

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 624 419**

51 Int. Cl.:

H03G 9/00

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **15.01.2014 PCT/US2014/011681**

87 Fecha y número de publicación internacional: **24.07.2014 WO14113471**

96 Fecha de presentación y número de la solicitud europea: **15.01.2014 E 14703670 (1)**

97 Fecha y número de publicación de la concesión europea: **15.03.2017 EP 2946469**

54 Título: **Sistema y procedimiento para optimizar la sonoridad y el rango dinámico a través de diferentes dispositivos de reproducción**

30 Prioridad:

21.01.2013 US 201361754882 P
05.04.2013 US 201361809250 P
16.05.2013 US 201361824010 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
14.07.2017

73 Titular/es:

DOLBY LABORATORIES LICENSING CORPORATION (50.0%)
1275 Market Street
San Francisco, CA 94103, US y
DOLBY INTERNATIONAL AB (50.0%)

72 Inventor/es:

RIEDMILLER, JEFFREY;
NORCROSS, SCOTT GREGORY y
ROEDEN, KARL JONAS

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 624 419 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

Sistema y procedimiento para optimizar la sonoridad y el rango dinámico a través de diferentes dispositivos de reproducción.

Referencia cruzada a las aplicaciones relacionadas

- 5 Esta solicitud reivindica prioridad a la solicitud de patente provisional de los Estados Unidos n.º 61/754.882 presentada el 21 de enero de 2013; a la solicitud de patente provisional de los Estados Unidos n.º 61/809.250 presentada el 5 de abril de 2013; y a la solicitud de patente provisional de los Estados Unidos n.º 61/824.010 presentada el 16 de mayo de 2013.

Campo de la invención

- 10 Una o más realizaciones se refieren, en general, al procesamiento de señales de audio, y más específicamente al procesamiento de flujos de bits de datos de audio con metadatos indicativos de las características de sonoridad y de rango dinámico del contenido de audio basadas en entornos y dispositivos de reproducción.

Antecedentes

- 15 El objeto tratado en la sección de antecedentes no debe ser asumido como técnica anterior únicamente como resultado de su mención en la sección de antecedentes. De forma similar, no debe asumirse que haya sido reconocido previamente en la técnica anterior un problema mencionado en la sección de antecedentes o asociado con el objeto de la sección de antecedentes. El objeto en la sección de antecedentes únicamente representa diferentes enfoques, que en sí mismos también pueden ser invenciones.

- 20 El rango dinámico de una señal de audio es, en general, la relación entre los valores más grandes y más pequeños posibles del sonido integrado en la señal, y habitualmente se mide como un valor de decibelio (en base 10). En muchos sistemas de procesamiento de audio, el control de rango dinámico (o compresión de rango dinámico, DRC) se utiliza para reducir el nivel de sonidos fuertes y/o amplificar el nivel de sonidos bajos para ajustarse a una fuente de rango dinámico amplio contenido en un rango dinámico registrado más estrecho que puede ser almacenado y reproducido con mayor facilidad mediante equipos electrónicos. Para contenido audio/visual (AV), se puede usar un nivel de referencia de diálogo para definir el punto "nulo" para compresión a través del mecanismo de la DRC. La DRC actúa para realzar el contenido por debajo del nivel de referencia del diálogo y recortar el contenido por encima del nivel de referencia.

- 30 En un sistema de codificación de audio conocido, se usan metadatos asociados con la señal de audio para establecer el nivel de DRC basado en el tipo y uso pretendido del contenido. El modo DRC establece la cantidad de compresión aplicada a la señal de audio y define el nivel de referencia de salida del decodificador. Tales sistemas pueden estar limitados a dos ajustes del nivel de DRC que están programados en el codificador y ser seleccionados por el usuario. Por ejemplo, un valor de dialnorm (diálogo normalizado) de -31 dB (Línea) se utiliza tradicionalmente para el contenido que se reproduce en un AVR o dispositivos con capacidad de rango dinámico completo y se utiliza un valor de dialnorm de -20 dB (RF) para el contenido reproducido en aparatos de televisión o dispositivos similares.
- 35 Este tipo de sistema permite que un único flujo de bits de audio se utilice en dos comunes pero muy diferentes escenarios de reproducción a través del uso de dos diferentes conjuntos de metadatos de DRC. Sin embargo, estos sistemas están limitados a los valores de dialnorm preestablecidos y no están optimizados para su reproducción en la amplia variedad de los diferentes dispositivos de reproducción y entornos de audición que ahora son posibles gracias a la llegada de medios digitales y tecnología de transmisión por Internet.

- 40 En los actuales sistemas de codificación de audio basados en metadatos, un flujo de datos de audio puede incluir tanto contenido de audio (p. ej., uno o más canales de contenido de audio) como metadatos indicativos de al menos una característica del contenido de audio. Por ejemplo, en un flujo de bits AC-3 hay varios parámetros de metadatos de audio que están específicamente destinados a ser utilizados para cambiar el sonido del programa enviado a un entorno de audición. Uno de los parámetros de metadatos es el parámetro dialnorm, que indica el nivel medio de la sonoridad del diálogo (o sonoridad media del contenido) que se produce en un programa de audio, y se utiliza para determinar el nivel de la señal de reproducción de audio.

- 50 Durante la reproducción de un flujo de bits que comprende una secuencia de diferentes segmentos de programa de audio (teniendo cada uno un parámetro dialnorm diferente), un decodificador AC-3 utiliza el parámetro dialnorm de cada segmento para realizar un tipo de procesamiento de la sonoridad que modifica el nivel de reproducción del segmento o sonoridad, de manera que la sonoridad percibida del diálogo del segmento esté en un nivel constante. Cada segmento de audio codificado (elemento) en una secuencia de elementos de audio codificados tendría (en general) un parámetro dialnorm diferente, y el decodificador escalaría el nivel de cada uno de los elementos de modo que el nivel de reproducción o sonoridad del diálogo para cada elemento sea el mismo o muy similar, aunque esto puede requerir la aplicación de diferentes cantidades de ganancia a los diferentes elementos durante la reproducción.

- 55 En algunas realizaciones, el parámetro dialnorm es establecido por un usuario, y no se genera automáticamente,

aunque existe un valor de dialnorm por defecto si el usuario no establece ningún valor. Por ejemplo, un creador de contenidos puede realizar mediciones de la sonoridad con un dispositivo externo a un codificador AC-3 y luego transferir el resultado (indicativo de la sonoridad del diálogo hablado de un programa de audio) al codificador para establecer el valor de dialnorm. De ese modo, existe confianza en el creador del contenido para establecer correctamente el parámetro dialnorm.

Hay varias razones diferentes por las que el parámetro dialnorm en un flujo de bits AC-3 puede ser incorrecto. En primer lugar, cada codificador AC-3 tiene un valor de dialnorm por defecto que se utiliza durante la generación del flujo de bits si el creador de contenido no establece un valor de dialnorm. Este valor por defecto puede ser sustancialmente diferente del nivel real de sonoridad del diálogo del audio. En segundo lugar, incluso si un creador de contenido mide la sonoridad y establece el valor de dialnorm en estos términos, un algoritmo de medida o medidor de la sonoridad se puede haber utilizado de modo que no se ajusta al procedimiento de medición de la sonoridad recomendado, dando como resultado un valor de dialnorm incorrecto. En tercer lugar, incluso si se ha creado un flujo de bits AC-3 con el valor de dialnorm medido y ajustado correctamente por el creador de contenido, se puede haber cambiado a un valor incorrecto por un módulo intermedio durante la transmisión y/o almacenamiento del flujo de bits. Por ejemplo, no es infrecuente en las aplicaciones de difusión de televisión para los flujos de bits AC-3 que se decodifiquen, se modifiquen y luego se vuelvan a codificar utilizando información de metadatos de dialnorm incorrectos. De ese modo, un valor de dialnorm incluido en un flujo de bits AC-3 puede ser incorrecto o inexacto y, por lo tanto, puede tener un impacto negativo en la calidad de la experiencia de audición.

Además, el parámetro dialnorm no indica el estado de procesamiento de la sonoridad de los datos de audio correspondientes (p. ej., qué tipo (o tipos) de procesamiento de la sonoridad que se ha realizado sobre los datos de audio). Además, los sistemas de la sonoridad y DRC actualmente desplegados, como los sistemas en Dolby Digital (DD) y sistemas Dolby Digital Plus (DD+), fueron diseñados para mezclar el contenido AV en la sala de estar de un consumidor o en un cine. Para adaptar dicho contenido para su reproducción en otros entornos y equipos de audición (p. ej., un dispositivo móvil), debe aplicarse un posprocesamiento "a ciegas" en el dispositivo de reproducción para adaptar el contenido AV a ese entorno de audición. En otras palabras, un posprocesador (o un decodificador) asume que el nivel de la sonoridad del contenido recibido está en un nivel particular (p. ej., -31 o -20 dB) y el posprocesador fija el nivel a un nivel objetivo fijado predeterminado adecuado para un dispositivo en particular. Si el nivel de la sonoridad asumido o el nivel objetivo predeterminado son incorrectos, el posprocesamiento puede tener el resultado opuesto al pretendido; es decir, el posprocesamiento puede hacer que el audio de salida sea menos deseable para un usuario.

Las realizaciones descritas no se limitan a su uso con un flujo de bits AC-3, un flujo de bits E-AC-3 o un flujo de bits Dolby E, sin embargo, por comodidad, tales flujos de bits se tratarán conjuntamente con un sistema que incluye los metadatos del estado de procesamiento de la sonoridad. Dolby, Dolby Digital, Dolby Digital Plus y Dolby E son marcas registradas de Dolby Laboratories Licensing Corporation. Dolby Laboratories ofrece aplicaciones propias de AC-3 y de E-AC-3 conocidas como Dolby Digital y Dolby Digital Plus, respectivamente.

En la publicación de solicitud de patente de Estados Unidos número US 2006/0002572 A1, una señal codificada transmite información de audio codificada y metadatos que pueden usarse para controlar la sonoridad y el rango dinámico de la información de audio durante su reproducción. Si los valores de estos parámetros de metadatos se establecen incorrectamente, pueden producirse fluctuaciones molestas en la sonoridad durante la reproducción. Este problema se supera detectando los valores de parámetros de metadatos incorrectos en la señal y reemplazando los valores incorrectos por valores corregidos.

Breve resumen de realizaciones

Las realizaciones se dirigen a un procedimiento según se expone en la reivindicación 1.

Las realizaciones se dirigen además a un aparato tal como el expuesto en la reivindicación 8.

Las realizaciones se dirigen todavía además a un medio legible por ordenador como el expuesto en la reivindicación 7.

Breve descripción de los dibujos

En los dibujos siguientes, números de referencia similares se utilizan para referirse a elementos similares. Aunque las figuras siguientes representan varios ejemplos, las aplicaciones descritas en el presente documento no se limitan a los ejemplos representados en las figuras.

La Fig. 1 es un diagrama de bloques de una realización de un sistema de procesamiento de audio configurado para realizar la optimización de la sonoridad y el rango dinámico, en algunas realizaciones.

La Fig. 2 es un diagrama de bloques de un codificador para uso en el sistema de la Fig. 1, en algunas realizaciones.

La Fig. 3 es un diagrama de bloques de un decodificador para uso en el sistema de la Fig. 1, en algunas realizaciones.

La Fig. 4 es un diagrama de una trama del AC-3, que incluye los segmentos en los que está dividida.

La Fig. 5 es un diagrama del segmento de Información de la Sincronización (SI) de una trama del AC-3, que incluye los segmentos en los que está dividido.

5 La Fig. 6 es un diagrama del segmento de la Información de Flujo de Bits (BSI) de una trama AC-3, que incluye los segmentos en los que está dividido.

La Fig. 7 es un diagrama de una trama del E-AC-3, que incluye los segmentos en los que está dividida.

La Fig. 8 es una tabla que ilustra ciertas tramas de un flujo de bits codificado y el formato de metadatos, en algunas realizaciones.

10 La información de la sonoridad asociada con un perfil de salida, y transmitir la información de la sonoridad determinada para el perfil de salida para uso en la mezcla del audio. En una realización, el paso de determinar la información de la sonoridad asociada con un perfil de salida puede incluir además analizar las características del perfil de salida, determinar los parámetros basados en las características, y transmitir la información de la sonoridad determinada comprende transmitir los parámetros determinados. La información de la sonoridad puede incluir parámetros de la sonoridad para o característicos de un perfil de salida. En una realización, el procedimiento puede comprender además determinar un flujo codificado de baja velocidad de bits a transmitir, en donde la información de la sonoridad comprende características para uno o más perfiles de salida.

15

Las realizaciones se dirigen además a un aparato o sistema que incluye procesar componentes que realizan los actos descritos en las realizaciones de los procedimientos de decodificación anteriores.

Breve descripción de los dibujos

20 En los dibujos siguientes, números de referencia similares se utilizan para referirse a elementos similares. Aunque las figuras siguientes representan varios ejemplos, las aplicaciones descritas en el presente documento no se limitan a los ejemplos representados en las figuras.

La Fig. 1 es un diagrama de bloques de una realización de un sistema de procesamiento de audio configurado para realizar la optimización de la sonoridad y el rango dinámico, en algunas realizaciones.

25 La Fig. 2 es un diagrama de bloques de un codificador para uso en el sistema de la Fig. 1, en algunas realizaciones.

La Fig. 3 es un diagrama de bloques de un decodificador para uso en el sistema de la Fig. 1, en algunas realizaciones.

La Fig. 4 es un diagrama de una trama del AC-3, que incluye los segmentos en los que está dividida.

30 La Fig. 5 es un diagrama del segmento de Información de la Sincronización (SI) de una trama AC-3, que incluye los segmentos en los que está dividido.

La Fig. 6 es un diagrama del segmento de la Información de Flujo de Bits (BSI) de una trama AC-3, que incluye los segmentos en los que está dividido.

La Fig. 7 es un diagrama de una trama del E-AC-3, que incluye los segmentos en los que está dividida.

35 La Fig. 8 es una tabla que ilustra ciertas tramas de un flujo de bits codificado y el formato de metadatos, en algunas realizaciones.

La Fig. 9 es una tabla que ilustra un formato de metadatos del estado de procesamiento de la sonoridad, en algunas realizaciones.

La Fig. 10 es un diagrama de bloques más detallado del sistema de procesamiento de audio de la Fig. 1 que se puede configurar para realizar la optimización de la sonoridad y el rango dinámico, en algunas realizaciones.

40 La Fig. 11 es una tabla que ilustra diferentes requisitos de rango dinámico para una variedad de dispositivos de reproducción y entornos de audición de fondo en un ejemplo práctico de uso.

La Fig. 12 es un diagrama de bloques de un sistema de optimización de rango dinámico, en una realización.

La Fig. 13 es un diagrama de bloques que ilustra una interfaz entre diferentes perfiles para una variedad de diferentes clases de dispositivos de reproducción, en una realización.

45 La Fig. 14 es una tabla que ilustra la correlación entre la sonoridad a largo plazo y el rango dinámico a corto plazo para una pluralidad de perfiles definidos, en una realización.

La Fig. 15 ilustra ejemplos de perfiles de la sonoridad para diferentes tipos de contenido de audio, en una

realización.

La Fig. 16 es un diagrama de flujo que ilustra un procedimiento para optimizar la sonoridad y el rango dinámico a través de dispositivos y aplicaciones de reproducción, en una realización.

Descripción detallada

5 Definiciones y Nomenclatura

A lo largo de esta descripción, incluidas las reivindicaciones, la expresión que realiza una operación "sobre" una señal o dato (p. ej., filtrar, escalar, transformar o aplicar ganancia a la señal o dato) se usa en un sentido amplio para indicar la realización de la operación directamente sobre la señal o dato, o sobre una versión procesada de la señal o dato (p. ej., sobre una versión de la señal que ha sido sometida a un filtrado preliminar o preprocesamiento antes de la realización de la operación). La expresión "sistema" se usa en un sentido amplio para indicar un dispositivo, sistema o subsistema. Por ejemplo, un subsistema que aplica un decodificador puede mencionarse como un sistema decodificador y un sistema que incluye dicho subsistema (p. ej., un sistema que genera X señales de salida en respuesta a múltiples entradas, en las que el subsistema genera M de las entradas y las otras entradas X - M se reciben desde una fuente externa) también se pueden mencionar como un sistema decodificador. El término "procesador" se usa en un sentido amplio para indicar un sistema o dispositivo programable o si no configurable (p. ej., con software o firmware) para realizar operaciones sobre datos (p. ej., audio, vídeo u otros datos de imagen). Ejemplos de procesadores incluyen una matriz de puertas programables por campo (u otro circuito integrado o conjunto de chips configurable), un procesador de señales digitales programado y/o, si no, configurado para realizar el procesamiento canalizado en audio u otros datos de sonido, un procesador programable de propósito general u ordenador, y un chip o conjunto de chips de microprocesador programable.

Las expresiones "procesador de audio" y "unidad de procesamiento de audio" se usan indistintamente, y en sentido amplio, para indicar un sistema configurado para procesar datos de audio. Ejemplos de unidades de procesamiento de audio incluyen pero no se limitan a codificadores (p. ej., transcodificadores), decodificadores, códecs, sistemas de preprocesamiento, sistemas de posprocesamiento y sistemas de procesamiento de flujo de bits (a veces mencionados como herramientas de procesamiento de flujo de bits). La expresión "metadatos del estado de procesamiento" (p. ej., como en la expresión "metadatos del estado de procesamiento de la sonoridad") se refiere a datos separados y diferentes de los datos de audio correspondientes (el contenido de audio de un flujo de datos de audio que incluye también metadatos del estado de procesamiento). El procesamiento del estado de metadatos está asociado con datos de audio, indica el estado de procesamiento de la sonoridad de los datos de audio correspondientes (p. ej., qué tipo (o tipos) de procesamiento ya se han realizado en los datos de audio) y opcionalmente también indica al menos una particularidad o característica de los datos de audio. En alguna realización, la asociación de los metadatos del estado de procesamiento con los datos de audio es sincrónica temporal. De este modo, los metadatos del estado de procesamiento actuales (más recientemente recibidos o actualizados) indican que los datos de audio correspondientes comprenden de forma contemporánea los resultados del tipo (o tipos) indicado del procesamiento de datos de audio. En algunos casos, los metadatos del estado de procesamiento pueden incluir el historial de procesamiento y/o algunos o todos los parámetros que se utilizan en y/o derivados de los tipos de procesamiento indicados. Además, los metadatos del estado de procesamiento pueden incluir al menos una particularidad o característica de los datos de audio correspondientes, que se han calculado o extraído de los datos de audio. Los metadatos del estado de procesamiento también pueden incluir otros metadatos que no están relacionados con o derivados de ningún procesamiento de los datos de audio correspondientes. Por ejemplo, datos de terceros, información de seguimiento, identificadores, información de propiedad o estándar, datos de anotación de usuario, datos de preferencia de usuario, etc., pueden ser añadidos por una unidad de procesamiento de audio en particular para pasar a otras unidades de procesamiento de audio.

