menos una herramienta SBR que se describe en el estándar MPEG ISO/IEC 14496-3:2009; y/o

35 incluyen uno o más parámetros usados tanto para la transposición espectral como para la transposición armónica.

5. Unidad de procesamiento de audio según una cualquiera de las reivindicaciones anteriores, en la que el procesamiento de replicación de banda espectral mejorada incluye transposición armónica y no transposición espectral.

40 6. Un procedimiento para la decodificación de una secuencia de bits de audio codificada, comprendiendo el procedimiento:

recibir al menos un bloque de una secuencia de bits de audio codificada;

demultiplexar al menos una parte del al menos un bloque de la secuencia de bits de audio codificada;

decodificar al menos una parte del al menos un bloque de la secuencia de bits de audio codificada;

45 en el que el al menos un bloque de la secuencia de bits de audio codificada incluye:

un elemento de relleno con un identificador que indica un inicio del elemento de relleno y datos de relleno después del identificador, caracterizado por que los datos de relleno incluyen:

al menos un indicador que identifica si debe realizarse o no el procesamiento de replicación de banda espectral mejorada, eSBR, sobre el contenido de audio del al menos un bloque de la secuencia de bits de audio codificada, y

5 metadatos de replicación de banda espectral mejorada que no incluyen uno o más parámetros usados tanto para la transposición espectral como para la transposición armónica, en el que los metadatos de replicación de banda espectral mejorada son metadatos configurados para habilitar al menos una herramienta eSBR que se describe en el estándar *Moving Picture Experts Group*, MPEG, *Unified Speech and Audio Coding*, USAC, y que no se describe en el estándar MPEG ISO/IEC 14496-3:2009; y comprendiendo además el procedimiento:

10 realizar procesamiento de replicación de banda espectral mejorada usando o en respuesta al al menos un indicador, en el que la replicación de banda espectral mejorada incluye transposición armónica, en el que el identificador es un entero sin signo de tres bits transmitiéndose primero el bit más significativo y que tiene un valor de 0x6.

15 7. Procedimiento según la reivindicación 6, en el que los datos de relleno incluyen una carga útil de extensión, la carga útil de extensión incluye datos de extensión de replicación de banda espectral, y la carga útil de extensión se identifica con un entero sin signo de cuatro bits transmitiéndose primero el bit más significativo y que tiene un valor de '1101' o '1110' y, opcionalmente,

en el que los datos de extensión de replicación de banda espectral incluyen:

una cabecera de replicación de banda espectral opcional,

datos de replicación de banda espectral después de la cabecera,

20 un elemento de extensión de replicación de banda espectral después de los datos de replicación de banda espectral, y en el que el indicador está incluido en el elemento de extensión de replicación de banda espectral.

25 8. Procedimiento según una cualquiera de las reivindicaciones 6-7, en el que el procesamiento de replicación de banda espectral mejorada es una transposición armónica, un valor del al menos un indicador indica que debe realizarse dicho procesamiento de replicación de banda espectral mejorada sobre el contenido de audio del al menos un bloque de la secuencia de bits de audio codificada, y otro valor del al menos un indicador indica que debe realizarse una transposición espectral, pero no dicha transposición armónica, sobre el contenido de audio del al menos un bloque de la secuencia de bits de audio codificada.

30 9. Medio o dispositivo de almacenamiento en el que se almacena un programa de ordenador legible por un ordenador programable de propósito general o especial, adaptado para configurar y operar el ordenador cuando el ordenador lee el medio o dispositivo de almacenamiento para realizar el procedimiento según una cualquiera de las reivindicaciones 6 a 8.

