

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 738 494**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 21/0316 (2013.01)

H03G 7/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.03.2014 PCT/US2014/031992**

87 Fecha y número de publicación internacional: **02.10.2014 WO14160849**

96 Fecha de presentación y número de la solicitud europea: **27.03.2014 E 14724887 (6)**

97 Fecha y número de publicación de la concesión europea: **24.04.2019 EP 2956936**

54 Título: **Metadatos para control de sonoridad y rango dinámico**

30 Prioridad:

29.03.2013 US 201361806570 P
26.03.2014 US 201414225950

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
23.01.2020

73 Titular/es:

APPLE INC. (100.0%)
One Apple Park Way
Cupertino CA 95014, US

72 Inventor/es:

BAUMGARTE, FRANK;
ALLAMANCHE, ERIC A. y
STROMMER, STEFAN K.O.

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 738 494 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

Metadatos para control de sonoridad y rango dinámico

5 **[0001]** Esta solicitud no provisional reivindica el beneficio de la fecha de presentación anterior de la Solicitud Provisional de EE. UU. n.º 61/806 570, presentada el 29 de marzo de 2013.

10 **[0002]** Un modo de realización de la invención se refiere en general a la codificación y descodificación de una señal de audio y el uso de los metadatos asociados con la señal codificada para mejorar la calidad de la reproducción de la señal descodificada en un dispositivo de usuario final de electrónica de consumo. También se describen otros modos de realización.

ANTECEDENTES

15 **[0003]** El contenido de audio digital aparece en muchos casos diferentes, incluyendo por ejemplo los archivos de música y películas. En la mayoría de los casos, una señal de audio se codifica para fines de reducción de la velocidad de datos, de modo que la transferencia o entrega del archivo o flujo de medios consume menos ancho de banda y es más rápida, lo cual permite que ocurran muchas otras transferencias simultáneamente. El archivo o flujo de medios se puede recibir en diferentes tipos de dispositivos de usuario final donde la señal de audio codificada se descodifica antes de ser presentada al consumidor a través de altavoces integrados o desmontables. Esto ha ayudado a alimentar el apetito de los consumidores por obtener medios digitales a través de Internet. Los creadores y distribuidores de programas de audio digital tienen varios estándares de la industria a su disposición, que pueden usarse para codificar y descodificar contenido de audio. Entre estos se incluye el estándar de compresión de audio digital (AC-3, E-AC-3), Revisión B, Documento A/52B, 14 de junio de 2005, publicado por Advanced Television Systems Committee, Inc. (el "Estándar ATSC"), Instituto de Estándares Europeos de Telecomunicaciones, ETSI TS 101154 Transmisión de video digital (DVB) basada en el flujo de transporte MPEG-2 en ISO/IEC 13818-7, Codificación de audio avanzada (AAC) ("MPEG-2 AAC Standard") e ISO/IEC 14496-3 ("MPEG-4 Audio"), publicado por la Organización Internacional de Normalización (ISO).

20 **[0004]** Existe una creciente variedad de dispositivos de usuario final para la reproducción de audio digital, incluyendo ordenadores de sobremesa, ordenadores portátiles, dispositivos de mano portátiles (por ejemplo., teléfonos inteligentes), televisores domésticos y sistemas de medios de comunicación en el vehículo. Estos dispositivos tienen diferentes rutas de señal analógica, altavoces y entornos acústicos. Además, el rango dinámico de una señal de audio varía entre los diferentes programas. Además, los productores (incluidos los creadores y, a veces, incluso, los distribuidores) de programas de audio digital a menudo desean aumentar la sonoridad media de sus programas, modificando digitalmente una señal de audio para que su sonoridad media sea superior en varios dB. Sin embargo, hacerlo también requiere que los niveles máximos de la señal de audio resultante se reduzcan para evitar el recorte (lo cual ocasiona una distorsión audible no deseada). Esto se logra mediante el control de rango dinámico (DRC), que comprime los altos y bajos de la señal de audio para que la señal de audio resultante pueda caber dentro de una envoltura más estrecha (evitando así el recorte). Sin embargo, todos estos factores crean un problema, ya que la sonoridad percibida por el usuario final puede variar significativamente, tanto en diferentes dispositivos de usuario final como entre programas consecutivos en el mismo dispositivo, lo cual lleva a una experiencia de usuario desagradable durante la reproducción.

35 **[0005]** Una herramienta de software como el programa SoundCheck™ de Apple Inc. ajusta automáticamente el volumen de reproducción de las canciones para obtener la misma sonoridad perceptible, de modo que, por ejemplo, una grabación de pop reciente que tiene un rango dinámico más pequeño pero una mayor sonoridad media es rechazada, en comparación con una canción más antigua de la década de 1970. Además, un programa de audio puede incluir una parte de metadatos que está asociada con la señal de audio codificada y que describe la señal de audio asociada. Los metadatos pueden incluir información que utiliza el software en el dispositivo del usuario final para controlar, por ejemplo, el nivel de diálogo, el DRC y cualquier mezcla descendente de la señal de audio descodificada, para cambiar la experiencia del consumidor durante la reproducción. El documento "Guía para los metadatos de Dolby", 1 de enero de 2005, se relaciona con los metadatos de Dolby y, en particular, con el control de rango dinámico (DRC). El documento describe varios perfiles de control de rango dinámico diferentes que se pueden aplicar a una señal de audio de entrada 5 para producir parámetros de DRC en el lado de la codificación. En el lado de la descodificación, estos parámetros DRC se reciben en los metadatos y luego se aplican a la señal de audio descodificada para comprimir la señal de audio.

45 **[0006]** El documento US 2006/002572 A1 divulga un procedimiento para corregir los metadatos que afecta a la sonoridad de reproducción 10 y al rango dinámico de la información de audio.

SUMARIO

50 **[0007]** *La invención se define mediante las reivindicaciones independientes. Las reivindicaciones dependientes definen modos de realización ventajosos.* Es deseable un enfoque sistemático pero aún flexible para controlar la

calidad del audio entregado a un consumidor, utilizando cualquier codificación/descodificación (códec) convencional y construcción de metadatos asociados.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

5 **[0008]** Los modos de realización de la invención se ilustran a modo de ejemplo y no a modo de limitación en las figuras de los dibujos adjuntos en los que referencias similares indican elementos similares. Debe observarse que las referencias a "un" modo de realización de la invención en esta divulgación no son necesariamente al mismo modo de realización, y significan al menos uno. Además, una figura dada puede usarse para ilustrar las características de más de un modo de realización de la invención, y no todos los elementos en la figura pueden requerirse para un modo de realización dado.

15 La **Fig. 1** es un diagrama de bloques de componentes relevantes de una etapa de codificación de audio digital o un aparato de codificación.

La **Fig. 2** muestra varios ejemplos de características de DRC que se pueden usar en la etapa de codificación.

20 La **Fig. 3** es un diagrama de bloques de componentes relevantes de una etapa de descodificador o aparato descodificador, y en particular uno en el que el procesamiento se realiza durante la reproducción de la señal de audio descodificada.

La **Fig. 4** muestra varios ejemplos de características de DRC que se pueden usar en la etapa de descodificación para producir nuevos valores de ganancia de DRC.

25 La **Fig. 5** es un diagrama de bloques de componentes en la etapa de descodificación que tiene un procesador de mezcla descendente que recibe una señal de audio multicanal descodificada.

30 La **Fig. 6** muestra varios ejemplos de características de DRC adecuadas para usar en una etapa de descodificación para producir nuevos valores de ganancia.

La **Fig. 7** es un diagrama de bloques de componentes relevantes de una etapa de codificación de audio digital o un aparato de codificación adecuado para un escenario de mezcla descendente particular.

DESCRIPCIÓN DETALLADA

35 **[0009]** Varios modos de realización de la invención se describen aquí como mejoras de metadatos en la codificación y descodificación de audio digital, que se utiliza para la normalización de sonoridad y el control de rango dinámico (DRC) durante la reproducción de un archivo de audio codificado o un flujo de audio codificado que tiene los metadatos especificados. Aunque se exponen numerosos detalles, se entiende que algunos modos de realización de la invención pueden ponerse en práctica sin estos detalles. En otros casos, no se han mostrado en detalle circuitos, estructuras y técnicas bien conocidos para no confundir la comprensión de esta descripción. Por ejemplo, aquí se describen ciertos detalles en el contexto de la codificación para la reducción de la velocidad de transmisión de bits de acuerdo con los estándares MPEG; los modos de realización de la invención también son aplicables a otras formas de codificación y descodificación de audio, incluida la compresión de datos sin pérdida, como Apple Lossless Audio Codec (ALAC).