La expresión "metadatos del estado de procesamiento de la sonoridad" (o "LPSM") indica metadatos del estado de procesamiento indicativos del estado de procesamiento de la sonoridad de los datos de audio correspondientes (p. ej., qué tipo (o tipos) de procesamiento de la sonoridad se han realizado sobre los datos de audio) y opcionalmente también al menos una particularidad o característica (p. ej., sonoridad) de los datos de audio correspondientes. Los metadatos del estado de procesamiento de la sonoridad pueden incluir datos (p. ej., otros metadatos) que no son metadatos del estado de procesamiento de la sonoridad (es decir, cuando se consideran solos). El término "acopla" o "acoplado" se usa para significar una conexión directa o indirecta.

Se describen sistemas y procedimientos para un codificador/decodificador de audio que normaliza no destructivamente la sonoridad y el rango dinámico del audio a través de diversos dispositivos que requieren o utilizan diferentes valores de la sonoridad objetivo y tienen diferentes capacidades de rango dinámico. Los procedimientos y componentes funcionales según algunas realizaciones envían información acerca del contenido de audio desde el codificador al decodificador para uno o más perfiles de dispositivo. Un perfil del dispositivo especifica la sonoridad y el rango dinámico objetivos deseados para uno o más dispositivos. El sistema es extensible, de manera que se pueden soportar nuevos perfiles de dispositivo con diferentes objetivos de sonoridad "nominal".

En una realización, el sistema genera ganancias adecuadas basadas en requisitos de control de la sonoridad y rango dinámico en el codificador o genera las ganancias en el decodificador, bajo control desde el codificador a

través de la parametrización de las ganancias originales para reducir la velocidad de los datos. El sistema del rango dinámico incluye dos mecanismos para aplicar el control de la sonoridad: un perfil de rango dinámico artístico que proporciona el control de los creadores de contenido sobre cómo se reproducirá el audio, y un mecanismo de protección separado para asegurar que no se produzcan sobrecargas para los diversos perfiles de reproducción. El sistema está configurado también para permitir que otros parámetros de metadatos (internos o externos) se utilicen para controlar adecuadamente las ganancias y/o perfiles de la sonoridad y del rango dinámico. El decodificador está configurado para soportar una entrada auxiliar de n canales que aprovechará los ajustes/procesamiento de la sonoridad y rango dinámico del lado del decodificador.

En algunas realizaciones, los metadatos del estado de procesamiento de la sonoridad (LPSM) están incrustados en uno o más campos reservados (o ranuras) de segmentos de metadatos de un flujo de bits de audio que también incluye datos de audio en otros segmentos (segmentos de datos de audio). Por ejemplo, al menos un segmento de cada trama de flujo de bits incluye LPSM, y al menos otro segmento de la trama incluye datos de audio correspondientes (es decir, datos de audio cuyo estado de procesamiento de la sonoridad y la sonoridad están indicados por el LPSM). En algunas realizaciones, el volumen de los datos de los LPSM puede ser suficientemente pequeño para ser transportados sin afectar a la velocidad de bits asignada para transportar los datos de audio.

La comunicación de metadatos del estado de procesamiento de la sonoridad en una cadena del procesamiento de datos de audio es particularmente útil cuando dos o más unidades de procesamiento de audio necesitan trabajar en tándem entre sí a lo largo de la cadena de procesamiento (o ciclo de vida del contenido). Sin la inclusión de metadatos del estado de procesamiento de la sonoridad en un flujo de bits de audio, pueden producirse problemas de procesamiento de los medios tales como degradaciones de la calidad, del nivel y espaciales, por ejemplo, cuando se utilizan dos o más codecs de audio en la cadena y se aplica nivelación de volumen de un solo terminal más de una vez durante el trayecto del flujo de bits hacia un dispositivo de consumo de los medios (o un punto de mezcla del contenido de audio del flujo de bits).

Sistema de procesamiento de metadatos de la sonoridad y rango dinámico

La Fig. 1 es un diagrama de bloques de una realización de un sistema de procesamiento de audio que puede ser configurado para realizar la optimización de la sonoridad y rango dinámico, en algunas realizaciones que utilizan ciertos componentes de procesamiento de metadatos (p. ej., preprocesamiento y posprocesamiento). La Fig. 1 ilustra un ejemplo de cadena de procesamiento de audio (un sistema del procesamiento de datos de audio), en el que uno o más de los elementos del sistema pueden ser configurados según una realización de la presente invención. El sistema 10 de la Fig. 1 incluye los elementos siguientes, acoplados juntos como se muestra: una unidad 12 de preprocesamiento, un codificador 14, una unidad 16 de análisis de señal y de corrección de metadatos, un transcodificador 18, un decodificador 20 y una unidad 24 de posprocesamiento. En variaciones del sistema mostrado, uno o más de los elementos se omiten o se incluyen unidades del procesamiento de datos de audio adicionales. Por ejemplo, en una realización, la unidad 22 de posprocesamiento es parte del decodificador 20 en lugar de ser una unidad separada.

En algunas aplicaciones, la unidad de preprocesamiento de la Fig. 1 está configurada para aceptar muestras de PCM (dominio del tiempo) que comprenden audio contenido como entrada 11, y para emitir muestras PCM procesadas. El codificador 14 puede ser configurado para aceptar las muestras PCM como entrada y emitir un flujo de bits de audio codificado (p. ej., comprimido) indicativo del contenido de audio. Los datos del flujo de bits que son indicativos del contenido de audio a veces se denominan "datos de audio" en el presente documento. En una realización, el flujo de bits de audio procedente del codificador incluye metadatos del estado de procesamiento de la sonoridad (y opcionalmente también otros metadatos) así como datos de audio.

La unidad 16 de análisis de señal y de corrección de metadatos puede aceptar uno o más flujos de bits de audio codificados como entrada y determinar (p. ej., validar) si los metadatos del estado de procesamiento en cada flujo de bits de audio codificado son correctos, al realizar análisis de señal. En algunas realizaciones, la validación puede ser realizada por un componente validador del estado, tal como el elemento 102 mostrado en la Fig. 2 y una técnica de validación de este tipo se describe más adelante en el contexto del validador 102 del estado. En algunas realizaciones, la unidad 16 está incluida en el codificador y la validación se realiza por la unidad 16 o por el validador 102. Si la unidad de análisis de señal y de corrección de los metadatos encuentra que los metadatos incluidos no son válidos, la unidad 16 de corrección de metadatos realiza el análisis de señal para determinar el valor (o valores) correcto y reemplaza el valor (o valores) incorrecto por el valor (o valores) correcto determinado. De este modo, cada flujo de bits de audio codificado procedente de la unidad de análisis de señal y de corrección de metadatos puede incluir metadatos del estado de procesamiento corregidos, así como datos de audio codificados. La unidad 16 de análisis de señal y de corrección de metadatos puede ser parte de la unidad 12 de preprocesamiento, del codificador 14, del transcodificador 18, del decodificador 20 o de la unidad 22 de posprocesamiento. Alternativamente, la unidad 16 de análisis de señal y de corrección de metadatos puede ser una unidad separada o parte de otra unidad en la cadena de procesamiento de audio.

El transcodificador 18 puede aceptar flujos de bits de audio codificados como entrada, y emitir flujos de bits de audio modificados (p. ej., codificados de forma diferente) en respuesta (p. ej., decodificando un flujo de entrada y volviendo a codificar el flujo decodificado en un formato de codificación diferente). El flujo de bits de audio procedente del

transcodificador incluye metadatos del estado de procesamiento de la sonoridad (y opcionalmente también otros metadatos) así como datos de audio codificados. Los metadatos pueden haber sido incluidos en el flujo de bits.

El decodificador 20 de la Fig. 1 puede aceptar flujos de bits de audio codificados (p. ej., comprimidos) como entrada, y emitir (en respuesta) flujos de muestras de audio PCM decodificadas. En una realización, la salida del decodificador es o incluye cualquiera de los siguientes: un flujo de muestras de audio y un flujo correspondiente de metadatos del estado de procesamiento de la sonoridad (y opcionalmente también otros metadatos) extraídos de un flujo de bits codificado de entrada; un flujo de muestras de audio, y un flujo correspondiente de bits de control determinados a partir de metadatos del estado de procesamiento de la sonoridad (y opcionalmente también otros metadatos) extraídos de un flujo de bits codificado de entrada; o un flujo de muestras de audio, sin un flujo correspondiente de metadatos del estado de procesamiento o bits de control determinados a partir de metadatos del estado de procesamiento. En este último caso, el decodificador puede extraer metadatos del estado de procesamiento de la sonoridad (y/u otros metadatos) del flujo de bits codificado de entrada y realizar al menos una operación sobre los metadatos extraídos (p. ej., validación), aunque no emite los metadatos extraídos o los bits de control determinados a partir de los mismos.

Configurando la unidad de posprocesamiento de la Fig. 1 según una realización de la presente invención, la unidad 22 de posprocesamiento está configurada para aceptar un flujo de muestras de audio PCM decodificadas, y llevar a cabo el posprocesamiento sobre el mismo (p. ej., nivelación del volumen del contenido de audio) utilizando metadatos del estado de procesamiento de la sonoridad (y opcionalmente también otros metadatos) recibidos con las muestras, o bits de control (determinados por el decodificador a partir de metadatos del estado de procesamiento de la sonoridad y opcionalmente también otros metadatos) recibidos con las muestras. La unidad 22 de posprocesamiento está también opcionalmente configurada para mezclar el contenido de audio posprocesado para su reproducción por uno o más altavoces. Estos altavoces pueden ser realizados en cualquier variedad de diferentes dispositivos de audición o elementos de equipos de reproducción, tales como ordenadores, televisores, sistemas estéreo (hogar o cine), teléfonos móviles, y otros dispositivos de reproducción portátiles. Los altavoces pueden ser de cualquier tamaño y nivel de potencia adecuados, y pueden proporcionarse en forma de controladores independientes, cajas de altavoces, sistemas de sonido envolvente, barras de sonido, audífonos, auriculares, etc.

Algunas realizaciones proporcionan una cadena de procesamiento de audio mejorada en la que las unidades de procesamiento de audio (p. ej., codificadores, decodificadores, transcodificadores y unidades de pre- y posprocesamiento) adaptan su procesamiento respectivo para que se aplique a datos de audio según un estado contemporáneo de los datos de los medios indicados por los metadatos del estado de procesamiento de la sonoridad recibidos respectivamente por las unidades de procesamiento de audio. La entrada 11 de datos de audio a cualquier unidad de procesamiento de audio del sistema 100 (p. ej., el codificador o transcodificador de la Figura 1) puede incluir metadatos del estado de procesamiento de la sonoridad (y opcionalmente también otros metadatos) así como datos de audio (p. ej., datos de audio codificados). Estos metadatos pueden haber sido incluidos en el audio de entrada por otro elemento u otra fuente según algunas realizaciones. La unidad de procesamiento que recibe el audio de entrada (con metadatos) puede estar configurada para realizar al menos una operación sobre los metadatos (p. ej., validación) o en respuesta a los metadatos (p. ej., procesamiento adaptativo del audio de entrada), y opcionalmente también para incluir en su audio de salida los metadatos, una versión procesada de los metadatos, o bits de control determinados a partir de los metadatos.

Una realización de la unidad de procesamiento de audio (o procesador de audio) está configurada para realizar el procesamiento adaptativo de datos de audio basado en el estado de los datos de audio como se indica por los metadatos del estado de procesamiento de la sonoridad correspondientes a los datos de audio. En algunas realizaciones, el procesamiento adaptativo es (o incluye) el procesamiento de la sonoridad (si los metadatos indican que el procesamiento de la sonoridad, o procesamiento similar a la misma, no se ha realizado ya sobre los datos de audio, pero no es (y no incluye) el procesamiento de la sonoridad (si los metadatos indican que dicho procesamiento de la sonoridad o un procesamiento similar, ya se ha realizado sobre los datos de audio). En algunas realizaciones, el procesamiento adaptativo es o incluye la validación de los metadatos (p. ej., realizada en una subunidad de validación de metadatos) para asegurar que la unidad de procesamiento de audio realiza otro procesamiento adaptativo de los datos de audio basado en el estado de los datos de audio como se indica mediante los metadatos del estado de procesamiento de la sonoridad. En algunas realizaciones, la validación determina la fiabilidad de los metadatos del estado de procesamiento de la sonoridad asociados con (p. ej., incluidos en un flujo de bits con) los datos de audio. Por ejemplo, si los metadatos son validados de ser fiables, entonces los resultados de un tipo de procesamiento de audio realizado previamente pueden ser reutilizados y se puede evitar la realización adicional del mismo tipo de procesamiento de audio. Por otra parte, si se comprueba que los metadatos han sido manipulados (o no fiables), entonces el tipo de procesamiento de los medios supuestamente realizado previamente (como se indica por los metadatos no fiables) puede repetirse por la unidad de procesamiento de audio, y/u otro procesamiento puede ser realizado por la unidad de procesamiento de audio sobre los metadatos y/o los datos de audio. La unidad de procesamiento de audio también puede configurarse para señalar otras unidades de procesamiento de audio aguas abajo en una cadena de procesamiento de medios mejorada de manera que los metadatos del estado de procesamiento de la sonoridad (p. ej., presentes en un flujo de bits de medios) son válidos si la unidad determina que los metadatos del estado de procesamiento son válidos (por ejemplo, basado en una coincidencia de un valor criptográfico extraído y un valor criptográfico de referencia).

Para la realización de la Fig. 1, el componente 12 de preprocesamiento puede ser parte del codificador 14, y el componente 22 de posprocesamiento puede ser parte del decodificador 22. Alternativamente, el componente 12 de preprocesamiento puede ser realizado en un componente funcional que está separado del codificador 14. De manera similar, el componente 22 de posprocesamiento puede estar realizado en un componente funcional que está separado del decodificador 20.

La Fig. 2 es un diagrama de bloques de un codificador 100 que se puede utilizar junto con el sistema 10 de la Fig. 1. Cualquiera de los componentes o elementos del codificador 100 puede ser aplicado como uno o más procesos y/o uno o más circuitos (p. ej., ASIC, FPGA u otros circuitos integrados), en hardware, software o una combinación de hardware y software. El codificador 100 comprende la memoria intermedia 110 de la trama, el reconocedor 111, el decodificador 101, el validador 102 del estado de audio, la etapa 103 de procesamiento de la sonoridad, la etapa 104 de selección del flujo de audio, el codificador 105, la etapa 107 de rellenador/formateador, la etapa 106 de generación de metadatos, el subsistema 108 de medición de la sonoridad del diálogo y de la memoria intermedia 109 de la trama, conectados como se muestra. Opcionalmente también, el codificador 100 incluye otros elementos de procesamiento (no mostrados). El codificador 100 (que es un transcodificador) está configurado para transformar un flujo de bits de audio de entrada (que, por ejemplo, puede ser uno de un flujo de bits AC-3, un flujo de bits E-AC-3 o un flujo de bits Dolby E) en un flujo de bits de audio de salida codificado (que, por ejemplo, puede ser otro de un flujo de bits AC-3, un flujo de bits E-AC-3 o un flujo de bits Dolby E) que incluye realizar un procesamiento de la sonoridad adaptativo y automatizado utilizando los metadatos del estado de procesamiento de la sonoridad incluidos en el flujo de bits de entrada. Por ejemplo, el codificador 100 puede ser configurado para transformar un flujo de bits Dolby E de entrada (un formato utilizado típicamente en instalaciones de producción y difusión pero no en dispositivos para el consumidor que reciben programas de audio que han sido difundidos por él) en un flujo de bits de audio de salida codificado (adecuado para la radiodifusión por dispositivos del consumidor) en formato AC-3 o E-AC-3.

El sistema de la Fig. 2 incluye también el subsistema 150 de envío de audio codificado (que almacena y/o envía los flujos de bits codificados procedentes del codificador 100) y el decodificador 152. Un flujo de bits de audio codificado procedente del codificador 100 puede ser almacenado por el subsistema 150 (p. ej., en forma de un DVD o disco BluRay), o transmitido por el subsistema 150 (que puede aplicar un enlace o red de transmisión), o puede ser almacenado y transmitido por el subsistema 150. El decodificador 152 está configurado para decodificar un flujo de bits de audio codificado (generado por el codificador 100) que recibe a través del subsistema 150, incluyendo extraer metadatos del estado de procesamiento de la sonoridad (LPSM) de cada trama del flujo de bits, y generar datos de audio decodificados. En una realización, el decodificador 152 está configurado para realizar un procesamiento de la sonoridad adaptativo sobre los datos de audio decodificados usando LPSM, y/o para enviar los datos de audio decodificados y LPSM a un posprocesador configurado para realizar un procesamiento de la sonoridad adaptativo sobre los datos de audio decodificados utilizando el LPSM. Opcionalmente, el decodificador 152 incluye una memoria intermedia, que almacena (p. ej., de manera no transitoria) el flujo de bits de audio codificado recibido del subsistema 150.

Varias aplicaciones del codificador 100 y del decodificador 152 están configuradas para realizar diferentes realizaciones descritas en el presente documento. La memoria intermedia 110 de la trama es una memoria intermedia acoplada para recibir un flujo de bits de audio de entrada codificado. En funcionamiento, la memoria intermedia 110 almacena (p. ej., de manera no transitoria) al menos una trama del flujo de bits de audio codificado, y una secuencia de las tramas del flujo de bits de audio codificado se presenta desde la memoria intermedia 110 al reconocedor 111. El reconocedor 111 está acoplado y configurado para extraer metadatos del estado de procesamiento de la sonoridad (LPSM) y otros metadatos de cada trama del audio de entrada codificado, para presentar al menos el LPSM al validador 102 del estado de audio, a la etapa 103 de procesamiento de la sonoridad, a la etapa 106 y al subsistema 108, para extraer datos de audio desde el audio de entrada codificado, presentar al decodificador 101. El decodificador 101 del codificador 100 está configurado para decodificar los datos de audio para generar datos de audio decodificados, y para presentar los datos de audio decodificados a la etapa 103 de procesamiento de la sonoridad, a la etapa 104 de selección de flujo, al subsistema 108, y opcionalmente también al validador 102 del estado.