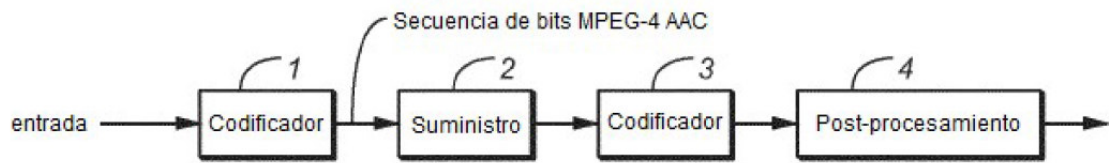


FIG. 1

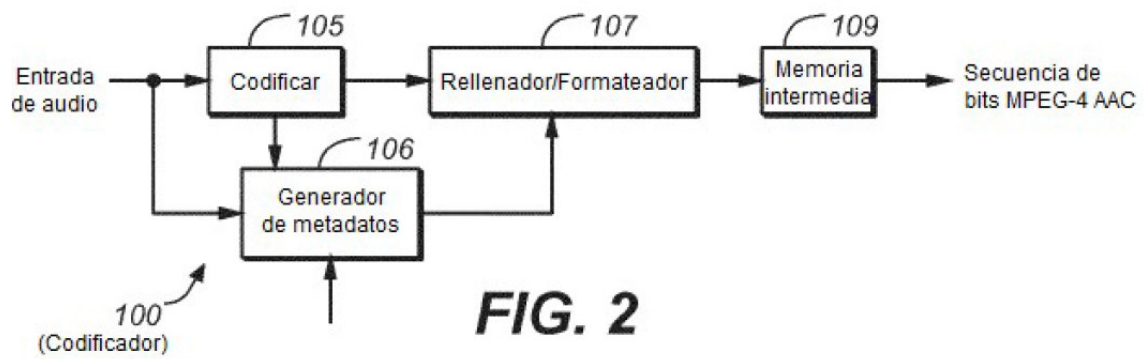


FIG. 2

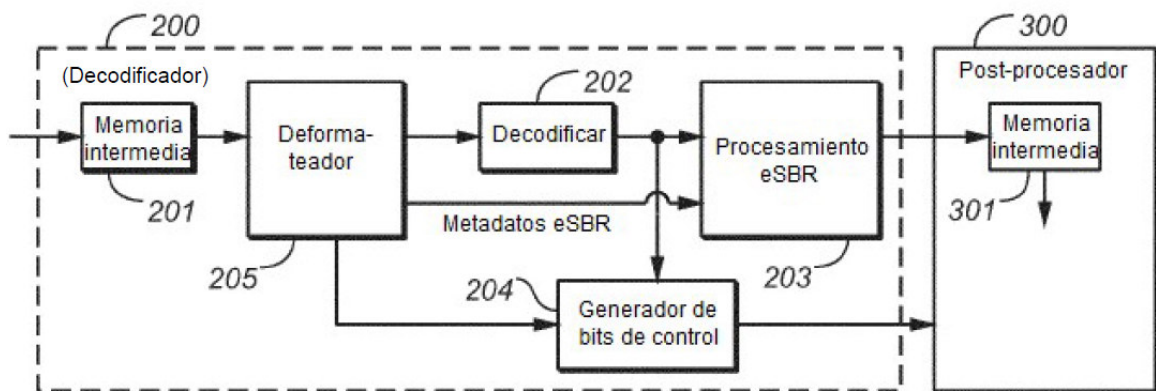


FIG. 3

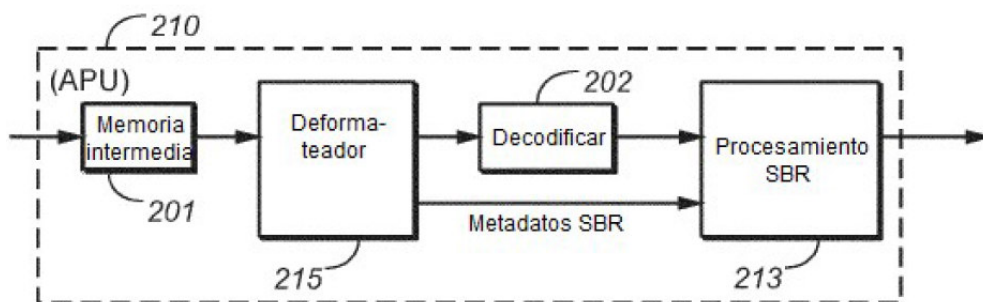


FIG. 4

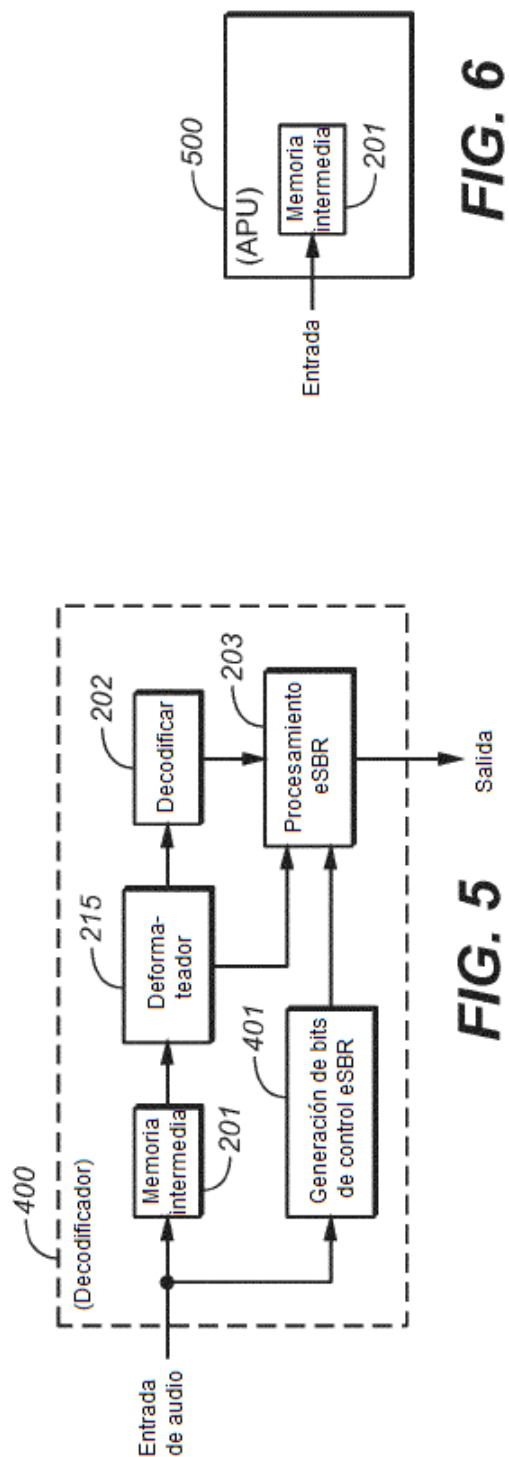


FIG. 6