45 **[0010]** Con referencia a la **Fig. 1**, se muestra un diagrama de bloques de una etapa de codificador o un aparato de codificador, de acuerdo con un modo de realización de la invención. Este diagrama de bloques se puede usar para describir no solo un aparato codificador de audio digital, sino también un procedimiento para codificar una señal de audio. Una señal de audio digital está codificada por un codificador 2, y luego se puede ensamblar con metadatos utilizando un multiplexor (mux) 8. El flujo de bits de audio codificado o audio codificado resultante contiene metadatos incorporados que incluyen un valor de ganancia de control de rango dinámico (DRC) por trama de audio digital. Aunque no se muestra, una alternativa para integrar los metadatos con la señal codificada es proporcionar los metadatos (incluidos en este caso los valores de ganancia DRC por trama) a través de un canal de metadatos separado al que se hace referencia en la **Fig. 1** como un canal de datos auxiliar 7. La codificación de la señal de audio puede implicar la codificación de reducción de la velocidad de datos con pérdida o la compresión de datos sin pérdida. Además, la señal de audio puede ser una secuencia de audio digital modulada por código de impulsos, PCM, u otra señal de audio de un solo canal o multicanal muestreada (también denominada aquí programa de audio) que puede incluir múltiples canales de audio, por ejemplo, seis canales 5.1 de sonido envolvente, 2 canales estéreo, etc.

60 **[0011]** Los metadatos incluyen un valor de ganancia DRC (por trama) que se calcula mediante un procesador DRC 4. Un normalizador o ajustador, representado por un símbolo multiplicador como se muestra, aplica un valor de ganancia de normalización de audio (por ejemplo, G_{LN}) a una señal de audio digital de entrada, para producir una señal normalizada. La señal normalizada es procesada (por el procesador 4 de DRC) para calcular un número de valores de ganancia de DRC de acuerdo con uno seleccionado de un número de características de DRC predefinidas. Las características de DRC pueden almacenarse dentro de la memoria como parte del procesador 4 de DRC, dentro del aparato de codificación. En la **Fig. 2** se dan ejemplos de las características de DRC, donde el nivel de entrada de DRC

a lo largo del eje x se refiere a un valor de sonoridad a corto plazo (también conocido aquí como nivel de entrada de DRC), y un rango de valores de ganancia a lo largo del eje y que deben aplicarse para lograr un efecto de compresión o expansión sobre la señal de audio de entrada. Los valores de ganancia en la **Fig. 2** también se conocen aquí como valores de ganancia DRC.

[0012] La normalización de audio es la aplicación de una cantidad constante de ganancia para una grabación de audio (también denominado un programa de audio o una señal de audio) para llevar un valor medio o de pico de la señal de audio a un nivel objetivo (la norma). Cuando se aplica la misma cantidad de ganancia a través de la señal completa o el programa de audio, la relación señal/ruido y la dinámica relativa de la señal no se modifican en su mayor parte. La normalización difiere de la compresión de rango dinámico o del control de rango dinámico (DRC), que aplica niveles de ganancia variables en el tiempo a la señal de audio de entrada, para ajustar el resultado dentro de un rango de mínimo a máximo. Los ejemplos incluyen la normalización de picos en la que la ganancia se cambia para llevar el valor de muestra digital más alto (por ejemplo, valor modulado por código de pulso, PCM) o pico de señal analógica a un nivel dado.

[0013] Otro tipo de normalización de audio se basa en una medición de la sonoridad de programa. Aquí, la ganancia aplicada (que se representa en la **Fig. 1** específicamente como la ganancia de normalización de la sonoridad de programa o G_{LN}) se cambia para llevar, por ejemplo, la sonoridad media en todo el programa de audio a un nivel objetivo. Entre los ejemplos de sonoridad de programa se incluye una medición de la potencia media, como el valor RMS, o una medición de la sonoridad percibida por el hombre, como la ofrecida por la Recomendación BS.1770 de la Unión Internacional de Telecomunicaciones (UIT) "Algoritmos para medir la sonoridad de programa y el nivel de audio de pico verdadero". Dependiendo del rango dinámico de la señal de entrada y del nivel objetivo, la normalización de la sonoridad de programa puede generar picos que son más altos que los límites de un medio de grabación o transmisión o reproducción determinados. En tales casos, se puede usar DRC para evitar el recorte en tales casos, lo cual a su vez puede alterar la relación señal/ruido y la dinámica relativa de la señal de audio.

[0014] El DRC puede reducir el volumen de los sonidos fuertes o amplificar los sonidos silenciosos, al reducir o "comprimir" el rango dinámico de una señal de audio. La compresión se usa comúnmente en la grabación, reproducción y transmisión de sonido. Una unidad de hardware electrónico, o software de audio, que se utiliza para aplicar la compresión a veces se denomina compresor. Los compresores a menudo tienen varios controles, que incluyen, por ejemplo, un umbral (por ejemplo, en dB), una relación o cantidad de reducción de ganancia (valor de ganancia), controles de ataque y liberación que varían la velocidad a la que se aplica la compresión y suavizan el efecto, y un control de consola duro/blando.

[0015] Una característica de DRC (como la frase se utiliza aquí) da la relación entre una medición a corto plazo de la sonoridad de una señal de audio de entrada (también denominada aquí "sonoridad[dB]" en la **Fig. 4** o nivel de entrada de DRC en la **Fig. 2**) calculada en un intervalo de tiempo que puede ser del orden de una trama o paquete de la señal de audio digital, por ejemplo, un intervalo de entre 5 y 100 ms, y un valor de ganancia que se puede aplicar a la señal de audio de entrada. Esto también puede denominarse perfil de compresor (donde los valores de ganancia sirven para comprimir o expandir la señal de audio de entrada). Varias características de DRC de ejemplo que se pueden usar en una etapa de codificación se muestran en la **Fig. 2**, mientras que la **Fig. 4** y la **Fig. 6** muestran características de DRC de ejemplo que se pueden usar en una etapa de descodificación (que se describirá más adelante).

[0016] Cualquier proceso adecuado puede ser utilizado para seleccionar la característica de DRC actual en la etapa de codificación. En respuesta a la recepción de la selección, el procesador DRC 4 accede a la característica de DRC almacenada y aplica el perfil al que se accede a la señal de entrada (en este caso, la señal de audio digital normalizada con ganancia) y, por lo tanto, genera valores de ganancia DRC por trama. El procesador DRC 4 puede generar los valores de ganancia DRC de la siguiente manera. La señal de audio normalizada se procesa para calcular una medición de sonoridad a corto plazo, por ejemplo, calculada en el orden de aproximadamente una trama de la señal de audio de entrada. La medición de sonoridad a corto plazo calculada se usa luego como entrada a una tabla de búsqueda que refleja una de las características de DRC, como las que se muestran en la **Fig. 2** o cualquier otra característica de DRC adecuada, para obtener un valor de ganancia de DRC. Por ejemplo, si la trama de audio es suave, es probable que se obtenga una ganancia DRC más positiva, y si es alta, es probable que se obtenga un valor de ganancia DRC más negativo. En un modo de realización, cada trama del audio digital de entrada puede asociarse con un valor de ganancia DRC generado independientemente. Los valores de ganancia de DRC se proporcionan luego como metadatos asociados con la señal de audio codificada, por ejemplo, integrados dentro del mismo flujo de bits. Dicho aparato de codificación puede ser parte de un servidor de medios que responde a una solicitud de, por ejemplo, un programa cliente que se ejecuta en un dispositivo de usuario final, para transmitir el programa de audio (por ejemplo, como parte de un archivo de música, por ejemplo, o un archivo de película) a través del Internet para el dispositivo del usuario final solicitante. De forma alternativa, el flujo de bits de audio codificado se puede proporcionar para el almacenamiento en cualquier forma de medio de grabación adecuado.