El validador 102 de estado está configurado para autenticar y validar LPSM (y opcionalmente otros metadatos) presentados a él. En algunas realizaciones, LPSM es (o está incluido en) un bloque de datos que se ha incluido en el flujo de bits de entrada (p. ej., según una realización de la presente invención). El bloque puede comprender un trozo criptográfico (un código de autenticación de mensaje basado en trozos o "HMAC") para procesar LPSM (y opcionalmente también otros metadatos) y/o los datos de audio subyacentes (proporcionados desde el decodificador 101 al validador 102). El bloque de datos puede estar firmado digitalmente en estas realizaciones, de manera que una unidad de procesamiento de audio aguas abajo pueda autenticar y validar de forma relativamente fácil los metadatos del estado de procesamiento.

Por ejemplo, el HMAC se utiliza para generar un resumen, y el valor (o valores) de protección incluidos en el flujo de bits de la invención pueden incluir el resumen. El resumen se puede generar de la siguiente manera para una trama AC-3: (1) Después de que los datos AC-3 y LPSM son codificados, los bytes de datos de la trama (frame_data #1 y frame_data #2 concatenados) y los bytes de datos LPSM se utilizan como entrada para la HMAC de la función de troceado. Otros datos que pueden estar presentes dentro de un campo auxdata no se toman en consideración para

calcular el resumen. Otros datos de este tipo pueden ser bytes que no pertenecen ni a los datos de AC-3 ni a los datos de LPSM. Los bits de protección incluidos en LPSM pueden no ser considerados para calcular el resumen HMAC. (2) Después de que se calcula el resumen, se escribe en el flujo de bits en un campo reservado para los bits de protección. (3) El último paso de la generación de la trama AC-3 completa es el cálculo de la verificación CRC. Esto se escribe muy al final de la trama y se tienen en consideración todos los datos pertenecientes a esta trama, incluidos los bits LPSM.

Otros procedimientos criptográficos que incluyen pero no se limitan a ninguno de uno o más procedimientos criptográficos no HMAC pueden ser usados para la validación de LPSM (p. ej., en el validador 102) para asegurar la transmisión y recepción seguras de LPSM y/o de los datos de audio subyacentes. Por ejemplo, puede realizarse la validación (usando un procedimiento criptográfico de este tipo) en cada unidad de procesamiento de audio que recibe una realización del flujo de bits de audio para determinar si los metadatos del estado de procesamiento de la sonoridad y los datos de audio correspondientes incluidos en el flujo de bits han experimentado (y/o han resultado de) el procesamiento específico de la sonoridad (como se indica por los metadatos) y no se han modificado después de la realización de tal procesamiento específico de la sonoridad.

El validador 102 del estado presenta datos de control a la etapa 104 de selección del flujo de audio, generador 106 de metadatos y al subsistema 108 de medición de la sonoridad del diálogo, para indicar los resultados de la operación de validación. En respuesta a los datos de control, la etapa 104 puede seleccionar (y pasar a través del codificador 105) bien: (1) la salida procesada de forma adaptativa de la etapa 103 de procesamiento de la sonoridad (p. ej., cuando los LPSM indican que los datos de audio procedentes del decodificador 101 no han experimentado un tipo específico de procesamiento de la sonoridad, y los bits de control procedentes del validador 102 indican que los LPSM son válidos); o bien (2) los datos de audio procedentes del decodificador 101 (p. ej., cuando los LPSM indican que los datos de audio procedentes del decodificador 101 ya han experimentado el tipo específico de procesamiento de la sonoridad que se realizaría por la etapa 103 y los bits de control del validador 102 indican que los LPSM son válidos). En una realización, la etapa 103 de procesamiento de la sonoridad corrige la sonoridad objetivo y al rango de sonoridad especificados.

La etapa 103 del codificador 100 está configurada para realizar un procesamiento de la sonoridad adaptativo sobre los datos de audio decodificados procedentes del decodificador 101, basado en una o más características de datos de audio indicadas por los LPSM extraídos por el decodificador 101. La etapa 103 puede ser un procesador de control de la sonoridad y del rango dinámico en tiempo real del dominio de transformación adaptativa. La etapa 103 puede recibir entrada del usuario (p. ej., valores de sonoridad /rango dinámico objetivos del usuario o valores dialnorm), u otra entrada de metadatos (p. ej., uno o más tipos de datos de terceros, información de seguimiento, identificadores, información de propiedad o estándar, datos de anotación de usuario, datos de preferencia del usuario, etc.) y/u otra entrada (p. ej., de un proceso de huellas digitales) y utilizar dicha entrada para procesar los datos de audio decodificados procedentes del decodificador 101.

El subsistema 108 de medición de la sonoridad del diálogo puede operar para determinar la sonoridad de los segmentos del audio decodificado (del decodificador 101) que son indicativos del diálogo (u otra voz), p. ej., utilizando los LPSM (y/u otros metadatos) extraídos por el decodificador 101, cuando los bits de control del validador 102 indican que los LPSM no son válidos. El funcionamiento del subsistema 108 de medición de la sonoridad del diálogo puede ser deshabilitado cuando los LPSM indican sonoridad determinada previamente de segmentos del diálogo (u otra voz) del audio decodificado (desde el decodificador 101) cuando los bits de control del validador 102 indican que los LPSM son válidos.

Existen herramientas útiles (p. ej., el medidor de sonoridad Dolby LM100) para medir el nivel de diálogo en contenido de audio conveniente y fácilmente. Algunas realizaciones de la APU (p. ej., la etapa 108 del codificador 100) se aplican para incluir (o para realizar las funciones de) tal herramienta para medir la sonoridad media de diálogo del contenido de audio de un flujo de bits de audio (p. ej., un flujo de bits AC-3 decodificado presentado a la etapa 108 desde el decodificador 101 del codificador 100). Si la etapa 108 se aplica para medir la sonoridad media real del diálogo de los datos de audio, la medición puede incluir una etapa de aislar segmentos del contenido de audio que contienen predominantemente voz. Los segmentos de audio que predominantemente son voz se procesan después según un algoritmo de medición de la sonoridad. Para datos de audio decodificados desde un flujo de bits AC-3, este algoritmo puede ser una medida de la sonoridad ponderada en K estándar (según la norma internacional ITU-R BS.1770). Alternativamente, pueden usarse otras medidas de la sonoridad (p. ej., aquellas basadas en modelos psicoacústicos de la sonoridad).

El aislamiento de segmentos de voz no es esencial para medir la sonoridad media del diálogo de los datos de audio. Sin embargo, mejora la precisión de la medida y proporciona resultados más satisfactorios desde la perspectiva de un oyente. Debido a que no todo el contenido de audio contiene diálogo (voz), la medida de la sonoridad de todo el contenido de audio puede proporcionar una aproximación suficiente del nivel de diálogo del audio, estuviera la voz presente.

El generador 106 de metadatos genera metadatos para ser incluidos mediante la etapa 107 en el flujo de bits codificado que saldrá del codificador 100. El generador 106 de metadatos puede pasar a través de la etapa 107 los LPSM (y/u otros metadatos) extraídos por el codificador 101 (p. ej., cuando los bits de control del validador 102

indican que los LPSM y/u otros metadatos son válidos), o generar nuevos LPSM (y/u otros metadatos) y presenta los nuevos metadatos a la etapa 107 (p. ej., cuando los bits de control del validador 102 indican que los LPSM y/u otros metadatos extraídos por el decodificador 101 no son válidos o pueden presentar a la etapa 107 una combinación de metadatos extraídos por el decodificador 101 y los metadatos recién generados. El generador 106 de metadatos puede incluir datos de la sonoridad generados por el subsistema 108, y al menos un valor indicativo del tipo de procesamiento de la sonoridad realizado por el subsistema 108, en los LPSM presenta a la etapa 107 para su inclusión en el flujo de bits codificado que saldrá del codificador 100. El generador 106 de metadatos puede generar bits de protección (que pueden consistir en o incluir un código de autenticación de mensaje basado en trozos o "HMAC") útiles para que al menos uno de entre descifrado, autenticación o validación de LPSM (y opcionalmente también otros metadatos) sean incluidos en el flujo de bits codificado y/o en los datos de audio subyacentes que se incluirán en el flujo de bits codificado. El generador 106 de metadatos puede proporcionar tales bits de protección a la etapa 107 para su inclusión en el flujo de bits codificado.

En una realización, el subsistema 108 de medida de la sonoridad del diálogo procesa los datos de audio procedentes del decodificador 101 para generar en respuesta a ello los valores de sonoridad (p. ej., valores de la sonoridad de diálogo con puerta y sin puerta) y valores de rango dinámico. En respuesta a estos valores, el generador 106 de metadatos puede generar metadatos del estado de procesamiento de la sonoridad (LPSM) para su inclusión (por el rellenedor/formateador 107) en el flujo de bits codificado que son procedentes del codificador 100. En una realización, la sonoridad puede calcularse basada en técnicas especificadas por las normas ITU-R BS.1770-1 e ITU-R BS.1770-2, u otras normas similares de medición de la sonoridad. La sonoridad con puerta puede ser sonoridad con puerta al diálogo o sonoridad con puerta relativa, o una combinación de estos tipos de sonoridad con puerta, y el sistema puede emplear bloques con puerta adecuados dependiendo de los requisitos de la aplicación y las restricciones del sistema.

De forma adicional, opcional o alternativa, los subsistemas 106 y/o 108 del codificador 100 pueden realizar un análisis adicional de los datos de audio para generar metadatos indicativos de al menos una característica de los datos de audio para su inclusión en el flujo de bits codificado que se va a emitir desde la etapa 107. El codificador 105 codifica (p. ej., realizando la compresión sobre el mismo) los datos de audio procedentes de la etapa 104 de selección y presenta el audio codificado a la etapa 107 para su inclusión en el flujo de bits codificado procedente de la etapa 107.

La etapa 107 multiplexa el audio codificado del codificador 105 y los metadatos (incluidos los LPSM) del generador 106 para generar que el flujo de bits codificado sea emitido desde la etapa 107, de manera que el flujo de bits codificado tenga el formato especificado por una realización. El memoria intermedia 109 de la trama es una memoria intermedia que almacena (p. ej., de manera no transitoria) al menos una trama del flujo de bits de audio codificado procedente de la etapa 107 y una secuencia de las tramas del flujo de bits de audio codificado se presenta entonces desde la memoria intermedia 109 como salida del codificador 100 hacia el sistema 150 de envío.

Los LPSM generados por el generador 106 de metadatos e incluido en el flujo de bits codificado por la etapa 107 es indicativo del estado de procesamiento de la sonoridad de los datos de audio correspondientes (p. ej., qué tipo (o tipos) de procesamiento de la sonoridad se han realizado sobre los datos de audio) y sonoridad (p. ej., sonoridad medida del diálogo, sonoridad con puerta y/o sin puerta, y/o rango dinámico) de los datos de audio correspondientes. En el presente documento, el "umbral de puerta" de la sonoridad y/o mediciones de nivel realizadas sobre datos de audio se refieren a un nivel específico o umbral de sonoridad en donde el valor (o valores) calculado que excede el umbral se incluye en la medición final (p. ej., ignorando valores de sonoridad a corto plazo inferiores a -60 dBFS en los valores finales medidos). El umbral de puerta en un valor absoluto se refiere a un nivel o sonoridad fijo, mientras que el umbral de puerta en un valor relativo se refiere a un valor que depende de un valor de medición "sin puerta" actual.

En algunas aplicaciones del codificador 100, el flujo de bits codificado guardado en la memoria 109 (y que sale hacia el sistema de envío 150) es un flujo de bits AC-3 o un flujo de bits E-AC-3, y comprende segmentos de datos de audio (p. ej., los segmentos AB0-AB5 de la trama mostrada en la Fig. 4) y segmentos de metadatos, donde los segmentos de datos de audio son indicativos de datos de audio, y cada uno de al menos algunos de los segmentos de metadatos incluye metadatos del estado de procesamiento de la sonoridad (LPSM). La etapa 107 inserta LPSM en el flujo de bits en el siguiente formato. Cada uno de los segmentos de metadatos que incluye LPSM está incluido en un campo "addbsi" del segmento de Información de Flujo de Bits ("BSI") de una trama del flujo de bits, o en un campo auxdata (p. ej., el segmento AUX mostrado en la Fig. 4) en el extremo de una trama del flujo de bits.

Una trama del flujo de bits puede incluir uno o dos segmentos de metadatos, cada uno de los cuales incluye LPSM, y si la trama incluye dos segmentos de metadatos, uno está presente en el campo addbsi de la trama y el otro en el campo AUX de la trama. Cada segmento de metadatos que incluye LPSM incluye un segmento de carga útil LPSM (o contenedor) que tiene el siguiente formato: una cabecera (p. ej., que incluya una palabra de sincronización que identifica el inicio de la carga útil LPSM, seguido por al menos un valor de identificación, p. ej., la versión del formato LPSM, longitud, período, cuenta y valores de asociación de flujo adicional indicados en la Tabla 2 de más adelante); y después de la cabecera, al menos un valor de indicación de diálogo (p. ej., parámetro "Canal (o canales) de diálogo" de la Tabla 2) que indica si los datos de audio correspondientes indican diálogo o no indican diálogo (p. ej., qué canales de datos de audio correspondientes indican diálogo); al menos un valor de cumplimiento de la

regulación de la sonoridad (p. ej., parámetro "Tipo de Regulación de la Sonoridad" de la Tabla 2) que indica si los datos de audio correspondientes cumplen con un conjunto indicado de regulaciones de la sonoridad; al menos un valor de procesamiento de la sonoridad (p. ej., uno o más de los parámetros "Indicador de corrección de la sonoridad con puerta al diálogo", "Tipo de corrección de la sonoridad" de la Tabla 2) que indica al menos un tipo de procesamiento de la sonoridad que se ha realizado sobre los datos de audio correspondientes; y al menos un valor de la sonoridad (p. ej., uno o más de los parámetros "Sonoridad con puerta relativa ITU", "Sonoridad con puerta de voz ITU", "Sonoridad 3s a corto plazo ITU (EBU 3341)", y "Pico real" de la Tabla 2) que indica al menos una característica de la sonoridad (p. ej., sonoridad pico o media) de los datos de audio correspondientes.

En algunas aplicaciones, cada uno de los segmentos de metadatos insertados por la etapa 107 en un campo "addbsi" o en un campo auxdata de una trama de flujo de bits tiene el siguiente formato: una cabecera principal (p. ej., que incluye una palabra de sincronización que identifica el inicio del segmento de metadatos, seguido por los valores de identificación, p. ej., la versión del elemento central, la longitud y el período, el recuento de elementos extendidos y los valores de asociación de flujos adicionales indicados en la Tabla 1 de más adelante); y después de la cabecera principal, al menos un valor de protección (p. ej., el resumen HMAC y los valores de Huella Digital de Audio de la Tabla 1) útil para al menos uno de descifrado, autenticación o validación de al menos uno de entre los metadatos del estado de procesamiento de la sonoridad o los datos de audio correspondientes); y también después de la cabecera principal, si el segmento de metadatos incluye LPSM, identificación ("ID") de la carga útil de los LPSM y los valores de tamaño de carga útil LPSM que identifican los siguientes metadatos como una carga útil de LPSM e indican el tamaño de la carga útil LPSM.

El segmento de la carga útil (o contenedor) de LPSM (p. ej., que tiene el formato especificado anteriormente) sigue los valores de ID de carga útil LPSM y de tamaño de carga útil LPSM.

En algunas realizaciones, cada uno de los segmentos de metadatos en el campo auxdata (o en el campo "addbsi") de una trama tiene tres niveles de estructura: una estructura de alto nivel, que incluye un indicador que indica si el campo auxdata (o el addbsi) incluye metadatos, al menos un valor ID que indique qué tipo (o tipos) de metadatos están presentes, y opcionalmente también un valor que indique cuántos bits de metadatos (p. ej., de cada tipo) están presentes (si hay metadatos presentes). Un tipo de metadatos que podrían estar presentes son LPSM, y otro tipo de metadatos que podrían estar presentes son metadatos de búsqueda de medios (p. ej., metadatos de Búsqueda de Medios Nielsen); una estructura de nivel intermedio, que comprende un elemento central para cada tipo identificado de metadatos (p. ej., cabecera principal, valores de protección y valores de ID de carga útil de LPSM y de tamaño de carga útil de LPSM, como se mencionó anteriormente, para cada tipo identificado de metadatos); y una estructura de bajo nivel, que comprende cada carga útil para un elemento central (p. ej., una carga útil de LPSM, si es identificado por el elemento central de que está presente y/o una carga útil de metadatos de otro tipo, si es identificado por el elemento central de que está presente).

Los valores de datos en una estructura de tres niveles de este tipo pueden estar anidados. Por ejemplo, el valor (o valores) de protección de una carga útil de LPSM y/u otra carga útil de metadatos identificada por un elemento central puede incluirse después de cada carga útil identificada por el elemento central (y, por lo tanto, después de la cabecera principal del elemento central). En un ejemplo, una cabecera principal podría identificar una carga útil de LPSM y otra carga útil de metadatos, los valores de ID de la carga útil y del tamaño de la carga útil para la primera carga útil (p. ej., la carga útil de los LPSM) podría seguir la cabecera principal, la propia primera carga útil podría seguir los valores de ID y de tamaño, el valor de ID de la carga útil y el tamaño de la carga útil para la segunda carga útil podría seguir la primera carga útil, la propia segunda carga útil podría seguir estos valores de ID y de tamaño, y los bits de protección para ambas cargas útiles (o para los valores del elemento central y ambas cargas útiles) podrían seguir la última carga útil.

En algunas realizaciones, si el decodificador 101 recibe un flujo de bits de audio generados según una realización de la invención con un trozo criptográfico, el decodificador está configurado para reconocer y recuperar el trozo criptográfico de un bloque de datos determinado a partir del flujo de bits, comprendiendo dicho bloque metadatos del estado de procesamiento de la sonoridad (LPSM). El validador 102 puede utilizar el trozo criptográfico para validar el flujo de bits recibidos y/o los metadatos asociados. Por ejemplo, el validador 102 considera que el LPSM es válido basado en una coincidencia entre un trozo criptográfico de referencia y el trozo criptográfico recuperado del bloque de datos, entonces puede desactivar la operación del procesador 103 sobre los datos de audio correspondientes y hace que la etapa 104 de selección pase a través (sin cambios) los datos de audio. Adicionalmente, opcionalmente o alternativamente, se pueden usar otros tipos de técnicas criptográficas en lugar de un procedimiento basado en un trozo criptográfico.