[0017] En un modo de realización, el procesador DRC 4 puede considerarse como que funciona en paralelo con el codificador 2, excepto que la entrada al procesador DRC 4 es una versión normalizada de la señal de audio que se introduce en el codificador 2. En el ejemplo de la **Fig. 1**, el valor de normalización G_{LN} puede ser un escaler fijo para un programa de audio completo y que se multiplica con las muestras de señales de la señal de audio digital. El valor

de ganancia de normalización G_{LN} se puede calcular "por adelantado" al pasar la señal de audio (normalizada) a través del procesador 4 de DRC. La ganancia de normalización se puede calcular (mediante el bloque etiquetado G_{LN} en la **Fig. 1**) basándose en un parámetro de sonoridad que se calculó previamente mediante un módulo de medición de audio 6, por ejemplo, la sonoridad de programa. Esto se puede hacer mediante el módulo de medición de audio 6 en un primer paso a través de la señal de audio de entrada, antes de que el procesador DRC 4 realice un segundo paso (a través de una versión normalizada de la señal de audio de entrada). A continuación, el cómputo de G_{LN} se basa, por ejemplo, en una relación de (i) una sonoridad medida del programa de audio digital original (por ejemplo, la sonoridad de programa calculada para todo el programa de audio original) y (ii) un objetivo o sonoridad deseada. Consulte a continuación para obtener detalles adicionales sobre cómo calcular un ejemplo de G_{LN} .

[0018] En un modo de realización adicional, el valor de ganancia de normalización de audio puede seleccionarse de forma automática basándose en un número de valores de sonoridad objetivo predeterminados que están asociados con un número de diferentes tipos de contenido de audio, respectivamente, en respuesta al tipo de contenido de audio en la señal de entrada de audio. Por ejemplo, si el programa de audio es música clásica, entonces se selecciona un valor de sonoridad objetivo diferente que si el programa objetivo fuera música pop, un diálogo o programa de entrevistas o una imagen en movimiento con acción.

[0019] La **Fig. 1** muestra la señal de entrada al módulo de medición de audio 6 que pasa a través de los siguientes tres bloques de procesamiento de señal: un multiplicador (para el ajuste del rango dinámico), un procesador de mezcla descendente 20 (para producir una mezcla descendente de una señal de audio multicanal) y un filtro de acondicionamiento 9. Sin embargo, este es solo un modo de realización que se ilustra en la **Fig. 1**. En otros modos de realización, uno o más de esos tres bloques de procesamiento de señales pueden omitirse. Por ejemplo, si la mezcla descendente no está indicada para la señal de audio (programa de audio), entonces el procesador de mezcla descendente 20 se puede omitir de la etapa de codificación, de modo que la señal de audio se sometería solo al ajuste de rango dinámico y al acondicionamiento (antes de ser procesado por el módulo de medición de audio 6). Como otro ejemplo, el filtro de acondicionamiento 9 puede omitirse, mientras que el ajuste de rango dinámico y mezcla descendente todavía están indicados, en cuyo caso la señal de audio se sometería solo a un ajuste de rango dinámico y una mezcla descendente (antes de ser procesada por el módulo de medición de audio 6).

[0020] Todavía en referencia a la **Fig. 1**, además de los beneficios de la DRC como metadatos, el aparato de codificación puede proporcionar también el índice de la característica de DRC seleccionado en un nuevo campo de transporte de metadatos, como más metadatos (asociada con la señal de audio codificada). En el ejemplo de la **Fig. 1**, se utiliza un canal de datos auxiliar 7 para proporcionar el índice, mientras que las ganancias de DRC están integradas con la señal de audio codificada. El canal auxiliar está separado del canal de datos principal en el que se entrega el audio codificado. Existen varios medios posibles para proporcionar los metadatos asociados con la señal de audio codificada. La **Fig. 1** solo muestra un mecanismo, donde los valores de ganancia DRC se han incorporado (mediante un multiplexor) con el audio codificado en el mismo flujo de bits (el canal de datos "principal"), mientras que los parámetros de sonoridad se suministran a través de un procesador de canal de datos auxiliar. 7. En otros medios, un ensamblador o formateador organiza los valores de ganancia DRC con la señal de audio codificada en tramas o paquetes dentro de un solo archivo, por ejemplo. En otro medio más, se puede agregar un compresor que reduce la velocidad de transmisión de bits de los valores de ganancia de DRC a medida que están integrados con la señal de audio en el mismo archivo o flujo de bits. En otro medio más, el procesador de canal de datos auxiliar 7 organiza los valores de ganancia de DRC en el canal de datos auxiliar (además de otros metadatos).

[0021] Además del índice de la característica de DRC actual, los metadatos pueden incluir también un valor de sonoridad de programa calculado por un módulo de medición de audio 6 en la etapa de codificación, y también, opcionalmente, un valor pico verdadero. Las mediciones de audio realizadas basándose en la señal de audio de entrada para calcular la sonoridad de programa y los valores máximos reales pueden estar de acuerdo con cualquier técnica conocida y adecuada, por ejemplo, de acuerdo con UIT-BS.1770-3. En un modo de realización adicional, como se ilustra en la **Fig. 1**, esta medición de audio puede usar los valores de ganancia de DRC calculados por el procesador DRC 4, para calcular los parámetros de sonoridad (por ejemplo, sonoridad de programa y valores máximos reales) que a continuación se proporcionan como otros metadatos. En otras palabras, se puede calcular un valor de sonoridad de programa y un valor pico verdadero para el programa de audio basándose en la versión ajustada o comprimida por DRC de la señal de audio de entrada (sin la normalización G_{LN}). Tenga en cuenta que, como se ilustra en la **Fig. 1**, la señal de audio comprimida puede ser una señal de audio multicanal que se procesa mediante un procesador de mezcla descendente 20, lo cual da como resultado una señal de audio de mezcla descendente particular que se produce y envía a una entrada del módulo de medición de audio 6. El proceso de mezcla descendente aquí es opcional, ya que el procesador de mezcla descendente 20 puede omitirse u saltarse, de modo que las mediciones de audio se realicen sobre la señal de audio comprimida sin mezcla descendente.

[0022] En un modo de realización adicional, para realizar la medición de audio (en el módulo de medición de audio 6), se utiliza un filtro de acondicionamiento opcional 9 que se ha configurado basándose en 1) una función o característica de la señal de audio de entrada y/o 2) una característica de un dispositivo de reproducción de usuario final que se espera que reciba y realice una reproducción del flujo de bits de audio codificado. Como se ve en la **Fig. 1**, la señal de audio de entrada puede ser preacondicionada por el filtro 9, antes de entrar en el módulo 6, donde se realizará un cálculo de sonoridad. El filtro de acondicionamiento 9 puede ser, por ejemplo, un filtro de paso alto, un

filtro de paso de banda o una combinación de los mismos, que modela la respuesta de frecuencia o la función de transferencia de las etapas digital a analógica y de altavoz (transductor) de un dispositivo de reproducción particular, por ejemplo, un dispositivo electrónico de consumo, como un ordenador de escritorio, un ordenador portátil, una tablet, un teléfono inteligente o un reproductor de medios digitales diseñado para recibir contenido digital a través de Internet y realizar la reproducción a través de un monitor de pantalla adjunto.

[0023] Volviendo ahora a la **Fig. 3**, este es un diagrama de bloques de componentes pertinentes de una etapa de descodificación o un aparato de descodificación. El diagrama de bloques en la **Fig. 3** proporciona soporte no solo para un aparato descodificador de audio digital o etapa de descodificación, sino también para un procedimiento de descodificación de audio, como se explica a continuación. El aparato descodificador mostrado tiene un descodificador 10 que recibe una señal de audio codificada que ha sido extraída de un flujo de bits de audio codificado por un demultiplexor 13. El flujo de bits puede incluir metadatos incorporados, por ejemplo, valores de ganancia DRC, un índice de la característica de DRC (utilizado en la etapa de codificación) y parámetros de sonoridad. Como se sugirió anteriormente con respecto al aparato de codificación, una alternativa es incluir algunos o todos los metadatos en un canal de datos auxiliares separado 7 (ver **Fig. 1**). En ambos casos, se proporciona una señal de audio codificada como entrada al descodificador 10 que, en respuesta, produce una señal de audio descodificada. El descodificador 10 puede ser complementario al codificador 2 de la **Fig. 1** para poder recuperar una señal que puede ser esencialmente similar a la señal de audio de entrada original de la etapa de codificación. La señal de audio descodificada pasa a través de los bloques de procesamiento de señal opcionales 14, 16 hasta que llega a un convertidor digital a analógico (DAC) 18 y un amplificador de potencia (no mostrado), donde se convierte en forma analógica y se aplica para activar un transductor (altavoz) 19.

[0024] El aparato de descodificación también tiene un procesador DRC_1 12 que recibe los metadatos asociados con la señal de audio codificada, en el que los metadatos incluyen los valores de ganancia de DRC que fueron calculados en la etapa de codificación. El procesador DRC_1 12 puede modificar esos valores de ganancia para producir valores de ganancia nuevos o modificados. Esta modificación puede basarse en parámetros locales, que incluyen la entrada del usuario y/o el rango dinámico de un transductor 19 (por ejemplo, un altavoz incorporado o un altavoz externo, inalámbrico o con cable), amplificador de potencia (no mostrado) y una combinación de convertidor digital a analógico (DAC) 19, que se está utilizando para la reproducción de la señal de audio descodificada. Los valores de ganancia modificados se aplican luego a la señal descodificada (este ajuste se representa con el símbolo del multiplicador), antes de pasar a la mezcla opcional y al bloque de procesamiento de audio adicional (bloques 14, 16), y luego al DAC 18. Los valores de ganancia de DRC modificados se aplican así a la señal de audio descodificada para producir la llamada señal de audio ajustada por DRC de la etapa de descodificación. Este último se puede combinar mediante un mezclador 14 con otras señales de audio ajustadas por DRC en la etapa de descodificación de otras fuentes de audio como se muestra (es decir, a través de otras unidades de ajuste de rango dinámico 15) antes de enviarlas al DAC 18.

[0025] En un modo de realización, los metadatos recibidos en la etapa de descodificación incluyen un índice de una característica de DRC actual o previamente seleccionada (de acuerdo con el cual los valores de ganancia de DRC se calcularon en una etapa de codificación, como en la **Fig. 1**). En un ejemplo, una copia de la característica de DRC seleccionada o actual puede almacenarse en el procesador DRC_1 12 y está asociada con dicho índice. Esto se conoce como una característica de DRC de la etapa de codificación. Por el contrario, el procesador DRC_1 12 también puede contener varias características denominadas DRC de etapa de descodificación, como se muestra. Uno de estos últimos se utiliza para producir los valores de ganancia nuevos o modificados, de la siguiente manera.

[0026] En un modo de realización, el procesador DRC_1 12 "invierte" las operaciones realizadas por el procesador DRC 4 de la etapa de codificación, a fin de obtener una sonoridad a corto plazo o nivel de entrada DRC (por ejemplo, en dB) a partir de la característica de DRC actual o seleccionada que se almacena en el procesador 12, aplicando un valor de ganancia de DRC recibido (de los metadatos recibidos) a la característica de DRC actual. Este valor de sonoridad a corto plazo recuperado se utiliza luego como entrada en una de las características de DRC de la etapa de descodificación seleccionada, para obtener un valor de ganancia nuevo o modificado. Aquí se hace referencia a este último como una etapa de descodificación o valor de ganancia de DRC del descodificador. Cada una de las características de DRC de la etapa de descodificación y de la etapa de codificación se puede almacenar en forma de tabla de búsqueda.

[0027] La selección de una característica de DRC de etapa de descodificación puede estar de acuerdo con uno o más de los siguientes: contexto de usuario (incluyendo por la noche, caminar, correr, en el vehículo o en el coche, y el auricular vs. altavoz incorporado), y rango dinámico de la ruta de la señal del altavoz. El aparato de descodificación de la **Fig. 3** en ese caso puede estar integrado dentro de un dispositivo de usuario final, como un teléfono móvil, un ordenador portátil o una tablet que a menudo se encuentra en contextos de usuario bastante diferentes. Dichos contextos pueden ser detectados por hardware y software de sensores electrónicos conocidos, que incluyen luz ambiental y sensores de posición, orientación o movimiento. El rango dinámico de la trayectoria de la señal del transductor puede determinarse de antemano por un fabricante del dispositivo de usuario final y puede almacenarse o proporcionarse de otra manera al procesador DRC_1 12 para su uso durante el proceso de descodificación. El rango dinámico del transductor puede incluir un rango dinámico del DAC 18. En el caso de que haya disponible un rango dinámico mayor, el procesador DRC_1 12 puede seleccionar una característica de DRC de la etapa de descodificación

que permite que se retenga más rango dinámico, en el rango esperado de sonoridad a corto plazo de la señal de audio descodificada.

5 [0028] En otro modo de realización, también representado en la Fig. 3, los metadatos recibidos en una etapa de descodificación pueden incluir además un número de parámetros de sonoridad que pueden ser uno o más de los siguientes: sonoridad de los programas, pico verdadero, rango de sonoridad, sonoridad máxima momentánea y sonoridad a corto plazo. El procesador DRC_1 puede utilizar uno o más de estos valores al generar el valor de ganancia DRC modificado o nuevo (por trama).

10 [0029] Como se describió anteriormente, el procesador DRC_1 12 en la etapa de descodificación (también denominado procesador DRC de descodificador) puede generar sus valores de ganancia DRC modificados usando una característica de DRC de etapa de descodificación seleccionada. Se sugirió anteriormente que esta última puede ser una de varias características de DRC predefinidas que se almacenan en el procesador 12 y a las que se pueden acceder de acuerdo con una entrada de usuario o una señal de control de contexto de usuario. La Fig. 4 ilustra varios ejemplos de las características de DRC de la etapa de descodificación que se pueden usar. Estos incluyen uno para uso nocturno, otro para uso en coche o en vehículo, y otro para un altavoz o transductor 19 que tiene un rango dinámico limitado.

20 [0030] De acuerdo con otro modo de realización de la invención, una combinación de componentes de la Fig. 1 y la Fig. 3 puede producir un sistema para la codificación y descodificación de una señal de audio. Una serie de características de control del rango dinámico del codificador (DRC) se almacenan en la etapa del codificador. Cada una de estas relaciona los valores de ganancia de DRC con los valores de sonoridad a corto plazo, donde cada valor de sonoridad puede referirse a la potencia media de una señal de audio en un intervalo predeterminado, por ejemplo, en el orden de aproximadamente una trama de audio digital, por ejemplo, entre 5 y 100 milisegundos de longitud. Una de las características de DRC del codificador se selecciona y luego se utiliza para producir valores de ganancia de la etapa del codificador. Estos se proporcionan luego como metadatos asociados con una señal de audio codificada. Varios ejemplos de tal etapa de codificador se dieron anteriormente y se describieron en relación con la Fig. 1.

30 [0031] La etapa del descodificador tiene un procesador que almacena no solo las características de DRC del codificador, sino también una serie de características de DRC del descodificador. Cada uno de estos últimos relaciona los valores de ganancia de DRC con los valores de sonoridad a corto plazo, que pueden ser similares a los calculados en la etapa del codificador para su uso en la determinación de los valores de ganancia de DRC del codificador. El procesador en la etapa del descodificador es para calcular los valores de ganancia de la etapa del descodificador utilizando los valores de ganancia de la etapa del codificador de la etapa del codificador. En un caso particular, el procesador en la etapa del descodificador utiliza un índice de una característica de DRC del codificador seleccionado junto con los valores de ganancia de la etapa del codificador basado en metadatos, para calcular sus valores de ganancia de la etapa del descodificador. Los ejemplos de esto se dieron anteriormente y se describieron en relación con la Fig. 3. En última instancia, la etapa del descodificador puede (cuando se usa como parte del procesamiento de reproducción en, por ejemplo, un dispositivo de usuario final) descodificar la señal de audio codificada y luego aplicar los valores de ganancia de la etapa del descodificador para lograr un ajuste de rango dinámico deseado (sobre la señal de audio descodificada).

45 [0032] De acuerdo con todavía otro modo de realización de la invención, un procedimiento para proporcionar audio codificado y metadatos asociados implica mezcla descendente de una señal de audio multicanal. Como se ve en la Fig. 1, un conjunto de parámetros de sonoridad se calcula mediante un módulo de medición de audio 6 basado en la señal de audio digital multicanal que ha pasado a través de un procesador de mezcla descendente 20, donde los parámetros describen el perfil de sonoridad de una mezcla descendente de la señal de audio multicanal. Por ejemplo, una señal de audio de seis o siete canales puede ser procesada por el procesador de mezcla descendente 20 en una señal estéreo o de dos canales; más en general, una mezcla descendente es una conversión de una señal de audio de N canales en uno que tiene M canales, donde N es mayor que M.