El codificador 100 de la Fig. 2 puede determinar (en respuesta a LPSM extraído por el decodificador 101) que una unidad de pos-/pre-procesamiento ha realizado un tipo de procesamiento de la sonoridad sobre los datos de audio a codificar (en los elementos 105, 106 y 107) y, en consecuencia, puede crear (en el generador 106) metadatos del estado de procesamiento de la sonoridad que incluyen los parámetros específicos utilizados en y/o derivados del procesamiento de la sonoridad realizado previamente. En algunas aplicaciones, el codificador 100 puede crear (e incluir en el flujo de bits codificado procedentes de él) metadatos del estado de procesamiento indicativos del historial de procesamiento en el contenido de audio siempre y cuando el codificador tenga en cuenta los tipos de procesamiento que se han realizado sobre el contenido de audio.

La Fig. 3 es un diagrama de bloques de un decodificador que se puede utilizar junto con el sistema 10 de la Fig. 1. Cualquiera de los componentes o elementos del decodificador 200 y del posprocesador 300 puede ser aplicado como uno o más procesos y/o uno o más circuitos (p. ej., los ASIC, FPGA u otros circuitos integrados), en hardware, software o una combinación de hardware y software. El decodificador 200 comprende la memoria intermedia 201 de la trama, el reconocedor 205, el decodificador 202 de audio, la etapa de validación (validador) 203 del estado de audio, y la etapa 204 de generación de bits de control, conectados como se muestra. El decodificador 200 puede incluir otros elementos de procesamiento (no mostrados). La memoria intermedia 201 de la trama (una memoria tampón) guarda (p. ej., de manera no transitoria) al menos una trama del flujo de bits de audio codificado recibido por el decodificador 200. Una secuencia de las tramas del flujo de bits de audio codificado se presenta desde la memoria intermedia 201 al reconocedor 205. El reconocedor 205 está acoplado y configurado para extraer metadatos del estado de procesamiento de la sonoridad (LPSM) y otros metadatos de cada trama del audio de entrada codificado, para presentar al menos el LPSM al validador 203 del estado de audio y a la etapa 204, para presentar los LPSM como salida (p. ej., al post-procesador 300), para extraer datos de audio del audio de entrada codificado y presenta los datos de audio extraídos al decodificador 202. El flujo de bits de audio codificado de entrada al decodificador 200 puede ser uno de un flujo de bits AC-3, un flujo de bits E-AC-3, o un flujo de bits Dolby E.

El sistema de la Fig. 3, incluye también el posprocesador 300. El posprocesador 300 comprende la memoria intermedia 301 de la trama y otros elementos de procesamiento (no mostrados) que incluyen al menos un elemento de procesamiento acoplado a la memoria intermedia 301. La memoria intermedia 301 de la trama guarda (p. ej., de manera no transitoria) al menos una trama del flujo de bits de audio decodificado recibido por el posprocesador 300 del decodificador 200. Los elementos de procesamiento del posprocesador 300 están acoplados y configurados para recibir y procesar de forma adaptativa una secuencia de las tramas del flujo de bits de audio decodificados procedente del memoria intermedia 301, utilizando metadatos (incluidos los valores de los LPSM) procedentes del decodificador 202 y/o bits de control procedentes de la etapa 204 del decodificador 200. En una realización, el posprocesador 300 está configurado para realizar un procesamiento de la sonoridad adaptativo sobre los datos de audio decodificados que utilizan los valores de los LPSM (p. ej., basado en el estado de procesamiento de la sonoridad, y/o una o más características de datos de audio, indicadas por los LPSM). Varias aplicaciones del decodificador 200 y del posprocesador 300 están configuradas para realizar diferentes realizaciones de los procedimientos según las realizaciones descritas en el presente documento.

El decodificador 202 de audio del decodificador 200 está configurado para decodificar los datos de audio extraídos por el reconocedor 205 para generar datos de audio decodificados y presentar los datos de audio decodificados como salida (p. ej., hacia el posprocesador 300). El validador 203 del estado está configurado para autenticar y validar los LPSM (y opcionalmente otros metadatos) presentados al mismo. En algunas realizaciones, LPSM es (o está incluido en) un bloque de datos que se ha incluido en el flujo de bits de entrada (p. ej., según una realización de la presente invención). El bloque puede comprender un trozo criptográfico (un código de autenticación de mensajes basado en trozos o "HMAC") para procesar los LPSM (y opcionalmente también otros metadatos) y/o los datos de audio subyacentes (proporcionados desde el reconocedor 205 y/o el decodificador 202 hasta el validador 203). El bloque de datos se puede firmar digitalmente en estas realizaciones, de modo que una unidad de procesamiento de audio aguas abajo pueda, de una manera relativamente fácil, autenticar y validar los metadatos del estado de procesamiento.

Otros procedimientos criptográficos que incluyen pero no se limitan a cualquiera de uno o más procedimientos criptográficos no HMAC pueden ser utilizados para la validación de los LPSM (p. ej., en el validador 203) para garantizar la transmisión y recepción seguras de los LPSM y/o de los datos de audio subyacentes. Por ejemplo, puede realizarse la validación (usando un procedimiento criptográfico de este tipo) en cada unidad de procesamiento de audio que recibe una realización del flujo de bits de audio de la invención para determinar si los metadatos del estado de procesamiento de la sonoridad y los datos de audio correspondientes incluidos en el flujo de bits han experimentado (y/o han resultado de) un procesamiento específico de la sonoridad (como se indica por los metadatos) y no han sido modificados después de la realización de dicho procesamiento de sonoridad específico.

El validador 203 del estado presenta datos de control para controlar el generador 204 de bits y/o presenta los datos de control como salida (p. ej., hacia el posprocesador 300), para indicar los resultados de la operación de validación. En respuesta a los datos de control (y opcionalmente también a otros metadatos extraídos del flujo de bits de entrada), la etapa 204 puede generar (y presenta al posprocesador 300) bien: bits de control que indican que los datos de audio decodificados procedentes del decodificador 202 han experimentado un tipo específico de procesamiento de la sonoridad (cuando los LPSM indican que los datos de audio procedentes del decodificador 202 han experimentado el tipo específico de procesamiento de la sonoridad, y los bits de control del validador 203 indican que los LPSM son válidos); o bits de control que indican que los datos de audio decodificados procedentes del decodificador 202 deben experimentar un tipo específico de procesamiento de la sonoridad (p. ej., cuando los LPSM indican que los datos de audio procedentes del decodificador 202 no han experimentado el tipo específico de procesamiento de la sonoridad, o cuando los LPSM indican que los datos de audio procedentes del decodificador 202 han experimentado el tipo específico de procesamiento de la sonoridad pero los bits de control del validador 203 indican que los LPSM no son válidos).

Alternativamente, el decodificador 200 presenta los LPSM (y cualesquiera otros metadatos) extraídos por el

decodificador 202 desde el flujo de bits de entrada hacia el posprocesador 300, y el posprocesador 300 realiza el procesamiento de la sonoridad sobre los datos de audio decodificados utilizando el LPSM, o realiza la validación del LPSM y después realiza el procesamiento de la sonoridad sobre los datos de audio decodificados utilizando los LPSM si la validación indica que los LPSM son válidos.

- 5 En algunas realizaciones, si el decodificador 201 recibe un flujo de bits de audio generados según una realización de la invención con un trozo criptográfico, el decodificador está configurado para reconocer y recuperar el trozo criptográfico de un bloque de datos determinado del flujo de bits, comprendiendo dicho bloque los metadatos del estado de procesamiento de la sonoridad (LPSM). El validador 203 puede usar el trozo criptográfico para validar el flujo de bits recibidos y/o los metadatos asociados. Por ejemplo, si el validador 203 considera que el LPSM es válido
- 10 basado en una coincidencia entre un trozo criptográfico de referencia y el trozo criptográfico recuperado del bloque de datos, entonces puede señalar a una unidad de procesamiento de audio aguas abajo (p. ej., el posprocesador 300, que puede ser o incluir una unidad de nivelación de volumen) que pasen a través (sin cambios) los datos de audio del flujo de bits. De forma adicional, opcional o alternativa, se pueden usar otros tipos de técnicas criptográficas en lugar del procedimiento basado en un trozo criptográfico.
- 15 En algunas aplicaciones del decodificador 100, el flujo de bits codificado recibido (y guardado en la memoria 201) es un flujo de bits AC-3 o un flujo de bits E-AC-3, y comprende segmentos de datos de audio (p. ej., los segmentos AB0-AB5 de la trama mostrada en la Fig. 4) y segmentos de metadatos, donde los segmentos de datos de audio son indicativos de datos de audio, y cada uno de al menos algunos de los segmentos de metadatos incluye metadatos del estado de procesamiento de la sonoridad (LPSM). La etapa 202 del decodificador está configurada para extraer
- 20 del flujo de bits los LPSM que tienen el siguiente formato. Cada uno de los segmentos de metadatos que incluye LPSM está incluido en un campo "addbsi" del segmento de la Información del Flujo de Bits ("BSI") de una trama del flujo de bits, o en un campo auxdata (p. ej., el segmento AUX mostrado en la Fig. 4) en el extremo de una trama del flujo de bits. Una trama del flujo de bits puede incluir uno o dos segmentos de metadatos, cada uno de los cuales incluye LPSM, y si la trama incluye dos segmentos de metadatos, uno está presente en el campo addbsi de la trama y el otro en el campo AUX de la trama. Cada segmento de metadatos que incluye LPSM incluye un segmento de carga útil (o contenedor) de LPSM que tiene el siguiente formato: una cabecera (p. ej., incluyendo una palabra de sincronización que identifica el inicio de la carga útil LPSM, seguida de valores de identificación, p. ej., la versión de formato LPSM, longitud, período, recuento, y los valores de asociación de flujos adicionales indicados en la Tabla 2 de más adelante); y después de la cabecera, al menos un valor de indicación de diálogo (p. ej., parámetro "Canal (o canales) de diálogo" de la Tabla 2) indicando si el dato de audio correspondiente indica diálogo o no indica diálogo (p. ej., qué canales de datos de audio correspondientes indican diálogo); al menos un valor de cumplimiento de la regulación de la sonoridad (p. ej., parámetro "Tipo de regulación de la sonoridad" de la Tabla 2) indicando si el dato de audio correspondiente cumple con un conjunto indicado de regulaciones de la sonoridad; al menos un valor de procesamiento de la sonoridad (p. ej., uno o más de los parámetros "indicador de Corrección de la Sonoridad con puerta de diálogo", "Tipo de Corrección de la Sonoridad" de la Tabla 2) indicando al menos un tipo de procesamiento de la sonoridad que se ha realizado sobre los datos de audio correspondientes; y al menos un valor de la sonoridad (p. ej., uno o más de los parámetros "Sonoridad con puerta relativa ITU", "Sonoridad con puerta de voz ITU", "Sonoridad 3s a corto plazo ITU (EBU 3341)" y "Pico Real" de la Tabla 2) indicando al menos una característica de la sonoridad (p. ej., sonoridad de pico o media) de los datos de audio correspondientes.
- 30
- 35
- 40 En algunas aplicaciones, la etapa 202 decodificadora está configurada para extraer, desde el campo "addbsi" o un campo auxdata de una trama del flujo de bits, cada segmento de metadatos que tiene el siguiente formato: una cabecera principal (p. ej., incluyendo una palabra de sincronización que identifica el inicio del segmento de metadatos, seguido por al menos un valor de identificación, p. ej., la versión, longitud y período del elemento Central, el recuento de elementos extendidos y los valores de asociación de flujos adicionales indicados en la Tabla 1 de más adelante); y después de la cabecera principal, al menos un valor de protección (p. ej., el resumen HMAC y los valores de Huella Digital de Audio de la Tabla 1) útil para al menos uno de entre descifrado, autenticación o validación de al menos uno de entre los metadatos del estado de procesamiento de la sonoridad o los datos de audio correspondientes); y también después de la cabecera principal, si el segmento de metadatos incluye valores de LPSM, de identificación ("ID") de carga útil LPSM y de tamaño de carga útil de LPSM que identifican los siguientes metadatos como una carga útil de LPSM e indican el tamaño de la carga útil LPSM. El segmento (p. ej., que tiene el formato especificado anteriormente) de la carga útil (o contenedor) LPSM sigue los valores de ID de la carga útil de LPSM y de tamaño de la carga útil de LPSM.
- 45
- 50

De manera más general, el flujo de bits de audio codificado generado por una realización tiene una estructura que proporciona un mecanismo para etiquetar los elementos y sub-elementos de metadatos como centrales (obligatorio) o expandidos (elementos opcionales). Esto permite que la velocidad de datos del flujo de bits (que incluye sus metadatos) escale a través de numerosas aplicaciones. Los elementos centrales (obligatorios) de la sintaxis del flujo de bits deben ser también capaces de señalar que los elementos expandidos (opcionales) asociados con el contenido de audio están presentes (en la banda) y/o en una ubicación remota (fuera de la banda).

55

- 60 En alguna realización, se requiere que el elemento (o elementos) central esté presente en cada trama del flujo de bits. Algunos sub-elementos de elementos centrales son opcionales y pueden estar presentes en cualquier combinación. No se requiere que los elementos expandidos estén presentes en cada trama (para limitar la sobrecarga de la velocidad de los bits). Así, los elementos expandidos pueden estar presentes en algunas tramas y

no en otras. Algunos sub-elementos de un elemento expandido son opcionales y pueden estar presentes en cualquier combinación, mientras que algunos sub-elementos de un elemento ampliado pueden ser obligatorios (es decir, si el elemento ampliado está presente en una trama del flujo de bits).

5 En algunas realizaciones, un flujo de bits de audio codificado que comprende una secuencia de segmentos de datos de audio y segmentos de metadatos es generado (p. ej., mediante una unidad de procesamiento de audio que incorpora la invención). Los segmentos de datos de audio son indicativos de datos de audio, cada uno de al menos algunos de los segmentos de metadatos incluye metadatos del estado de procesamiento de la sonoridad (LPSM) y los segmentos de datos de audio son multiplexados por división de tiempo con los segmentos de metadatos. En algunas realizaciones de esta clase, cada uno de los segmentos de metadatos tiene un formato que se describirá en el presente documento. En un formato, el flujo de bits codificado es un flujo de bits AC-3 o un flujo de bits E-AC-3, y cada uno de los segmentos de metadatos que incluye LPSM está incluido (p. ej., por la etapa 107 del codificador 100) como información adicional del flujo de bits en el campo "addbsi" (mostrado en la Fig. 6) del segmento de la Información de Flujo de Bits ("BSI") de una trama del flujo de bits, o en un campo auxdata de una trama del flujo de bits. Cada una de las tramas incluye un elemento central en el campo addbsi de la trama que tiene el formato mostrado en la Tabla 1 de la Fig. 8.

10 En un formato, cada uno de los campos addbsi (o auxdata) que contiene LPSM contiene una cabecera principal (y opcionalmente también elementos centrales adicionales), y después de la cabecera principal (o la cabecera principal y otros elementos centrales), los siguientes valores LPSM (parámetros): un ID de carga útil (identificando los metadatos como LPSM) según los valores de los elementos centrales (p. ej., como se especifica en la Tabla 1); un tamaño de carga útil (que indica el tamaño de la carga útil de LPSM) según el ID de carga útil; y datos LPSM (según el valor del ID de carga útil y del tamaño de la carga útil) que tienen el formato indicado en la Tabla 2 de la Fig. 9.

15 En un segundo formato de un flujo de bits codificado, el flujo de bits es un flujo de bits AC-3 o un flujo de bits E-AC-3, y cada uno de los segmentos de metadatos que incluye LPSM está incluido (p. ej., mediante la etapa 107 del codificador 100) bien en: un campo "addbsi" (mostrado en la Fig. 6) del segmento de Información de Flujo de Bits ("BSI") de una trama del flujo de bits; o en un campo auxdata (p. ej., el segmento AUX mostrado en la Fig. 4) en el extremo de una trama del flujo de bits. Una trama puede incluir uno o dos segmentos de metadatos, cada uno de los cuales incluye LPSM, y si la trama incluye dos segmentos de metadatos, uno está presente en el campo addbsi de la trama y el otro en el campo AUX de la trama. Cada segmento de metadatos que incluye LPSM tiene el formato especificado anteriormente con referencia a las Tablas 1 y 2 anteriores (es decir, incluye los elementos centrales especificados en la Tabla 1, seguidos por los ID de carga útil (identificando los metadatos como LPSM) y los valores de tamaño de carga útil especificados anteriormente, seguido de la carga útil (los datos LPSM que tienen el formato indicado en la Tabla 2).

20 En otro, el flujo de bits codificado es un flujo de bits Dolby E, y cada uno de los segmentos de metadatos que incluye LPSM está en las primeras N ubicaciones de muestra del intervalo de banda de protección Dolby E. Un flujo de bits Dolby E que incluye un segmento de metadatos de este tipo, que incluye LPSM, p. ej., incluye un valor indicativo de longitud de carga útil LPSM señalado en la palabra Pd del preámbulo SMPTE 337M (la tasa de repetición de palabra Pa SMPTE 337M puede permanecer idéntica a la velocidad de fotogramas asociada).

25 En un formato en donde el flujo de bits codificado es un flujo de bits E-AC-3, cada uno de los segmentos de metadatos que incluye LPSM está incluido (p. ej., mediante la etapa 107 del codificador 100) como información de flujo de bits adicional en el campo "addbsi" del segmento Información de Flujo de Bits ("BSI") de una trama del flujo de bits. Los aspectos adicionales de la codificación de un flujo de bits E-AC-3 con LPSM en este formato se describen como sigue: (1) durante la generación de un flujo de bits E-AC-3, mientras que el codificador E-AC-3 (que inserta los valores LPSM en el flujo de bits) está "activo", para cada trama (trama de sincronización) generada, el flujo de bits debe incluir un bloque de metadatos (incluido LPSM) llevado en el campo addbsi de la trama. Los bits necesarios para transportar el bloque de metadatos no deben incrementar la velocidad de bits del codificador (longitud de la trama); (2) cada bloque de metadatos (que contiene LPSM) debe contener la siguiente información: loudness_correction_type_flag: donde "1" indica que la sonoridad de los datos de audio correspondientes era corregida aguas arriba del codificador y "0" indica que la sonoridad era corregida por un corrector de la sonoridad incrustado en el codificador (p. ej., el procesador 103 de la sonoridad del codificador 100 de la Fig. 2); speech_channel: indica qué canal (o canales) fuente contiene la voz (durante los 0,5 segundos anteriores). Si no se detecta ninguna voz, se indicará como tal; speech_loudness: indica la sonoridad de voz integrada de cada canal de audio correspondiente que contiene la voz (durante los 0,5 segundos anteriores); ITU_loudness: indica la sonoridad ITU BS.1770-2 integrada de cada canal de audio correspondiente; gain: ganancia (o ganancias) compuesta de la sonoridad para la inversión en un decodificador (para demostrar reversibilidad).

30 Mientras que el codificador E-AC-3 (que inserta los valores de LPSM en el flujo de bits) está "activo" y está recibiendo una trama AC-3 con un indicador de "confianza", el controlador de la sonoridad en el codificador (p. ej., el procesador 103 de la sonoridad del codificador 100 de la Fig. 2) es derivado. Los valores de dialnorm y de DRC de una fuente "de confianza" se pasan a través (p. ej., por el generador 106 del codificador 100) al componente codificador E-AC-3 (p. ej., la etapa 107 del codificador 100). La generación del bloque LPSM continúa y el loudness_correction_type_flag se establece en "1". La secuencia de derivación del controlador de la sonoridad se sincroniza con el inicio de la trama AC-3 decodificada donde aparece el indicador "de confianza". La secuencia de

derivación del controlador de la sonoridad se aplica de la siguiente manera: el control `leveler_amount` es decrementado desde un valor de 9 a un valor de 0 sobre 10 períodos de bloque de audio (es decir, 53,3 ms) y el control `leveler_back_end_meter` se coloca en modo de derivación (esta operación debe dar como resultado una transición fluida). El término derivación "confiable" del nivelador implica que el valor de `dialnorm` del flujo de bits de la fuente también se reutiliza en la salida del codificador (p. ej., si el flujo de bits de la fuente "de confianza" tiene un valor de `dialnorm` de -30 entonces la salida del codificador debería utilizar -30 para el valor de `dialnorm` de salida).

Mientras que el codificador E-AC-3 (que inserta los valores de LPSM en el flujo de bits) está "activo" y está recibiendo una trama AC-3 sin el indicador "de confianza", el controlador de la sonoridad incrustado en el codificador (p. ej., el procesador 103 de la sonoridad del codificador 100 de la Fig. 2) está activo. La generación del bloque de LPSM continúa y el `loudness_correction_type_flag` se establece en "0". La secuencia de activación del controlador de la sonoridad se sincroniza con el inicio de la trama AC-3 decodificada donde desaparece el indicador "de confianza". La secuencia de activación del controlador de la sonoridad se aplica de la siguiente manera: el control `leveler_amount` se incrementa desde un valor de 0 a un valor de 9 a lo largo de 1 período del bloque de audio (es decir, 5,3 ms) y el control `leveler_back_end_meter` se coloca en modo "activo" (esta operación da como resultado una transición fluida e incluye un reinicio de integración `back_end_meter`); y durante la codificación, una interfaz gráfica de usuario (GUI) indicaba a un usuario los siguientes parámetros: "Programa de Audio de Entrada: [Confiable/No confiable]"- el estado de este parámetro se basa en la presencia del indicador "de confianza" dentro de la señal de entrada; y "Corrección de la sonoridad en tiempo real: [Activada/Desactivada]" - el estado de este parámetro se basa en si está activo este controlador de la sonoridad incorporado en el codificador.

Cuando se decodifica un flujo de bits AC-3 o E-AC-3 que tiene LPSM (en el formato descrito) incluido en el campo "addbsi" del segmento de la Información de Flujo de Bits ("BSI") de cada trama del flujo de bits, el decodificador reconoce los datos del bloque de LPSM (en el campo `addbsi`) y pasa los valores LPSM extraídos a una interfaz gráfica de usuario (GUI). El conjunto de valores de LPSM extraídos se actualiza en cada trama.

En otro formato más, el flujo de bits codificado es un flujo de bits AC-3 o un flujo de bits E-AC-3, y cada uno de los segmentos de metadatos que incluye LPSM está incluido (p. ej., por la etapa 107 del codificador 100) como información adicional del flujo de bits en el campo "addbsi" (mostrado en la Fig. 6) del segmento de Información de Flujo de Bits ("BSI") (o en el segmento Aux) de una trama del flujo de bits. En este formato (que es una variación del formato descrito anteriormente con referencia a las Tablas 1 y 2), cada uno de los campos `addbsi` (o Aux) que contiene LPSM contiene los siguientes valores LPSM: los elementos centrales especificados en la Tabla 1, seguidos por la ID de la carga útil (que identifica los metadatos como LPSM) y los valores del tamaño de la carga útil, seguido por la carga útil (datos de LPSM) que tiene el siguiente formato (similar a los elementos indicados en la Tabla 2 anterior): versión de la carga útil de LPSM: un campo de 2 bits que indica la versión de la carga útil de LPSM; `dialchan`: un campo de 3 bits que indica si los canales Izquierdo, Derecho y/o Central de los datos de audio correspondientes contienen diálogo hablado. La asignación de bits del campo `dialchan` puede ser la siguiente: el bit 0, que indica la presencia de diálogo en el canal izquierdo, se almacena en el bit más significativo del campo `dialchan`; y el bit 2, que indica la presencia de diálogo en el canal central, se almacena en el bit menos significativo del campo `dialchan`. Cada bit del campo `dialchan` se establece en "1" si el canal correspondiente contiene diálogo hablado durante los 0,5 segundos anteriores del programa; `loudregtyp`: un campo de 3 bits que indica qué regulación de la sonoridad estándar cumple con la sonoridad del programa. El ajuste del campo "loudregtyp" a "000" indica que LPSM no indica el cumplimiento de la regulación de la sonoridad. Por ejemplo, un valor de este campo (p. ej., 000) puede indicar que no se indica el cumplimiento con una norma de regulación de la sonoridad, otro valor de este campo (p. ej., 001) puede indicar que los datos de audio del programa cumplen con la norma ATSC A/85, y otro valor de este campo (p. ej., 010) puede indicar que los datos de audio del programa cumplen con la norma EBU R128. En el ejemplo, si el campo se establece en cualquier valor distinto de "000", los campos `loudcorrldialgat` y `loudcorrtyp` deben seguir en la carga útil; `loudcorrldialgat`: un campo de un bit que indica si se ha aplicado la corrección de sonoridad con puerta al diálogo. Si se ha corregido la sonoridad del programa con el umbral de puerta del diálogo, el valor del campo `loudcorrldialgat` se establece en "1". De lo contrario, se establece en "0"; `loudcorrtyp`: un campo de un bit que indica el tipo de corrección de la sonoridad aplicada al programa. Si la sonoridad del programa se ha corregido con un proceso de corrección de la sonoridad de análisis por anticipación de infinidad (basada en archivo), el valor del campo `loudcorrtyp` se establece en "0". Si se ha corregido la sonoridad del programa con una combinación de medición de la sonoridad en tiempo real y control del rango dinámico, el valor de este campo se establece en "1"; `loudrelgate`: un campo de un bit que indica si existen datos de sonoridad con puerta relativa (ITU). Si el campo `loudrelgate` se establece en "1", un campo `ituloudrelgat` de 7 bits debe seguir en la carga útil; `loudrelgat`: un campo de 7 bits que indica la sonoridad del programa con puerta relativa (ITU). Este campo indica la sonoridad integrada del programa de audio, medida según la ITU-R BS.1770-2 sin ningún ajuste de ganancia debido a la compresión `dialnorm` y de rango dinámico que se aplica. Los valores de 0 a 127 se interpretan como -58 LKFS a +5,5 LKFS, en pasos de 0,5 LKFS; `loudspchgate`: un campo de un bit que indica si existen datos de sonoridad de puerta de voz (ITU). Si el campo `loudspchgate` se establece en "1", un campo `loudspchgat` de 7 bits debe seguir en la carga útil; `loudspchgat`: un campo de 7 bits que indica la sonoridad del programa de puerta de voz. Este campo indica que la sonoridad integrada de todo el programa de audio correspondiente, medido según la fórmula (2) de ITU-R BS.1770-3 y sin ningún ajuste de ganancia debido a la compresión `dialnorm` y de rango dinámico que se aplica. Los valores de 0 a 127 se interpretan como -58 a +5,5 LKFS, en pasos de 0,5 LKFS; `loudstrm3se`: un campo de un bit que indica si existen datos de sonoridad a corto plazo (3 segundos). Si el campo se

- establece en "1", un campo loudstrm3s de 7 bits debe seguir en la carga útil; loudstrm3s: un campo de 7 bits que indica la sonoridad sin puerta de los 3 segundos anteriores del programa de audio correspondiente, medida según la ITU-R BS.1771-1 y sin ningún ajuste de ganancia debido al dialnorm y a la compresión del rango dinámico que se aplica. Los valores de 0 a 256 se interpretan como -116 LKFS a +11,5 LKFS, en pasos de 0,5 LKFS; truepk: un campo de un bit que indica si existen datos de sonoridad de pico real. Si el campo truepk se ajusta en "1", un campo truepk de 8 bits debería seguir en la carga útil; y truepk: un campo de 8 bits que indica el valor de la muestra de pico real del programa, medido según el Anexo 2 de la ITU-R BS.1770-3 y sin ningún ajuste de ganancia debido al dialnorm y a la compresión del rango dinámico que se aplica. Los valores de 0 a 256 se interpretan como -116 LKFS a +11,5 LKFS en 0,5 pasos de LKFS.
- 5
- 10 En algunas realizaciones, el elemento central de un segmento de metadatos en un campo auxdata (o campo "addbsi") de una trama de un flujo de bits AC-3 o de un flujo de bits E-AC-3 comprende una cabecera principal (que opcionalmente incluye valores de identificación, p. ej., versión de elemento central), y después la cabecera principal: valores indicativos de si los datos de la huella digital están (u otros valores de protección están) incluidos en los metadatos del segmento de metadatos, valores indicativos de si existen datos externos (relativos a los datos de audio correspondientes a los metadatos del segmento de metadatos), el ID de la carga útil y los valores de tamaño de la carga útil para cada tipo de metadatos (p. ej., LPSM y/o metadatos de un tipo distinto de LPSM) identificados por el elemento central y los valores de protección para al menos un tipo de metadatos identificados por el elemento central. La carga útil (o cargas útiles) del segmento de metadatos sigue la cabecera principal y se anidan (en algunos casos) dentro de los valores del elemento central.
- 15
- 20 Sistema de Sonorización y Rango Dinámico Optimizado
- La codificación de los metadatos y el esquema de transporte seguro descritos anteriormente se utilizan junto con un sistema escalable y extensible para optimizar la sonoridad y el rango dinámico a través de diferentes dispositivos de reproducción, aplicaciones y entornos de audición, tal como se ilustra en la Fig. 1. En una realización, el sistema 10 está configurado para normalizar los niveles de la sonoridad y el rango dinámico del audio 11 de entrada a través de varios dispositivos que requieren diferentes valores de sonoridad objetivo y tienen diferentes capacidades de rango dinámico. Para normalizar los niveles de sonoridad y rango dinámico, el sistema 10 incluye diferentes perfiles de dispositivos con el contenido de audio, y la normalización se hace basada en estos perfiles. Los perfiles pueden ser incluidos por una de las unidades de procesamiento de audio en las cadenas de procesamiento de audio y los perfiles incluidos pueden ser utilizados por una unidad de procesamiento aguas abajo en la cadena de procesamiento de audio para determinar la sonoridad y el rango dinámico objetivos deseados para un dispositivo objetivo. Componentes de procesamiento adicionales pueden proporcionar o procesar información para la gestión de perfiles de dispositivos (que incluyen (pero que no se limitan a) los siguientes parámetros: rango de banda nula, umbral de pico real, rango de sonoridad, constante de tiempo rápida/lenta (coeficientes) y máx. realce), control de ganancia y funciones de generación de ganancia de banda ancha y/o multibanda.
- 25
- 30
- 35 La Fig. 10 ilustra un diagrama más detallado del sistema de la Fig. 1 para un sistema que proporciona control de sonoridad y rango dinámico optimizados, en algunas realizaciones. Para el sistema 321 de la Fig. 10, la etapa del codificador comprende un componente codificador central 304 que codifica la entrada 303 de audio en un formato digital adecuado para su transmisión al decodificador 312. El audio se procesa de manera que pueda reproducirse en varios entornos de audición diferentes, cada uno de los cuales puede requerir diferentes ajustes objetivo de sonoridad y/o rango dinámico. Así, como se muestra en la Fig. 10, el decodificador emite una señal digital que es transformada en formato analógico mediante un transformador 316 de digital a analógico para reproducirse a través de una diversidad de diferentes tipos de controladores que incluyen altavoces 320 de rango completo, altavoces miniatura 322 y auriculares 324. Estos controladores ilustran sólo algunos ejemplos de posibles controladores de reproducción y se puede usar cualquier transductor o controlador de cualquier tamaño y tipo adecuados. Además, los controladores/transductores 320-324 de la Fig. 10 se pueden incorporar en cualquier dispositivo de reproducción adecuado para su uso en cualquier entorno de audición correspondiente. Los tipos de dispositivos pueden incluir, p. ej., AVR, televisores, equipo estéreo, ordenadores, teléfonos móviles, tabletas, reproductores de MP3, etc.; y los entornos de audición pueden incluir, p. ej., auditorios, hogares, automóviles, cabinas de audición, etc.
- 40
- 45
- 50 Dado que la gama de entornos de reproducción y tipos de controladores puede variar desde contextos privados muy pequeños hasta espacios públicos muy grandes, el margen de posibles y óptimas configuraciones de sonoridad y rango dinámico de la reproducción puede variar considerablemente dependiendo del tipo de contenido, niveles de ruido de fondo, y similares. Por ejemplo, en un entorno de cine en casa, se puede reproducir un amplio rango dinámico a través de equipos de sonido envolvente y un contenido de rango dinámico más estrecho se puede reproducir a través de un sistema de televisión normal (tal como un tipo LCD/LED de pantalla plana), mientras que puede utilizarse un modo de rango dinámico muy estrecho para ciertas condiciones de audición (p. ej., de noche o en un dispositivo con estrictas limitaciones de potencia acústica de salida, p. ej., la salida de altavoces internos o auriculares de un teléfono móvil/tableta) cuando no se desean grandes variaciones de nivel. En contextos de audición de portátiles o de móviles, como el uso de pequeños altavoces de ordenador o de base, o auriculares/audífonos, el rango dinámico óptimo de la reproducción puede variar dependiendo del entorno. Por ejemplo, en un entorno tranquilo, el rango dinámico óptimo puede ser mayor en comparación con un entorno ruidoso. Las realizaciones del sistema de procesamiento de audio adaptativo de la Fig. 10 variarán el rango dinámico para hacer que el contenido de audio sea más inteligible dependiendo de los parámetros, tales como el
- 55
- 60

entorno del dispositivo de audición y el tipo de dispositivo de reproducción.

La Fig. 11 es una tabla que ilustra diferentes requisitos de rango dinámico para una variedad de dispositivos de reproducción y entornos de audición de fondo en un ejemplo práctico de uso. Pueden derivarse requisitos similares para la sonoridad. Los diferentes requisitos de rango dinámico y de sonoridad generan diferentes perfiles que son utilizados por el sistema de optimización 321. El sistema 321 incluye un componente 302 de medida de la sonoridad y del rango dinámico que analiza y mide la sonoridad y el rango dinámico del audio de entrada. En una realización, el sistema analiza el contenido global del programa para determinar el parámetro global de la sonoridad. En este contexto, sonoridad se refiere a la sonoridad del programa a largo plazo o la sonoridad media de un programa, en donde un programa es una sola unidad de contenido de audio, como una película, un espectáculo en televisión, un anuncio o un contenido similar de un programa. La sonoridad se utiliza para proporcionar una indicación del perfil artístico del rango dinámico que es utilizado por los creadores de contenidos para controlar cómo se reproducirá el audio. La sonoridad está relacionada con el valor de los metadatos de dialnorm que representa la sonoridad media del diálogo de un único programa (p. ej., película, espectáculo, anuncio, etc.). El rango dinámico a corto plazo cuantifica las variaciones en las señales durante un período de tiempo mucho más corto que la sonoridad del programa. Por ejemplo, el rango dinámico a corto plazo puede medirse en el orden de segundos, mientras que la sonoridad del programa puede medirse durante un período de minutos o incluso de horas. El rango dinámico a corto plazo proporciona un mecanismo de protección que es independiente de la sonoridad del programa para asegurar que no se produzca sobrecarga para varios perfiles de reproducción y tipos de dispositivos. En una realización, el objetivo de sonoridad (sonoridad del programa a largo plazo) se basa en la sonoridad del diálogo y el rango dinámico a corto plazo se basa en la sonoridad con puerta y/o sin puerta relativa. En este caso, algunos componentes de DRC y de la sonoridad en el sistema se tienen en cuenta en el contexto con respecto al tipo de contenido y/o tipos y características de dispositivos de destino. Como parte de esta capacidad de tener en cuenta el contexto, el sistema está configurado para analizar una o más características del dispositivo de salida para determinar si el dispositivo es un miembro de grupos o de dispositivos particulares que están optimizados para ciertas condiciones de reproducción de DRC y de sonoridad, tales como dispositivos de tipo AVR, televisores, ordenadores, dispositivos portátiles, etc.

Un componente de preprocesamiento analiza el contenido del programa para determinar la sonoridad, picos, picos reales y períodos tranquilos para crear metadatos únicos para cada perfil de una pluralidad de perfiles diferentes. En una realización, la sonoridad puede ser una sonoridad con puerta al diálogo y/o una sonoridad con puerta relativa. Los diferentes perfiles definen varios modos de DRC (Control de Rango Dinámico) y de sonoridad objetivo en los que se generan diferentes valores de ganancia en el codificador dependiendo de las características del contenido de audio fuente, sonoridad objetivo deseada y del tipo y/o entorno del dispositivo de reproducción. El decodificador puede ofrecer diferentes modos de DRC y de la sonoridad objetivo (habilitados por los perfiles mencionados anteriormente) y puede incluir DRC y sonoridad objetivo deshabilitada/desactivada que permite un listado completo de rangos dinámicos sin compresión de la señal de audio y sin normalización de la sonoridad, DRC deshabilitado/desactivado y normalización de la sonoridad con un objetivo de modo de línea de -31 LKFS para reproducción en sistemas de cine en casa, que proporciona una compresión moderada del rango dinámico a través de valores de ganancia generados (específicamente para este modo de reproducción y/o perfil de dispositivo) en el codificador con normalización de la sonoridad con un objetivo de -31 LKFS; modo RF para la reproducción a través de altavoces de TV, que proporciona una gran cantidad de compresión del rango dinámico con normalización de la sonoridad con un objetivo de -24, -23 o -20 LKFS, un modo intermedio para la reproducción en ordenadores o dispositivos similares, que proporcionan compresión con normalización de la sonoridad a un objetivo de -14 LKFS, y el modo portátil, que proporciona una muy alta compresión del rango dinámico con un objetivo de normalización de la sonoridad de -11 LKFS. Los valores de la sonoridad objetivo de -31, -23/-20, -14 y -11 LKFS están destinados a ser ejemplos de diferentes perfiles de reproducción/dispositivos que pueden definirse para el sistema en algunas realizaciones, y puede usarse cualquier otro valor adecuado de la sonoridad objetivo, y el sistema genera valores de ganancia adecuados específicamente para estos modos de reproducción y/o perfil de dispositivo. Además, el sistema es extensible y adaptable para que se puedan acomodar diferentes dispositivos de reproducción y entornos de audición definiendo un nuevo perfil en el codificador o en otro lugar y cargado en el codificador. De esta forma, se pueden generar nuevos y únicos perfiles de reproducción/dispositivos para soportar dispositivos de reproducción mejorados o diferentes para futuras aplicaciones.