55 [0033] La señal de audio multicanal se codifica y se proporciona la señal codificada junto con los metadatos asociados con la misma, donde los metadatos en este caso incluyen el conjunto calculado de los parámetros de sonoridad que describen el perfil de sonoridad de la mezcla descendente. Los metadatos también pueden incluir un diseño de canal base de la señal de audio multicanal. Por ejemplo, el diseño del canal base puede proporcionar detalles de una señal de audio multicanal de sonido envolvente 5.1 mediante la identificación de cada uno de los seis canales, por ejemplo, frontal central, frontal derecho, frontal izquierdo, envolvente izquierdo, envolvente derecho y subwoofer. La señal de audio multicanal codificada y sus metadatos asociados pueden luego ser recibidos por una etapa de descodificación (vea la Fig. 5 como se describe a continuación), en la cual la señal de audio codificada se descodifica y procesa para producir una mezcla descendente, de acuerdo con el conjunto de parámetros de sonoridad y la disposición del canal base incluidos en los metadatos. Los metadatos también podrían incluir las ganancias de mezcla descendente que se usaron en la etapa de codificación para calcular la mezcla descendente particular para la cual se incluyeron los parámetros de sonoridad en los metadatos. Tenga en cuenta que en un modo de realización de este tipo, no es necesario realizar ningún ajuste para el rango dinámico en la etapa de descodificación (durante la reproducción a través de un altavoz local). Dicho esquema puede implementarse, por ejemplo, adoptando la convención donde, si los

metadatos recibidos en la etapa de descodificación no contienen valores de ganancia DRC, entonces no se realiza ningún ajuste de DRC sobre la señal descodificada.

[0034] La **Fig. 5** representa otro modo de realización de la invención en la que los metadatos asociados con una señal de audio multicanal codificada podrían incluir además valores de ganancia de DRC, además de parámetros de sonoridad tales como el pico verdadero y la sonoridad de programa con respecto a una mezcla descendente. Esta figura representa una etapa de descodificación de alguna manera similar a la de la **Fig. 3** en el sentido de que puede contener un desmultiplexor 13 seguido de un descodificador 13, un procesamiento de audio adicional opcional 16 y el DAC 18. También en este caso, los metadatos recibidos asociados con la señal de audio de codificación proporcionan valores de ganancia de DRC. Deben usarse en la etapa de descodificación para realizar ajustes de rango dinámico en la señal de audio descodificada, antes de procesar la señal de audio descodificada para producir la mezcla descendente. En otras palabras, el descodificador 10 descodifica una señal de audio multicanal codificada, y luego uno o más de los canales descodificados se ajustan con los nuevos valores de ganancia de DRC o con los valores de ganancia de DRC originales (estos últimos se han obtenido del desmultiplexor 13). Esto ocurre antes de que un módulo 20 de mezcla descendente produzca una mezcla descendente particular. El módulo de mezcla descendente 20 puede usar valores de ganancia de mezcla descendente que también se extraen de los metadatos (como se sugirió anteriormente). En otro modo de realización, la señal de audio descodificada se ajusta para el rango dinámico utilizando nuevos valores de ganancia que son calculados por el procesador DRC_1 12 utilizando, por ejemplo, procedimientos similares a los descritos anteriormente en relación con la **Fig. 3**. Los parámetros locales que pueden afectar a los nuevos valores de ganancia aquí pueden ser similares e incluir la configuración de volumen del usuario y un interruptor de activación/desactivación de la normalización de la sonoridad. Estos parámetros locales pueden gobernar la selección de una característica particular de varias características de DRC de etapa de descodificación, para producir nuevos valores de ganancia de DRC.

[0035] Como una alternativa al enfoque descrito en el párrafo anterior, los valores de ganancia de DRC producidos por el procesador DRC_1 12 (en la etapa de descodificación) pueden diseñarse para aplicarse a la salida del módulo de mezcla descendente 20. Para ilustrar un modo de realización de este tipo, la etapa de descodificación en la **Fig. 5** se puede modificar de modo que la entrada del multiplicador que se muestra allí (cuya otra entrada reciba los valores de ganancia DRC del procesador DRC_1 12) se coloque a la salida del módulo de mezcla descendente 20 (y a la entrada de la DAC 18). El procesador DRC_1 12 en este caso puede producir valores de ganancia de DRC modificados basados en los valores de ganancia de DRC originales recibidos a través de metadatos (desde una etapa de codificación), o puede pasar a través de los valores de ganancia de DRC originales, al bloque multiplicador o ajustador. Una etapa de codificación correspondiente, para producir la forma particular de los valores de ganancia DRC originales para este caso, puede ser como se muestra en la **Fig. 7**.

[0036] Con referencia a la **Fig. 7**, se muestra un aparato codificador que puede ser similar al de la **Fig. 1**, excepto al menos en el siguiente aspecto. Aquí, las ganancias de DRC se calculan (mediante el procesador 4 de DRC) de tal manera que luego se pueden aplicar directamente, en el aparato de descodificador, a una mezcla descendente. En contraste con la **Fig. 1**, el procesador DRC 4 de la **Fig. 7** también funciona con una señal de mezcla descendente, que es producida por un procesador de mezcla descendente 20 que recibe la señal de audio de entrada multicanal original y en respuesta alimenta al normalizador (o multiplicador) cuyo otra entrada recibe el valor de ganancia de normalización G_{LN} . El codificador 2 aún podría procesar la señal de audio de entrada original de la misma manera que en la **Fig. 1**, para generar la señal codificada.

[0037] Volviendo a la **Fig. 5**, la etapa de descodificación en el mismo puede recibir su señal de audio codificada y metadatos desde una etapa de codificación (no mostrada) en la que se produjo una mezcla descendente particular de una señal de audio de entrada multicanal. Los valores de ganancia de DRC actuales se calcularon en la etapa de codificación, utilizando la señal de audio multicanal de entrada y/o la mezcla descendente particular de la misma, y luego se proporcionaron como metadatos junto con la versión codificada de la señal de audio multicanal (por ejemplo, al ensamblar la señal de audio codificada con los valores de ganancia de DRC actuales como metadatos dentro de un único flujo de bits o archivo codificado, utilizando por ejemplo un multiplexor 8 como en la **Fig. 1**). Otros medios para proporcionar la señal de audio codificada junto con un conjunto calculado de parámetros de sonoridad y un diseño de canal base, y opcionalmente los parámetros DRC actuales (todos los cuales son metadatos asociados con la señal de audio codificada) son posibles. Entre los ejemplos se incluye un ensamblador o formateador que organiza los valores de ganancia DRC junto con la señal de audio en tramas o paquetes dentro de un solo archivo, un multiplexor que produce un flujo de bits que contiene los valores de ganancia DRC que han sido multiplexados en el tiempo con tramas de la señal de audio codificada, un compresor que reduce la velocidad de datos de los valores de ganancia de DRC integrados con la señal de audio codificada en el mismo archivo o en el mismo flujo de bits, y un procesador de canal de datos auxiliar que organiza los valores de ganancia de DRC en un canal de datos auxiliar que está separado de un canal de datos principal (donde este último contiene la señal de audio codificada que se entrega a la etapa de descodificación de la **Fig. 5**).

[0038] Tenga en cuenta que, como se sugirió anteriormente, si los metadatos asociados con la señal de audio multicanal codificada no especifican la realización de ajustes de rango dinámico, o simplemente no contienen valores de ganancia DRC o parámetros DRC (por ejemplo, sin mencionar un índice válido de la característica de DRC), a continuación la señal de audio descodificada se procesa para producir la mezcla descendente pero sin realizar ningún

ajuste de rango dinámico en la señal de audio descodificada. Esto podría ser como si en la **Fig. 5**, el procesador DRC_1 12 proporcionara que todos sus nuevos valores de ganancia con un valor de 0 dB, lo cual significa que no se realizará ningún ajuste en el rango dinámico de la señal de audio descodificada.

5 **[0039]** En el caso de que el procesador DRC_1 12 no reciba parámetros DRC como metadatos, los nuevos valores de ganancia pueden ser producidos por el procesador 12 de acuerdo con una característica de DRC de etapa de descodificación que puede seleccionarse a partir de los ejemplos representados en la **Fig. 6** (u otras características adecuadas de DRC). La **Fig. 6** muestra tres características diferentes de DRC donde cada una está asociada con una configuración o nivel de volumen de usuario diferente. Como se puede ver, a medida que aumenta el volumen, la cantidad de compresión definida por la característica de DRC aumenta cuando aumenta la sonoridad a corto plazo.

10 **[0040]** Los ejemplos de códecs que pueden beneficiarse de las técnicas descritas aquí incluyen estándares de MPEG y ATSC como AAC y AC-3, aunque otros estándares o enfoques que contienen mecanismos para controlar el volumen y el rango dinámico de una señal de audio descodificada también pueden beneficiar.

15

Ejemplo de mediciones de audio que pueden almacenarse como metadatos

20 **[0041]** El módulo de medición de audio 6 puede ser una rutina de software que va a ser ejecutada por un procesador, o una disposición de circuitos de lógica de procesamiento de audio digital cableada, que calcula o proporciona uno o más parámetros de sonoridad para un archivo de audio digital dado. La rutina se puede utilizar en una variedad de productos de audio, como reproductores multimedia, para la normalización de la sonoridad del contenido musical. Las mediciones de audio calculadas se pueden almacenar como metadatos en un archivo de audio codificado, durante un proceso de codificación de audio digital. Por ejemplo, MPEG actualmente proporciona campos de flujo de bits en los que se pueden almacenar dichos metadatos. Entre los usos actuales de dichos campos se incluye el almacenamiento de la sonoridad de referencia, las ganancias del control de rango dinámico (DRC) por trama de audio digital y los factores de ponderación de mezcla descendente. De acuerdo con un modo de realización de la invención, se define una nueva "caja" en la "extensión de descripción de la muestra" del marco los sistemas de audio MPEG-4, para almacenar los metadatos (como se describe más adelante).

30 **[0042]** La sonoridad de programa es una medición de audio que puede ser una estimación de la sonoridad media de todo el contenido de un archivo de audio digital. Se puede calcular un ejemplo de acuerdo con ITU-BS.1770-3. La sonoridad de programa se puede calcular en una etapa de codificación, después de haber aplicado la compresión del rango dinámico a una señal de contenido de audio, por ejemplo, ver la **Fig. 1** donde la señal de audio no normalizada se ajusta para el rango dinámico antes de ser precondicionada por el filtro 9 y alimentada al módulo de medición de audio 6 (mientras se salta el procesador de mezcla descendente 20). En otro modo de realización, la sonoridad de programa puede ser calculada por el módulo de medición de audio 6 en una etapa de codificación, para una mezcla descendente de la señal de contenido de audio, por ejemplo, ver la **Fig. 1**.

35 **[0043]** El valor pico real es una medición de audio que puede ser la magnitud de muestreo máxima de un flujo de bits de audio del archivo de audio (por ejemplo, a una frecuencia de sobremuestreo 4x). Se puede calcular un ejemplo según ITU-BS.1770-3.

40 **[0044]** El rango de sonoridad puede ser una medición de audio basada en ITU BS.1770 o de acuerdo con una especificación de la Unión Europea de Radiodifusión (EBU). Mide la distribución estadística de la sonoridad para un tamaño de bloque determinado de audio digital (por ejemplo, bloques de 400 ms) y genera la diferencia de un percentil bajo y alto de la distribución de sonoridad para describir el rango dinámico. Otras mediciones de audio que indican el rango de sonoridad son posibles.

Mejoras de metadatos

50 **[0045]** Un modo de realización de la invención aquí es una nueva "caja" en la parte "extensión de descripción de muestra" del marco de sistemas MPEG-4 que puede estar llena con metadatos estáticos para cada pista (programa de audio), por ejemplo, sonoridad de programa, sonoridad de anclaje, pico real y rango de sonoridad. El contenido adicional del programa por pista o por audio del nuevo cuadro dentro de la Extensión de Descripción de Muestra MPEG-4 puede incluir: máx. sonoridad momentánea, como en una ventana de 0,4 segundos, una sonoridad máxima a corto plazo, como en una ventana de 0,3 segundos, asignación de canales que define la disposición de canales para sistemas de reproducción, incluidos los canales de altura y otros, asignación de canales DRC, índice de características de DRC, coeficientes de mezcla descendente, sonoridad de programa de la mezcla descendente estéreo, sonoridad de anclaje la mezcla descendente estéreo y pico verdadero de la mezcla descendente estéreo. Mientras que otros canales de datos auxiliares para pasar los metadatos al procesador de reproducción son posibles como se describió anteriormente, el enfoque particular aquí puede tener las siguientes ventajas; los metadatos estáticos están disponibles sin descodificar el flujo de bits de audio; adición de sonoridad de anclaje (también conocida como dialnorm) para soportar la normalización del volumen de contenido de películas/TV; el conocimiento de la característica de DRC utilizada en el codificador puede ayudar a predecir el efecto de las ganancias de DRC; el conocimiento de la característica de DRC se puede utilizar para modificar la característica de DRC en el descodificador; los coeficientes

65

de mezcla descendente se pueden definir de una manera preparada para el futuro que puede soportar formatos de audio multicanal superiores a 5.1; y un mejor control sobre el recorte y la sonoridad de mezcla.

Uso de metadatos

5 **[0046]** El uso de la sonoridad de programa o la sonoridad de anclaje es adecuado para la normalización de la sonoridad. La sonoridad de anclaje en general se basa en segmentos de voz extraídos y puede aplicarse solo al contenido de películas/programas de televisión.

10 **[0047]** En cuanto al control de rango dinámico (DRC), varios valores de metadatos pueden hacerse disponibles que describen aspectos del rango dinámico del contenido de audio grabado (véase la tabla siguiente). El tamaño del rango dinámico puede ser útil para ajustar el DRC durante la reproducción, por ejemplo, el DRC es menos agresivo si el rango dinámico es pequeño o incluso se puede desactivar el DRC. Además, se puede establecer un rango dinámico objetivo dependiendo de la entrada del usuario, la configuración de volumen y el rango dinámico DAC y el rango dinámico del altavoz, y seleccionar una característica de DRC para que el rango se reduzca al objetivo. Esto también puede tener en cuenta una limitación de rango dinámico razonable para espacios más pequeños (entornos de escucha). Los valores pico verdadero y de sonoridad máxima pueden ser útiles para estimar el margen, por ejemplo, cuando la normalización de sonoridad produce una ganancia positiva [dB] o cuando se necesita margen para evitar el recorte de la mezcla descendente. A continuación, la característica de DRC se puede ajustar para acercarse a un objetivo de altura.

[0048] Ejemplo de metadatos que describe aspectos del rango dinámico

Rango de sonoridad	Tamaño del rango derivado del histograma de sonoridad
Sonoridad de programa	Este valor se puede combinar con uno de los valores pico a continuación para estimar el rango dinámico
Pico verdadero	Podría usar la diferencia (Pico verdadero) - (Sonoridad de programa) como estimación de rango dinámico. El pico verdadero puede cambiar significativamente después de que el contenido de audio original o el programa de audio haya pasado por el códec
Sonoridad momentánea máx. (por ej., 0,4 s)	Una descripción del pico de sonoridad, que no cambiará mucho después del códec. El tamaño de la ventana puede ser compatible con ITU BS.1770. Reste la sonoridad de programa para obtener una estimación del rango dinámico.
Sonoridad máx. a corto plazo (por ejemplo, 3 s)	Una descripción del pico de sonoridad, que no cambiará mucho después del códec. El tamaño de la ventana es más largo que el de ITU BS.1770. Reste la sonoridad de programa para obtener una estimación del rango dinámico.

25 **[0049]** Un procesador DRC se usa en la etapa de codificador para generar valores de ganancia utilizando una seleccionada de las características de DRC predefinidas. El índice de la característica de DRC seleccionada puede transmitirse en la nueva caja MPEG-4. Los valores de ganancia (por trama) pueden transmitirse en campos existentes (durante la compresión ligera y/o pesada).

30 **[0050]** Como se ve por ejemplo en la Fig. 1. la ganancia de normalización de sonoridad de programa G_{LN} se aplica en la entrada al procesador DRC 4 para asegurar la alineación nivel apropiado con respecto a la característica de DRC seleccionada. La ganancia de normalización se puede calcular basándose en el valor de sonoridad de programa L_{PL} que es producido por el módulo de medición de audio 6, y del valor de sonoridad objetivo DRC L_{CTL} [dB] como $G_{LN} = L_{CTL} - L_{PL}$ [dB] con, por ejemplo, $L_{CTL} = -31$ dB.