En una realización, los valores de ganancia pueden calcularse en cualquier componente de procesamiento adecuado del sistema 321, tal como en el codificador 304, en el decodificador 312 o en el transcodificador 308, o cualquier componente de preproceso asociado con el codificador o cualquier componente pos-proceso asociado con el decodificador.

La Fig. 13 es un diagrama de bloques que ilustra una interfaz entre diferentes perfiles para una variedad de diferentes clases de dispositivos de reproducción, en una realización. Como se muestra en la Fig. 13, un codificador 502 recibe la entrada de audio 501 y uno de varios perfiles 506 posibles diferentes. El codificador combina los datos de audio con el perfil seleccionado para generar un archivo de flujo de bits de salida que se procesa en componentes decodificadores presentes en o asociados con el dispositivo de reproducción de destino. Para el ejemplo de la Fig. 13, los diferentes dispositivos de reproducción pueden ser un ordenador 510, un teléfono móvil 512, un AVR 514 y un televisor 516, aunque también son posibles muchos otros dispositivos de salida. Cada uno de los dispositivos 510-516 incluye o está acoplado a altavoces (incluyendo controladores y/o transductores) tales como

los controladores 320-324. La combinación de los niveles de potencia, de procesamiento, y tamaños de los dispositivos de reproducción y altavoces asociados generalmente impone qué perfil es el más óptimo para ese objetivo en particular. De este modo, los perfiles 506 pueden estar definidos específicamente para la reproducción a través de AVR, TV, altavoces de móviles, auriculares de móviles, etc. También pueden definirse para modos o condiciones de funcionamiento específicos, como modo tranquilo, modo nocturno, exterior, interior, etc. Los perfiles mostrados en la Fig. 13 son sólo modos de ejemplo y se puede definir cualquier perfil adecuado, incluidos perfiles a medida para objetivos y entornos específicos.

Aunque la Fig. 13 ilustra una realización en la que un codificador 502 recibe los perfiles 506 y genera los parámetros adecuados para el procesamiento de la sonoridad y del DRC, debe señalarse que los parámetros generados basados en un perfil y contenido de audio pueden realizarse sobre cualquier unidad de procesamiento de audio adecuada tal como un codificador, decodificador, transcodificador, preprocesador, posprocesador, etc. Por ejemplo, cada dispositivo de salida 510-516 de la Fig. 13 tiene o está acoplado a un componente decodificador que procesa los metadatos en el flujo de bits en el archivo 504 enviado desde el codificador 502 para permitir la adaptación de la sonoridad y del rango dinámico para que coincidan con el dispositivo o tipo de dispositivo del dispositivo de salida objetivo.

En una realización, el rango dinámico y la sonoridad del contenido de audio se optimizan para cada posible dispositivo de reproducción. Esto se logra manteniendo la sonoridad a largo plazo hacia el objetivo y controlando el rango dinámico a corto plazo para optimizar la experiencia de audio (controlando la dinámica de la señal, picos de muestra y/o picos reales) para cada uno de los modos de reproducción objetivos. Se definen diferentes elementos de metadatos para la sonoridad a largo plazo y el rango dinámico a corto plazo. Como se muestra en la Fig. 10, el componente 302 analiza toda la señal de audio de entrada (o partes de la misma, tal como el componente de voz, si procede) para derivar las características relevantes para ambos de estos componentes DR separados. Esto permite definir diferentes valores de ganancia para ganancias artísticas frente a valores de ganancia de clip (protección contra la sobrecarga).

Estos valores de ganancia para la sonoridad a largo plazo y el rango dinámico a corto plazo se comparan entonces con un perfil 305 para generar parámetros que describen los valores de ganancia de control de la sonoridad y del rango dinámico. Estos parámetros se combinan con la señal de audio codificada procedente del codificador 304 en un multiplexor 306, o componente similar para la creación de un flujo de bits que se transmite a través del transcodificador 308 hasta una etapa decodificadora. La entrada del flujo de bits a la etapa decodificadora es demultiplexada en el demultiplexor 310. Entonces se decodifica en el decodificador 312. El componente 314 de ganancia aplica las ganancias correspondientes al perfil adecuado para generar datos de audio digital que se procesan después a través de la unidad 416 de DACS para reproducirse a través de los dispositivos de reproducción adecuados y controladores o transductores 320-324.

La Fig. 14 es una tabla que ilustra la correlación entre la sonoridad a largo plazo y el intervalo dinámico a corto plazo para una pluralidad de perfiles definidos, en una realización. Como se muestra en la Tabla 4 de la Fig. 14, cada perfil comprende un conjunto de valores de ganancia que imponen la cantidad de compresión de rango dinámico (DRC) aplicada en el decodificador del sistema o en cada dispositivo de destino. Cada uno de los N perfiles, indicados Perfiles 1 - N , establece parámetros de la sonoridad a largo plazo particulares (p. ej., dialnorm) y parámetros de compresión de sobrecarga imponiendo valores de ganancia correspondientes aplicados en la etapa del decodificador. Los valores de ganancia de DRC para los perfiles pueden ser definidos por una fuente externa que es aceptada por el codificador, o pueden ser generados internamente dentro del codificador como valores de ganancia por defecto si no se proporcionan valores externos.

En una realización, los valores de ganancia para cada perfil están incorporados en palabras de ganancia de DRC que se calculan basándose en un análisis de ciertas características de la señal de audio, tales como pico, pico real, sonoridad a corto plazo del diálogo o sonoridad a corto plazo global o una combinación (híbrida) de ambas, para calcular ganancias estáticas basadas en un perfil elegido (es decir, característica o curva de transferencia) así como las constantes de tiempo necesarias para aplicar el ataque rápido/lento y la liberación rápida/lenta de las ganancias del DRC final para cada perfil de dispositivo posible y/o sonoridad objetivo. Como se ha indicado anteriormente, estos perfiles pueden ser preestablecidos en el codificador, en el decodificador o generados externamente y llevados al codificador a través de metadatos externos desde el creador de contenidos.

En una realización, los valores de ganancia pueden ser ganancia de banda ancha que aplica la misma ganancia sobre todas las frecuencias del contenido de audio. Alternativamente, la ganancia puede estar comprendida por valores de ganancia de bandas múltiples, de manera que se aplican valores de ganancia diferentes a diferentes frecuencias o bandas de frecuencia del contenido de audio. En el caso multicanal, cada perfil puede constituir una matriz de valores de ganancia que indica ganancias para diferentes bandas de frecuencia en lugar de un solo valor de ganancia.

Con referencia a la Fig. 10, en una realización, la información relativa a las propiedades o características del entorno de audición y/o las capacidades y configuraciones de los dispositivos de reproducción, es proporcionada por la etapa decodificadora a la etapa codificadora por el enlace 330 de retroalimentación. La información 332 del perfil también se introduce en el codificador 304. En una realización, el decodificador analiza los metadatos en el flujo de bits para

determinar si un parámetro de la sonoridad para un primer grupo de dispositivos de reproducción de audio está disponible en el flujo de bits. Si es así, transmite los parámetros aguas abajo para su uso en la mezcla del audio. De lo contrario, el codificador analiza algunas características de los dispositivos para derivar los parámetros. Estos parámetros se envían después a un componente de mezclado aguas abajo para su reproducción. El codificador determina también un dispositivo de salida (o un grupo de dispositivos de salida que incluyen el dispositivo de salida) que mezclaría el flujo de audio recibido. Por ejemplo, el dispositivo de salida puede ser determinado como un teléfono celular o perteneciente a un grupo como dispositivos portátiles. En una realización, el decodificador utiliza el enlace 330 de retroalimentación para indicar al codificador el dispositivo de salida o el grupo de dispositivos de salida determinado. Para esta retroalimentación, un módulo conectado al dispositivo de salida (p. ej., un módulo en una tarjeta de sonido conectada a auriculares o conectada a altavoces en un ordenador portátil) puede indicar al decodificador una identidad del dispositivo de salida o una identidad de un grupo de dispositivos que incluye el dispositivo de salida. El decodificador transmite esta información al codificador a través del enlace 330 de retroalimentación. En una realización, el decodificador desempeña el decodificador determina los parámetros de sonoridad y de DRC. En una realización, el decodificador determina los parámetros de sonoridad y de DRC. En esta realización, en lugar de transmitir la información sobre el enlace 330 de retroalimentación, el decodificador utiliza la información sobre el dispositivo o el grupo de dispositivos de salida determinado para determinar los parámetros de sonoridad y de DRC. En otra realización, otra unidad de procesamiento de audio determina los parámetros de sonoridad y de DRC y el decodificador transmite la información a esa unidad de procesamiento de audio en lugar del decodificador.

La Fig. 12 es un diagrama de bloques de un sistema de optimización del rango dinámico, en una realización. Como se muestra en la Fig. 12, un codificador 402 recibe el audio de entrada 401. El audio codificado se combina en el multiplexor 409 con los parámetros 404 generados a partir de una curva 422 de compresión seleccionada y de un valor de dialnorm 424. El flujo de bits resultante es transmitido al desmultiplexor 411 que genera señales de audio que son decodificadas por el decodificador 406. Los parámetros y valores de dialnorm son usados por la unidad 408 de cálculo de ganancia para generar niveles de ganancia que controlan el amplificador 410 para la amplificación de la salida del decodificador. La Fig. 12 ilustra cómo se parametriza el control del rango dinámico y se inserta en el flujo de bits. La sonoridad también se puede parametrizar e insertar en el flujo de bits utilizando componentes similares. En una realización, también se puede proporcionar al decodificador un control de nivel de referencia de salida (no mostrado). Aunque la figura ilustra los parámetros de sonoridad y rango dinámico que se determinan e insertan en el codificador, una determinación similar puede realizarse en otras unidades de procesamiento de audio como un preprocesador, un decodificador y un posprocesador.

La Fig. 15 ilustra ejemplos de perfiles de la sonoridad para diferentes tipos de contenido de audio, en una realización. Como se muestra en la Fig. 15, las curvas de ejemplo 600 y 602 trazan la sonoridad de entrada (en LKFS) contra la ganancia centrada alrededor de 0 LKFS. Diferentes tipos de contenido presentan diferentes curvas, como se muestra en la Fig. 15 en la que la curva 600 puede representar la voz y la curva 602 puede representar el contenido de película estándar. Como se muestra en la Fig. 15, el contenido de voz está sujeto a una mayor cantidad de ganancia con relación al contenido de la película. La Fig. 15 se pretende que sea un ejemplo de curvas de perfil representativas para algunos tipos de contenido de audio, y también se pueden usar otras curvas de perfil. Algunos aspectos de las características del perfil, tal como se muestra en la Fig. 15, se utilizan para derivar los parámetros relevantes para el sistema de optimización. En una realización, estos parámetros incluyen: ancho de banda nulo, relación de corte, relación de realce, realce máximo, ataque de FS, decaimiento de FS, demora, límite de pico y sonoridad de nivel objetivo. Se pueden utilizar otros parámetros además o alternativamente a al menos algunos de estos parámetros dependiendo de los requisitos de aplicación y de las restricciones del sistema.

La Fig. 16 es un diagrama de flujo que ilustra un procedimiento para optimizar la sonoridad y el rango dinámico a través de dispositivos de reproducción y aplicaciones, en una realización. Aunque la figura ilustra la optimización de la sonoridad y del rango dinámico como se realiza en el codificador, una optimización similar puede realizarse en otras unidades de procesamiento de audio como en un preprocesador, en un decodificador y en un posprocesador. Como se muestra en el proceso 620, el procedimiento se inicia con una etapa codificadora que recibe una señal de entrada de una fuente (603). El codificador o un componente de preprocesamiento determina entonces si la señal de fuente se ha sometido o no a un proceso que logra una sonoridad y/o rango dinámico de destino (604). La sonoridad objetivo corresponde a la sonoridad a largo plazo y puede definirse externamente o internamente. Si la señal fuente no experimentó el proceso para alcanzar la sonoridad y/o rango dinámico de destino, el sistema lleva a cabo una operación adecuada de control de la sonoridad y/o rango dinámico (608); de lo contrario, si la señal de la fuente experimentó esta operación de control de la sonoridad y/o rango dinámico, el sistema entra en un modo de derivación para omitir las operaciones de control de la sonoridad y/o rango dinámico para permitir que el proceso original imponga la sonoridad y/o rango dinámico (606) a largo plazo adecuados. Los valores de ganancia adecuados bien para el modo 606 de derivación o el modo 608 llevado a cabo (que pueden ser valores de ganancia de banda ancha única o valores de ganancia de múltiples bandas dependientes de la frecuencia) se aplican después en el decodificador (612).

Formato del flujo de bits

Como se ha indicado anteriormente, el sistema para optimizar la sonoridad y el rango dinámico emplea un formato de metadatos extensible seguro para asegurar que los metadatos y el contenido de audio transmitidos en el flujo de

bits entre el codificador y el decodificador, o entre la fuente y los dispositivos de mezcla/reproducción no se han separado entre sí ni tampoco dañado durante la transmisión a través de redes u otros equipos propietarios, tales como interfaces de proveedores de servicios, etc. Este flujo de bits proporciona un mecanismo para señalar los componentes codificador y/o decodificador que adaptan la sonoridad y el rango dinámico de la señal de audio para adecuar el contenido de audio y las características del dispositivo de salida a través de la información adecuada del perfil. En una realización, el sistema está configurado para determinar un flujo de bits codificado de velocidad de bits baja que se transmite entre el codificador y el decodificador, y la información de la sonoridad codificada a través de los metadatos comprende características para uno o más perfiles de salida. A continuación sigue una descripción de un formato de flujo de bits para su uso con un sistema de optimización de la sonoridad y rango dinámico en una realización.

Un flujo de bits codificado por AC-3 comprende metadatos y de uno a seis canales de contenido de audio. El contenido de audio son datos de audio que han sido comprimidos mediante la codificación de audio perceptual. Los metadatos incluyen varios parámetros de metadatos de audio que se pretenden utilizar para cambiar el sonido de un programa que se envía a un entorno de audición. Cada trama de un flujo de bits de audio codificado por AC-3 contiene contenido de audio y metadatos para 1536 muestras de audio digital. Para una frecuencia de muestreo de 48 kHz, esto representa 32 milisegundos de audio digital o una velocidad de 31,25 tramas por segundo de audio.

Cada trama de un flujo de bits de audio codificado por E-AC-3 contiene contenido de audio y metadatos para 256, 512, 768 o 1536 muestras de audio digital, dependiendo de si la trama contiene uno, dos, tres o seis bloques de datos de audio respectivamente. Para una frecuencia de muestreo de 48 kHz, esto representa 5,333, 10,667, 16 o 32 milisegundos de audio digital, respectivamente, o una velocidad de 189,9, 93,75, 62,5 o 31,25 tramas por segundo de audio, respectivamente.

Como se indica en la Fig. 4, cada trama de AC-3 está dividida en secciones (segmentos), que incluyen: una sección de Información de Sincronización (SI) que contiene (como se muestra en la Fig. 5) una palabra de sincronización (SW) y la primera de dos palabras de corrección de error (CRC1); una sección de Información del Flujo de Bits (BSI) que contiene la mayoría de los metadatos; seis Bloques de Audio (AB0 a AB5) que contienen contenido de datos de audio comprimido (y también pueden incluir metadatos); bits residuales (W) que contienen cualquier bit no utilizado que queda después de comprimir el contenido de audio; una sección de información Auxiliar (AUX) que puede contener más metadatos; y la segunda de las dos palabras de corrección de errores (CRC2).

Como se indica en la Fig. 7, cada trama E-AC-3 está dividida en secciones (segmentos), que incluyen: una sección de Información de Sincronización (SI) que contiene (como se muestra en la Fig. 5) una palabra de sincronización (SW); una sección de Información de Flujo de Bits (BSI) que contiene la mayoría de los metadatos; entre uno y seis Bloques de Audio (AB0 a AB5) que contienen contenido de datos de audio comprimido (y también pueden incluir metadatos); bits residuales (W) que contienen cualesquiera bits no utilizados que quedan después de comprimir el contenido de audio; una sección de información Auxiliar (AUX) que puede contener más metadatos; y una palabra de corrección de errores (CRC).

En un flujo de bits AC-3 (o E-AC-3) hay varios parámetros de metadatos de audio que se pretenden usar específicamente para cambiar el sonido del programa enviado a un entorno de audición. Uno de los parámetros de metadatos es el parámetro dialnorm, que se incluye en el segmento BSI.

Como se muestra en la Fig. 6, el segmento BSI de una trama AC-3 incluye un parámetro de cinco bits ("dialnorm") que indica el valor de dialnorm para el programa. Un parámetro de cinco bits ("dialnorm2") que indica el valor de dialnorm para un segundo programa de audio llevado en la misma trama AC-3 se incluye si el modo de codificación de audio ("acmod") de la trama AC-3 es "0", lo que indica que se está utilizando una configuración de canal dualmono o "1+1".

El segmento BSI también incluye un indicador ("addbsie") que indica la presencia (o ausencia) de información adicional de flujo de bits después del bit "addbsie", un parámetro ("addbsil") que indica la longitud de cualquier información adicional de flujo de bits según el valor de "addbsil", y hasta 64 bits de información adicional de flujo de bits ("addbsi") según el valor de "addbsil". El segmento de BSI puede incluir otros valores de metadatos no mostrados específicamente en la Fig. 6.