35 **[0051]** De acuerdo con un modo de realización de la invención, los valores de ganancia de la DRC extraídos se cambian en la etapa de descodificación, para en efecto alcanzar DRC personalizado que puede adaptarse a varias condiciones, por ejemplo cambiando a una característica de DRC diferente (de la que fue utilizada en la etapa de codificación). El procesamiento en la reproducción ahora tiene conocimiento de qué característica de DRC se aplicó en la etapa de codificación, en virtud de poder entender el significado del índice extraído. Entre las condiciones locales que pueden justificar tales cambios se incluyen: modo nocturno; ambiente ruidoso (por ejemplo, ruido dentro de un coche en movimiento); limitaciones del sistema de reproducción (por ejemplo, un altavoz interno de un ordenador portátil, tablet o teléfono inteligente a diferencia de un altavoz externo o auriculares); preferencia del usuario; y rango dinámico del contenido. Consulte la Fig. 4 para ver ejemplos de las características de DRC útiles en tales situaciones.

40 **[0052]** En un modo de realización, las características de DRC disponibles deben basarse en los niveles de entrada/salida en estado estacionario del compresor, para una entrada sinusoidal a 1 kHz. Esto mantiene la compatibilidad con los compresores que utilizan la estimación de la sonoridad ponderada k. Se supone aquí que la característica de DRC se aplica a la señal de audio normalizada de sonoridad. Esto es importante para tener la banda muerta de DRC en el nivel correcto (si corresponde) y produce resultados más consistentes para el contenido con varios niveles de sonoridad, especialmente si dicho contenido se reproduce con la normalización de sonoridad activada.

Mezcla descendente

5 **[0053]** La mezcla descendente se refiere a la manipulación de audio donde se mezclan varios canales de audio distintos para producir un número menor de canales. La mezcla descendente puede ser controlada aquí por la facilidad de producción del programa de audio si es necesario. Por ejemplo, algunos contenidos pueden requerir una mayor atenuación de los canales envolventes antes de mezclarlos, para mantener la inteligibilidad.

10 **[0054]** En la actualidad DVB y MPEG requieren el uso de DRC cuando se genera una mezcla descendente, si se ha establecido DRC_presentation_mode. Esto puede dar como resultado una pérdida de rango dinámico en la mezcla descendente. En contraste, para mantener el rango dinámico cuando sea apropiado, un modo de realización de la invención aquí es un esquema adaptativo donde la compresión DRC solo se requiere para la mezcla descendente durante el alto volumen de reproducción como se muestra en la **Fig. 5**. La aplicación de DRC se adapta para reducir solo el nivel de los segmentos más ruidosos (ver, por ejemplo, la **Fig. 6**). Además, el valor pico verdadero de la mezcla descendente estéreo, que puede haberse calculado en la etapa de codificación mediante el módulo de medición de audio 6 (ver **Fig. 1**) y que se proporcionó como metadatos, ahora puede extraerse en la etapa de descodificación y usarse para calcular cuánta compresión DRC se necesita para aplicar a la señal descodificada antes de la mezcla descendente. Tenga en cuenta que aunque la **Fig. 5** muestra un modo de realización en la que los valores de ganancia DRC del procesador 12 se aplican a la señal descodificada antes de transferir la señal descodificada al procesador de mezcla descendente 20, una alternativa aquí es aplicar los valores de ganancia DRC después de la mezcla descendente, por ejemplo, directamente a la salida del procesador de mezcla descendente 20.

25 **[0055]** Si se utiliza la mezcla descendente y la compresión DRC es independientemente activa, la característica de DRC puede modificarse si es necesario para que se logre suficiente espacio libre para la mezcla descendente. Esta solución proporciona más flexibilidad. Además, la mezcla estéreo descendente se puede normalizar para la sonoridad en la etapa de descodificación, utilizando, por ejemplo, los valores ponderados K en relación con los valores de escala completa (LKFS) (que se recibieron como metadatos). Estos valores LKFS son parámetros de sonoridad que se calcularon en la etapa de codificación mediante el módulo de medición de audio 6 basándose en una versión de mezcla descendente de la señal de audio digital multicanal original (consulte la **Fig. 1**). Esto garantiza que el contenido multicanal y el contenido estéreo se reproduzcan en el mismo nivel, a través de un sistema estéreo.

35 **[0056]** Como se explicó anteriormente, un modo de realización de la invención puede ser un medio legible por máquina (tal como la memoria microelectrónica) que tiene almacenado en el mismo instrucciones, que programan uno o más componentes de procesamiento de datos (denominados genéricamente aquí como un "procesador") para realizar las operaciones de procesamiento de audio digital descritas anteriormente, incluyendo codificación, descodificación, mediciones de sonoridad, filtrado, mezcla, adición, inversión, comparaciones y toma de decisiones. Tales instrucciones pueden ser parte de un programa de aplicación de reproductor de medios. En otros modos de realización, algunas de esas operaciones pueden ser realizadas por componentes de hardware específicos que contienen lógica cableada (por ejemplo, bloques de filtros digitales dedicados, máquinas de estado). Dichas operaciones pueden realizarse de forma alternativa mediante cualquier combinación de componentes de procesamiento de datos programados y
40 componentes de circuitos fijos cableados.

REIVINDICACIONES

1. Un procedimiento para descodificar audio durante el procesamiento de reproducción, que comprende:

5 recibir una señal de audio codificada;

recibir metadatos asociados con la señal de audio codificada, con los metadatos incluyen una pluralidad de valores de ganancia control de rango dinámico, DRC, y un índice de una característica de DRC previamente seleccionada de acuerdo con la cual los valores de ganancia de DRC se calcularon cuando se estaba generando la señal de audio codificada;

10 descodificar la señal de audio codificada para producir una señal de audio descodificada;

seleccionar una característica de DRC actual de entre una pluralidad de características de DRC disponibles, de acuerdo con el índice, y

15 producir nuevos valores de ganancia DRC basados en la característica de DRC actual

i) aplicando la pluralidad de valores de ganancia DRC de los metadatos a la característica de DRC actual, invirtiendo así las operaciones realizadas previamente por un procesador DRC en una etapa de codificación para obtener una pluralidad de niveles de entrada de DRC o sonoridad a corto plazo,

20 ii) seleccionando una característica de DRC de la etapa de descodificación de entre una pluralidad de características de DRC de la etapa de descodificación, y

25 iii) aplicar la pluralidad de niveles de sonoridad a corto plazo o entrada de DRC, obtenida en i), a la característica de DRC de la etapa de descodificación seleccionada para generar los nuevos valores de ganancia de DRC; y

30 aplicar los nuevos valores de ganancia DRC a la señal de audio descodificada para producir una señal de audio ajustada por DRC durante el procesamiento de la reproducción.

2. El procedimiento según la reivindicación 1, en el que los metadatos recibidos incluyen además una pluralidad de valores seleccionados del grupo que consiste en: sonoridad de programa, pico verdadero, rango de sonoridad, sonoridad momentánea máxima y valores de sonoridad a corto plazo.

3. El procedimiento según la reivindicación 1, en el que seleccionar la característica de DRC de la etapa de descodificación de entre la pluralidad de características de DRC disponibles se basa además en uno o más de los siguientes: contextos de usuario; y el rango dinámico de la ruta de la señal del hablante, en el que los contextos del usuario incluyen por la noche, caminar, correr y el coche.

4. Un aparato descodificador de audio digital, que comprende:

45 un dispositivo de reproducción de medios digitales que tiene un descodificador (10), un procesador DRC (12) y un multiplicador, con el descodificador para recibir una señal de audio codificada y producir una señal de audio descodificada;

el procesador DRC para recibir los metadatos que están asociados con la señal de audio codificada, en el que los metadatos incluyen una pluralidad de valores de ganancia de control de rango dinámico, DRC, y un índice de una característica de DRC previamente seleccionada de acuerdo con la cual se calcularon los valores de ganancia DRC cuando se estaba generando la señal de audio codificada, con el procesador DRC para seleccionar una característica de DRC actual, de entre una pluralidad de características de DRC almacenadas, de acuerdo con el índice, y producir nuevos valores de ganancia DRC basados en

50 i) aplicando la pluralidad de valores de ganancia DRC de los metadatos a la característica de DRC actual, invirtiendo así las operaciones realizadas previamente por un procesador DRC en una etapa de codificación para obtener una pluralidad de niveles de entrada de DRC o sonoridad a corto plazo,

55 ii) seleccionar una característica de DRC de la etapa de descodificación de entre una pluralidad de características de DRC de la etapa de descodificación almacenada, y

60 iii) aplicar la pluralidad de niveles de sonoridad a corto plazo o entrada de DRC, obtenida en i), a la característica de DRC de la etapa de descodificación seleccionada para generar los nuevos valores de ganancia de DRC; y

65

en el que el multiplicador tiene que aplicar los nuevos valores de ganancia de DRC a la señal de audio descodificada para producir una señal de audio ajustada por DRC.