Aspectos de una o más realizaciones descritas en el presente documento pueden aplicarse en un sistema de audio que procesa señales de audio para transmisión a través de una red que incluye uno o más ordenadores o dispositivos de procesamiento que ejecutan instrucciones de software. Cualquiera de las realizaciones descritas se pueden usar solas o juntas entre sí en cualquier combinación. Aunque diversas realizaciones pueden haber sido motivadas por varias deficiencias con la técnica anterior, que pueden ser discutidas o aludidas en uno o más lugares de la memoria descriptiva, las realizaciones no se refieren necesariamente a ninguna de estas deficiencias. En otras palabras, diferentes realizaciones pueden abordar diferentes deficiencias que pueden ser tratadas en la memoria descriptiva. Algunas realizaciones sólo pueden abordar parcialmente algunas deficiencias o sólo una deficiencia que puede ser tratada en la memoria descriptiva, y algunas realizaciones pueden no abordar ninguna de estas deficiencias.

Aspectos de los sistemas descritos en el presente documento pueden aplicarse en un entorno informático adecuado de red de procesamiento de sonido para procesar archivos de audio digitales o digitalizados. Partes del sistema de audio adaptativo pueden incluir una o más redes que comprenden cualquier número deseado de máquinas individuales, incluyendo uno o más enrutadores (no mostrados) que sirven para almacenar en memoria intermedia y enrutar los datos transmitidos entre los ordenadores. Una red de este tipo puede estar construida sobre varios protocolos de red diferentes, y puede ser Internet, una Red de Área Ancha (WAN), una Red de Área Local (LAN) o cualquier combinación de las mismas.

Uno o más de los componentes, bloques, procesos u otros componentes funcionales pueden ser aplicados a través de un programa informático que controla la ejecución de un dispositivo informático del sistema basado en un procesador. También debe señalarse que las diversas funciones descritas en el presente documento pueden describirse utilizando cualquier número de combinaciones de hardware, firmware y/o como datos y/o instrucciones incorporadas en diversos medios legibles por máquina o legibles por ordenador, en términos de su comportamiento, transferencia de registro, componente lógico y/u otras características. Los medios legibles por ordenador en los que tales datos y/o instrucciones formateadas pueden ser incorporados incluyen, pero no se limitan a, medios de almacenamiento físico (no transitorios) no volátiles, en diversas formas, tales como medios de almacenamiento ópticos, magnéticos o semiconductores.

A menos que el contexto requiera claramente lo contrario, a lo largo de la descripción y de las reivindicaciones, las palabras "comprender", "comprendiendo", y similares, deben interpretarse en un sentido inclusivo en oposición a un sentido exclusivo o exhaustivo; es decir, en un sentido de "incluir, pero no de limitarse a". Las palabras que usan el término singular o plural incluyen también el término plural o singular, respectivamente. Además, los términos o expresiones "en el presente documento", "a continuación", "arriba", "abajo" y palabras de significado similar se refieren a esta solicitud en su conjunto y no a ninguna parte en particular de esta solicitud. Cuando la palabra "o" se utiliza en referencia a una lista de dos o más elementos, esa palabra cubre todas las siguientes interpretaciones de la palabra: cualquiera de los elementos de la lista, todos los elementos de la lista y cualquier combinación de los elementos de la lista.

Aunque se han descrito una o más aplicaciones a modo de ejemplo y en virtud de realizaciones específicas, debe entenderse que una o más aplicaciones no están limitadas a las realizaciones descritas. Por el contrario, se pretende cubrir diversas modificaciones y disposiciones similares que resultarían evidentes para los expertos en la técnica. Por lo tanto, el alcance de las reivindicaciones adjuntas debe recibir la interpretación más amplia de modo que abarque todas esas modificaciones y disposiciones similares.

REIVINDICACIONES

1. Un procedimiento, que comprende:
 recibir, en un flujo de bits que comprende datos de audio, metadatos asociados con los datos de audio;
 analizar los metadatos en el flujo de bits para determinar si un parámetro de la sonoridad para un primer grupo de dispositivos de reproducción de audio (510, 512, 514, 516) está disponible en el flujo de bits;
 sensible para determinar que el parámetro está presente en el primer grupo, utilizando el parámetro y los datos de audio para mezclar el audio; y
 sensible para determinar que el parámetro de la sonoridad no está presente en el primer grupo, analizar una o más características del primer grupo, y determinar el parámetro basado en una o más características,
 en donde una o más características comprenden niveles de ganancia para diferentes perfiles de los datos de audio.
2. El procedimiento de la reivindicación 1, en donde los niveles de ganancia definen al menos uno de: un perfil artístico de rango dinámico que controla cómo se reproducirán los datos de audio para un programa definido, y un perfil de rango dinámico a corto plazo que proporciona protección de sobrecarga para partes del programa definido.
3. El procedimiento de la reivindicación 1 en donde usando el parámetro y los datos de audio para mezclar el audio comprende transmitir el parámetro y los datos de audio a un módulo aguas abajo que mezcla el audio para su reproducción.
4. El procedimiento de la reivindicación 1, en donde utilizando el parámetro y los datos de audio para mezclar el audio comprende mezclar los datos de audio basados en el parámetro y en los datos de audio.
5. El procedimiento de la reivindicación 1, que comprende además:
 determinar un dispositivo de salida que mezclaría el flujo de audio recibido; y
 determinar si el dispositivo de salida pertenece o no al primer grupo de dispositivos de reproducción de audio, en donde el paso de analizar los metadatos en el flujo para determinar si el parámetro de sonoridad del primer grupo de dispositivos de reproducción de audio están disponibles, se ejecuta después del paso de determinar que el dispositivo de salida pertenece al primer grupo de dispositivos de reproducción de audio.
6. El procedimiento de la reivindicación 5, en donde la etapa de determinar que el dispositivo de salida pertenece al primer grupo de dispositivos de reproducción de audio comprende:
 recibir una indicación de un módulo conectado al dispositivo de salida que indica una identidad del dispositivo de salida o que indica una identidad de un grupo de dispositivos que incluyen el dispositivo de salida; y
 determinar que el dispositivo de salida pertenece al primer grupo de dispositivos de reproducción de audio basados en la indicación recibida.
7. Un medio legible por ordenador que comprende un programa de ordenador, estando configurado el programa de ordenador para controlar la ejecución de un dispositivo informático basado en un procesador de modo que realice el procedimiento de una cualquiera de las reivindicaciones anteriores.
8. Un aparato, que comprende:
 una interfaz configurada para recibir un flujo de bits que incluye datos de audio y metadatos asociados con los datos de audio;
 un analizador acoplado a la interfaz y configurado para analizar los metadatos en el flujo de bits para determinar si está disponible en el flujo de bits un parámetro de la sonoridad para un primer grupo de dispositivos de reproducción de audio (510, 512, 514, 516); y
 un componente de mezcla configurado para usar los parámetros y los datos de audio para mezclar el audio sensible al analizador que determina que los parámetros están presentes en el primer grupo, el componente de mezcla configurado además para analizar una o más características del primer grupo y determinar el parámetro basado en una o más características sensibles al analizador que determina que los parámetros de la sonoridad no están presentes en el primer grupo,
 en donde una o más características comprenden niveles de ganancia para diferentes perfiles de los datos de audio.
9. El aparato de la reivindicación 8, en donde el componente de mezcla que utiliza los parámetros y los datos de audio para mezclar audio comprende el componente de mezclado que transmite los parámetros y datos de audio hasta un módulo aguas abajo que mezcla el audio para su reproducción.
10. El aparato de la reivindicación 9, en donde el componente de mezcla que utiliza el parámetro y los datos de audio para mezclar el audio comprende el componente de mezcla que mezcla los datos de audio basados en el parámetro y en los datos de audio.
11. El aparato de la reivindicación 10 que comprende además un segundo componente configurado para determinar un dispositivo de salida que mezclaría el flujo de audio recibido y determinaría si el dispositivo de salida pertenece o no al primer grupo de dispositivos de reproducción de audio; en donde el analizador analiza los metadatos en el flujo para determinar si el parámetro de la sonoridad del primer grupo de dispositivos de reproducción de audio está disponible después de que el segundo componente determine si el dispositivo de salida pertenece o no al primer grupo de dispositivos de reproducción de audio.

12. El aparato de la reivindicación 11 que comprende además una interfaz configurada para recibir una indicación desde un módulo conectado al dispositivo de salida que indica una identidad del dispositivo de salida o que indica una identidad de un grupo de dispositivos que incluye el dispositivo de salida, y para determinar que el dispositivo de salida pertenece al primer grupo de dispositivos de reproducción de audio basado en la indicación recibida

5

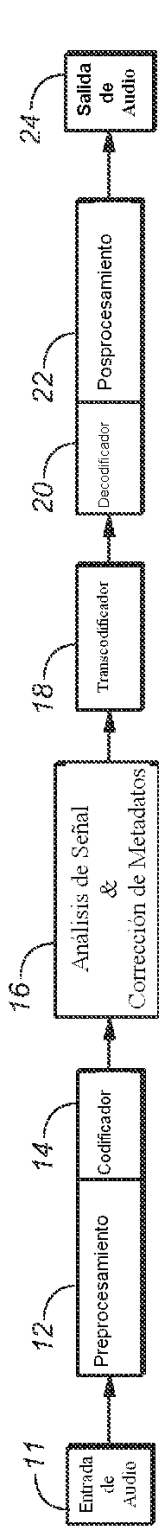


FIG. 1

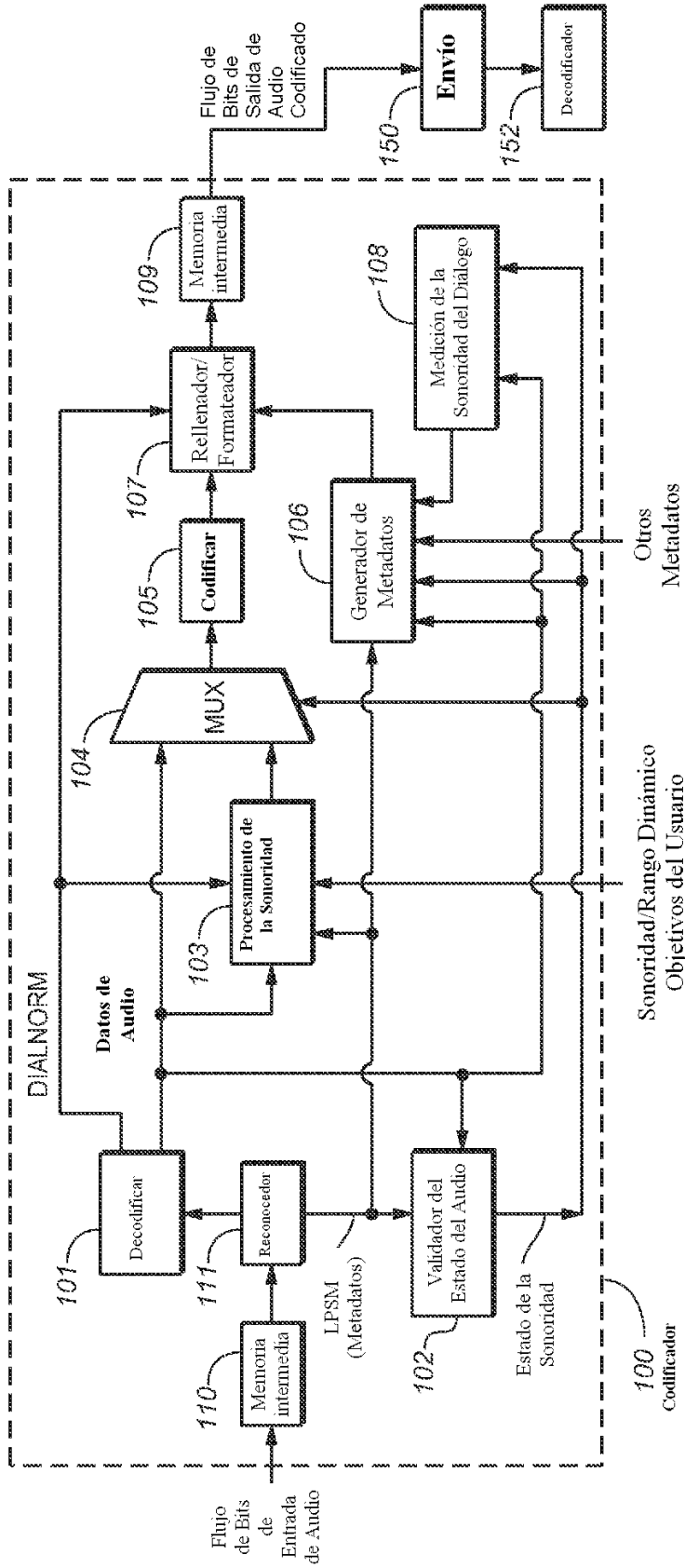


FIG. 2

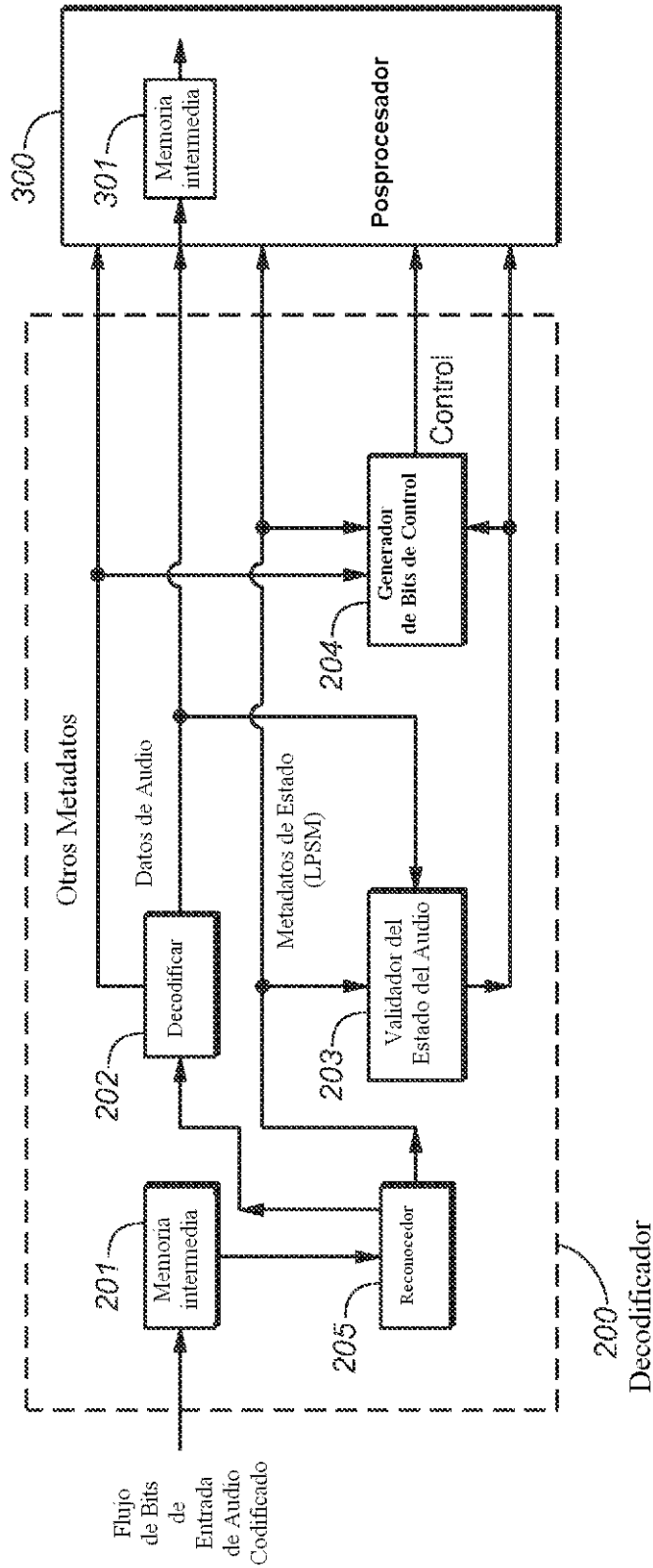


FIG. 3



FIG. 4

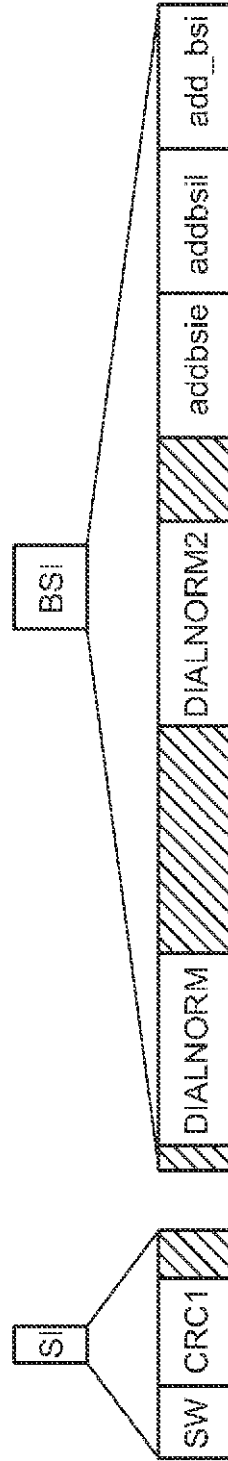


FIG. 5

FIG. 6

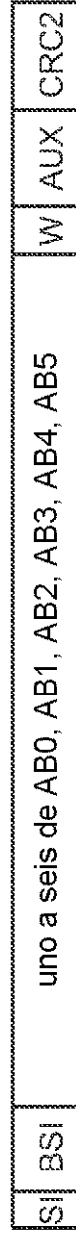


FIG. 7

Parámetro	Descripción	Obligatorio (M) /Opcional (O)
SYNC [ID]		M
Versión del elemento central		M
Longitud del elemento central		M
Período del elemento central (xxx)		M
Recuento del elemento ampliado	Indica el número de elementos metadatos ampliados asociados con el elemento central. Este valor puede incrementarse/decrementarse a medida que el flujo de bits pasa desde producción a través de la distribución y emisión final.	M
Asociación de flujo adicional	Describe con qué flujo (o flujos) adicional (o adicionales) está asociado el elemento central	M
Firma (compendio HMAC)	Resumen de HMAC de 256 bits (utilizando algoritmo SHA-2) calculada sobre los datos de audio, el elemento central, y todos los elementos ampliados, de la trama completa.	M
Cuenta atrás del límite de PGM	El campo solo aparece para algún número de tramas a la cabeza o a la cola de un archivo/flujo de programa de audio. De este modo, un cambio en la versión del elemento central podría usarse para señalar la inclusión de este parámetro.	O
Huella Digital de Audio	Huella Digital de Audio tomada sobre algún número de muestras de audio representadas por el campo del período del elemento central.	O
Huella Digital de Video	Huella Digital de Video tomada sobre algún número de muestras (si acaso) de video comprimidas representadas por el campo del período del elemento central.	O
URL/UUID	Este campo se define para llevar una URL y/o una UUID (puede ser innecesaria para la huella digital) que mencione una ubicación externa de contenido de programa adicional (esencia) y/o metadatos asociados con el flujo de bits.	O

TABLA 1

FIG. 8

Parámetro LPSM [Sonoridad Inteligente]	Descripción
Versión LPSM	
Período LPSM (xxx)	Aplicable solo a campos (xxx)
Recuento LPSM	
Asociación del flujo adicional de LPSM	
Canal (o canales) de diálogo	Indica qué combinación de canales de audio L, C & R contienen voz sobre los anteriores 0,5 segundos. Cuando la voz no está presente en ninguna combinación L, C o R, entonces este parámetro indicará "no diálogo"
Tipo de Regulación de la Sonoridad	Indica que el flujo de datos de audio asociados con arreglo a un conjunto específico de regulaciones (p. ej., ATSC A/85 o EBU R128)
Indicador de Corrección de la Sonoridad con puerta de Diálogo	Indica si el flujo de audio asociado se ha corregido basado en el umbral de puerta de diálogo
Tipo de Corrección de la Sonoridad	Indica si el flujo de audio asociado se ha corregido con un análisis por anticipación de infinidad (basada en archivo) o con un controlador de sonoridad y rango dinámico en tiempo real (RT).
Sonoridad de puerta relativa ITU (INF)	Indica la sonoridad integrada ITU-R BS. 1770-2 del flujo de audio asociado sin metadatos aplicados (p. ej., 7 bits: de -58 LKFS a +5,5 LKFS, en pasos de 0,5 LKFS)
Sonoridad de puerta de voz ITU (INF)	Indica la sonoridad integrada ITU-R BS. 1770-2 del flujo de audio asociado sin metadatos aplicados (p. ej., 7 bits: de -58 LKFS a +5,5 LKFS, en pasos de 0,5 LKFS)
Sonoridad a 3 s a corto plazo ITU (EBU 3341)	Indica la sonoridad ITU sin puerta de 3 segundos del flujo de audio asociado sin metadatos aplicados (ventana deslizante) @ velocidad de inserción ~10 Hz (p. ej., 8 bits: 116 a +11,5 LKFS, en pasos de 0,5 LKFS)
Valor de Pico Real	Indica el valor de pico Real ITU-R BS.1770 Anexo 2 (dB del TP) del flujo de audio asociado sin metadatos aplicados (es decir, el valor más amplio sobre el período de la trama señalizada en el campo del período del elemento) 116 a +11,5 LKFS, en pasos de 0,5 LKFS

TABLA 2

FIG. 9A

Número de Estados Únicos	Obligatorio (M) / Opcional (O)	Velocidad de Inserción (Período de parámetro de actualización)
	M	
	M	
	M	
	M	
8	M	~0,5 segundos (típico)
8	M	Trama
2	O (presente solo si Loudness_Regulation_Type indica que el audio correspondiente NO SE HA CORREGIDO)	Trama
2	O (presente solo si Loudness_Regulation_Type indica que el audio correspondiente NO SE HA CORREGIDO)	Trama
128	O	1 s
128	O	1 s
256	O	0,1 s
256	O	0,5 s

TABLA 2

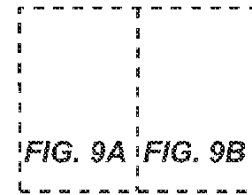


FIG. 9

FIG. 9B

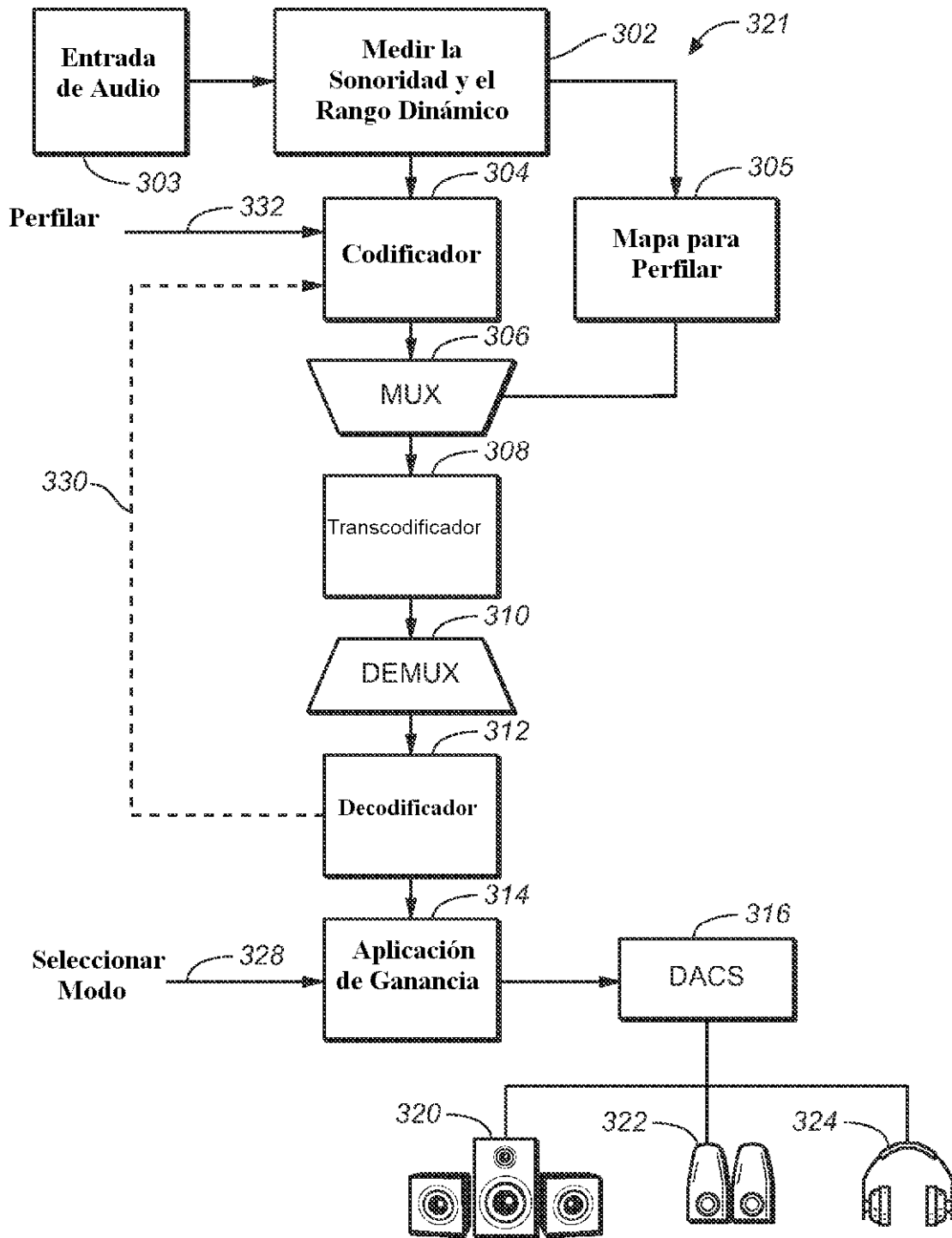


FIG. 10

ALTAVOCES	ESTADO DE LA REPRODUCCIÓN	
	TRANQUILA	RUIDOSA
PEQUEÑOS/MINI	RANGO DINÁMICO BAJO	RANGO DINÁMICO BAJO
AURICULARES	RANGO DINÁMICO MEDIO	RANGO DINÁMICO BAJO
GRANDES	RANGO DINÁMICO ALTO	RANGO DINÁMICO BAJO

TABLA 3

FIG. 11

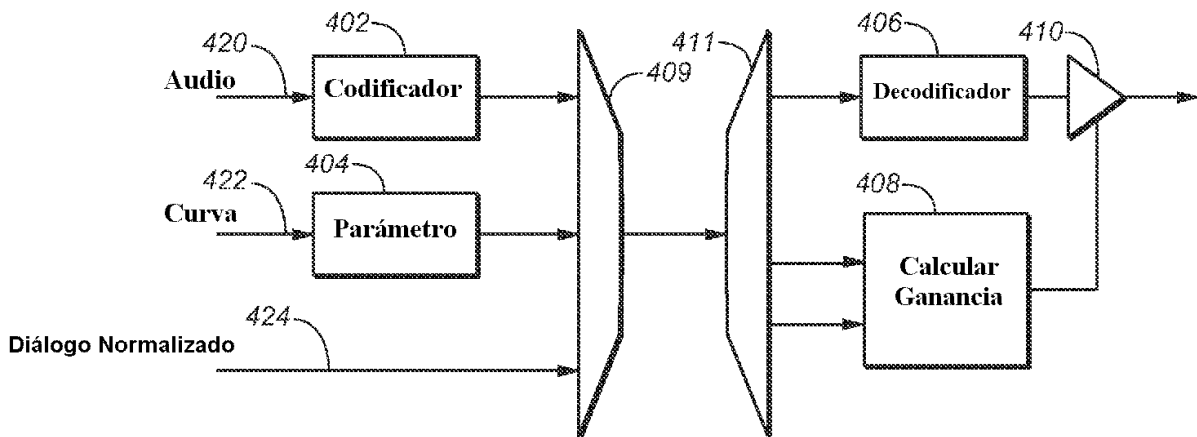


FIG. 12

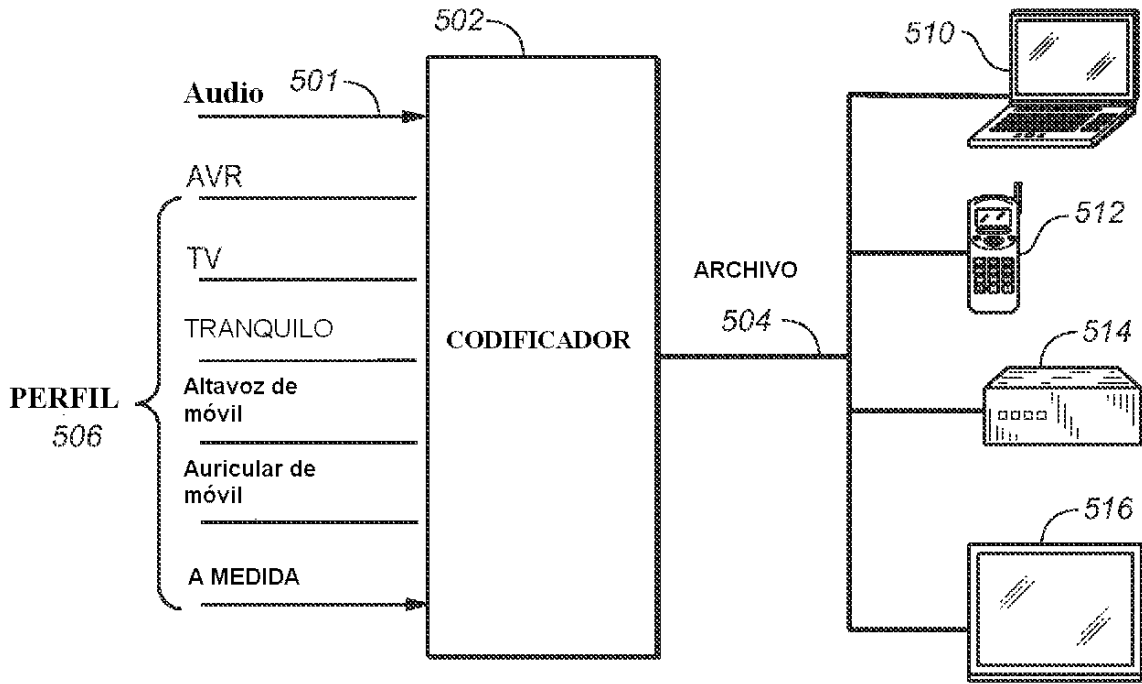


FIG. 13

SONORIDAD A LARGO PLAZO	GANANCIAS DEL DRC	COMPRESIÓN POR SOBRECARGA
1	Perfil 1	1
2	Perfil 2	2
3	Perfil 3	3
⋮	⋮	⋮
N	Perfil N	N

TABLA 4

FIG. 14

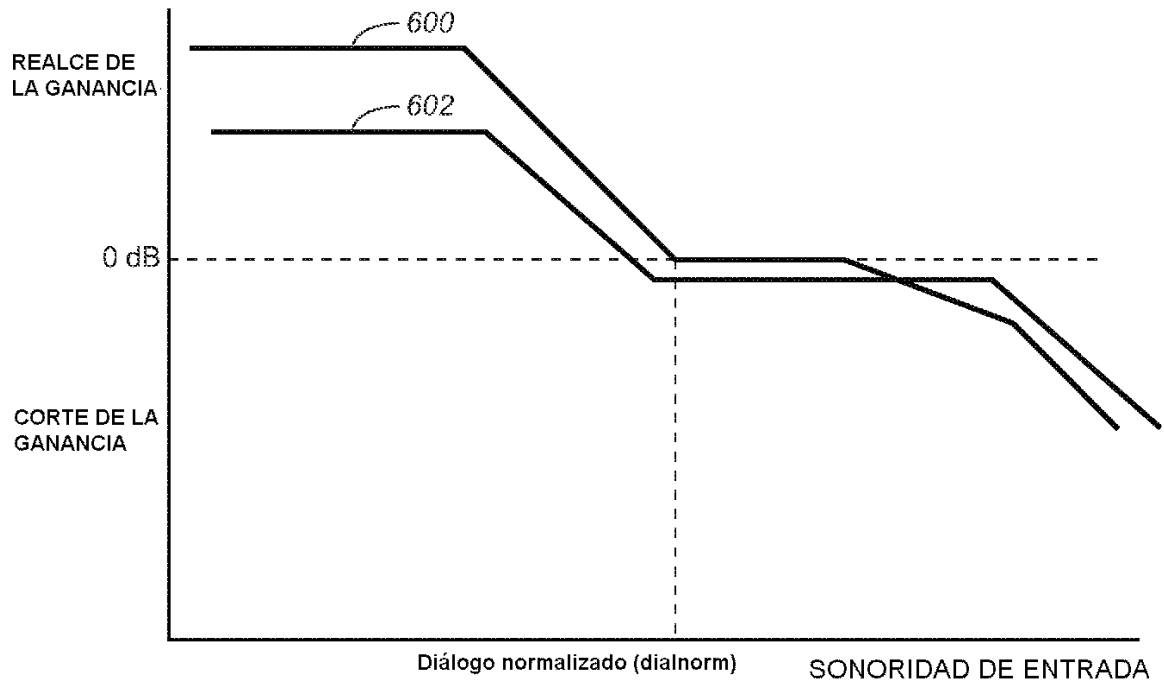


FIG. 15

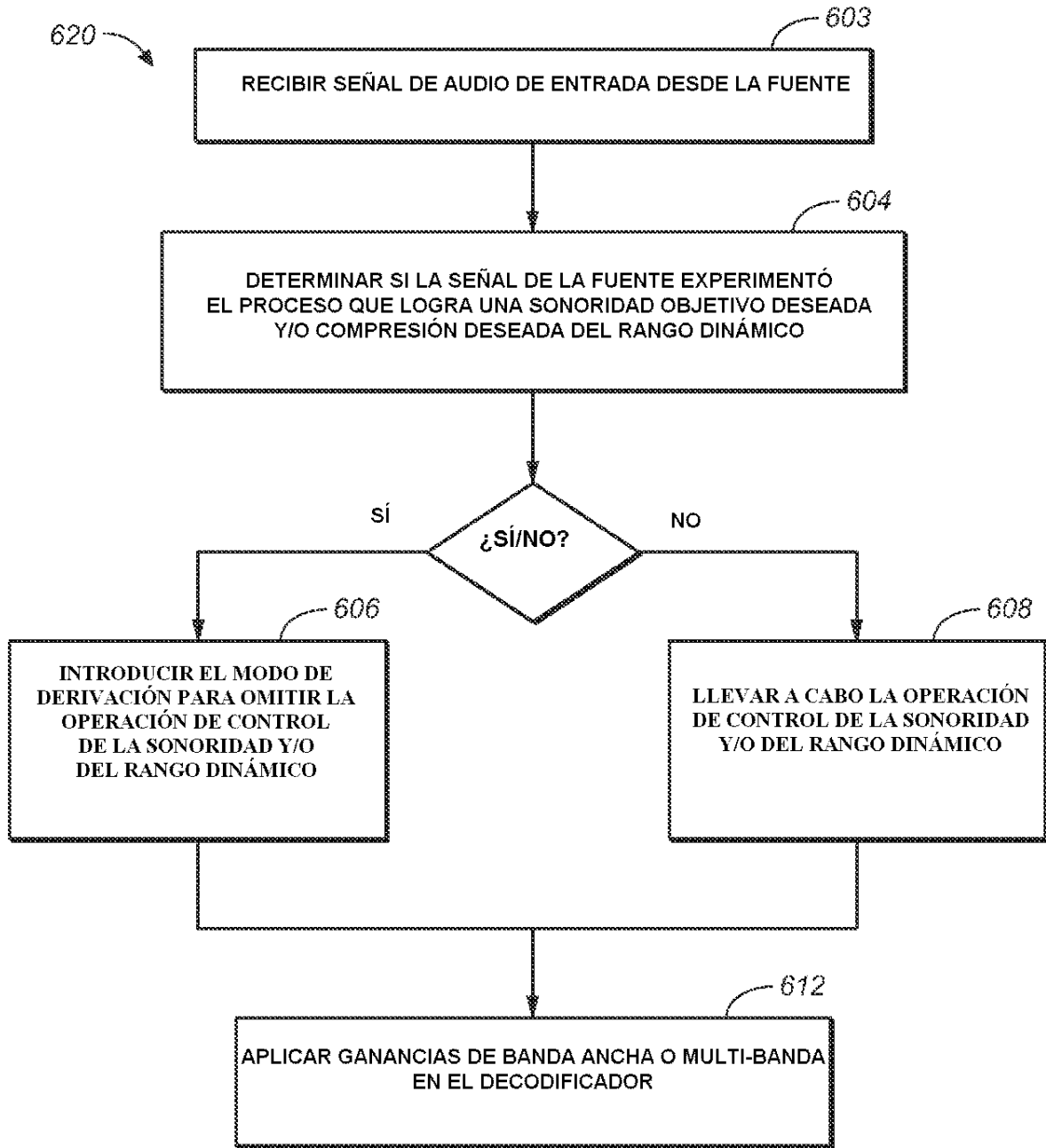


FIG. 16