- 5 **5.** El aparato de la reivindicación 4, en el que el descodificador (10), el procesador DRC (12) y el multiplicador son parte de un dispositivo de usuario final que comprende además un convertidor digital a analógico, DAC (18), para convertir la señal de audio ajustada por el DRC. en forma analógica durante la reproducción de la señal de audio codificada.
- 10 **6.** El aparato de la reivindicación 5, que comprende además un procesador de mezcla descendente (20) para realizar una conversión de mezcla descendente sobre la señal de audio ajustada por DRC, antes de la conversión en forma analógica, basándose en los valores de ganancia de mezcla descendente extraídos de los metadatos.
- 15 **7.** El aparato de la reivindicación 6, en el que el procesador DRC (12) selecciona la característica de DRC de la etapa de descodificación de entre la pluralidad de características de DRC almacenadas basándose en uno o más de los siguientes: configuración del volumen de reproducción; contexto del usuario, incluyendo por la noche, caminando, corriendo o en auto; rango dinámico DAC; y rango dinámico del altavoz.
- 20 **8.** El aparato de la reivindicación 6, en el que el procesador DRC (12) extrae de los metadatos un valor pico verdadero de una mezcla estéreo descendente de la señal de audio codificada, y varía la aplicación de la compresión DRC que se aplicará a la señal descodificada antes de la conversión de mezcla descendente en función del valor pico verdadero, para evitar recortes en el altavoz.
- 25 **9.** El aparato de la reivindicación 5, en el que el procesador DRC (12) selecciona la característica de DRC de la etapa de descodificación de entre la pluralidad de características de DRC almacenadas basándose en uno o más de los siguientes: configuración del volumen de reproducción; contexto del usuario, incluyendo por la noche, caminando, corriendo o en auto; rango dinámico DAC; y rango dinámico del altavoz.
- 30 **10.** El procedimiento según la reivindicación 1, que comprende además realizar una conversión de mezcla descendente sobre la señal de audio ajustada por DRC, antes de la conversión de la señal ajustada de DRC en forma analógica, basándose en los valores de ganancia de mezcla descendente extraídos de los metadatos.
- 35 **11.** El procedimiento según la reivindicación 10, en el que la selección de la característica de DRC de la etapa de descodificación de entre la pluralidad de características de DRC almacenadas se basa en la determinación de uno o más de los siguientes: configuración del volumen de reproducción; contexto del usuario, incluyendo por la noche, caminando, corriendo o en auto; rango dinámico del convertidor digital a analógico; y rango dinámico del altavoz.
- 40 **12.** El procedimiento según la reivindicación 10, que comprende además controlar cuándo se aplican los nuevos valores de ganancia de DRC a la señal de audio descodificada para producir la señal de audio ajustada de DRC, de modo que al mezclar, la señal de audio se ajusta a DRC solo cuando el volumen de reproducción está por encima de un umbral y solo en la medida necesaria para reducir el nivel de un segmento más alto en la señal de audio descodificada.
- 45 **13.** El procedimiento según la reivindicación 12, que comprende además:
 extraer de los metadatos un valor pico verdadero de una mezcla descendente estéreo de la señal de audio codificada; y
- 50 usar el valor pico verdadero para estimar cuánta compresión de DRC se aplicará a la señal de audio descodificada antes de la conversión de mezcla descendente.
- 55 **14.** El procedimiento según la reivindicación 1, en el que la selección de la característica de DRC de la etapa de descodificación de entre la pluralidad de características de DRC almacenadas se basa en la determinación de uno o más de los siguientes: configuración del volumen de reproducción; contexto del usuario, incluyendo por la noche, caminando, corriendo o en auto; rango dinámico del convertidor digital a analógico; y rango dinámico del altavoz.

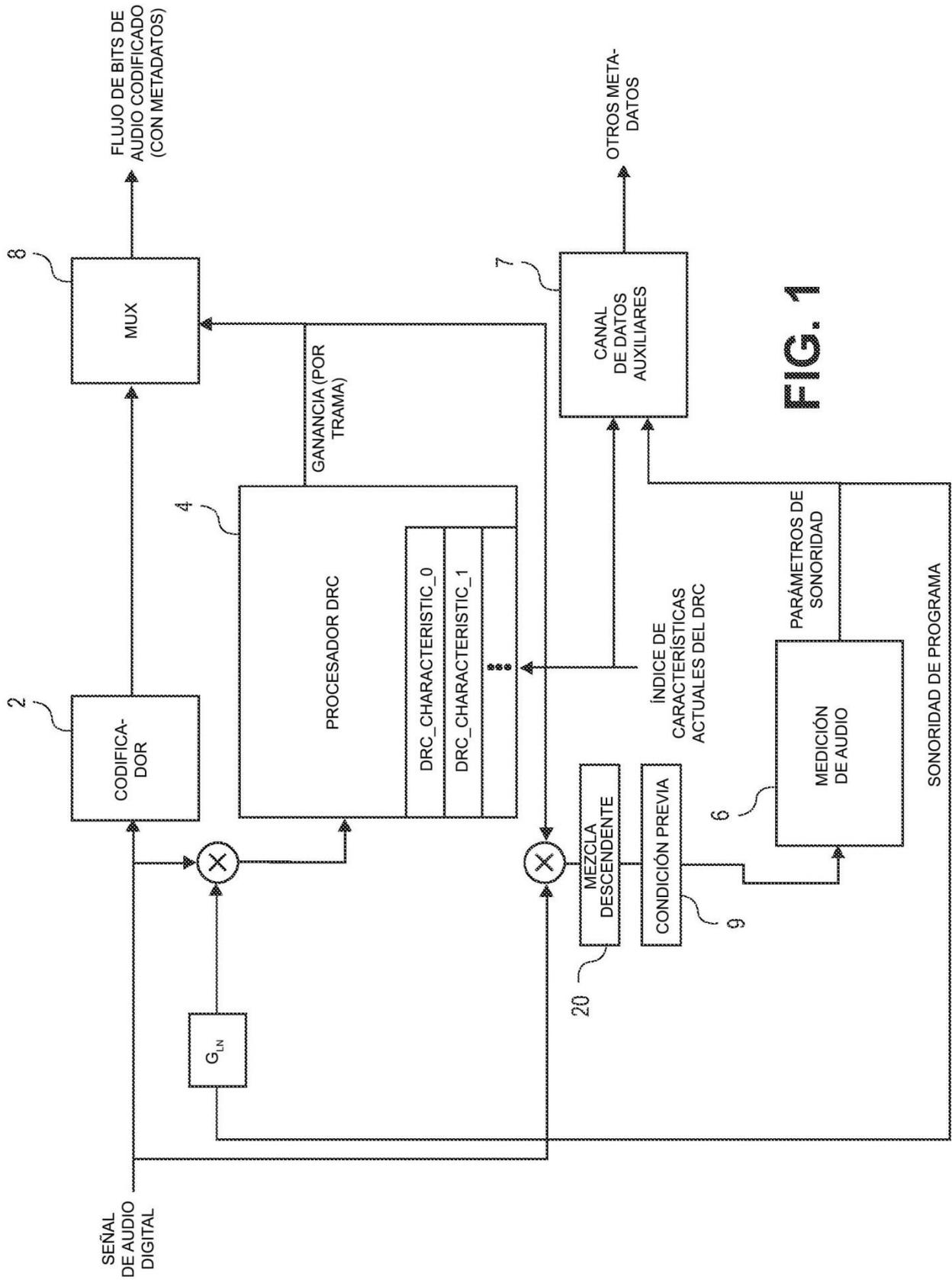


FIG. 1

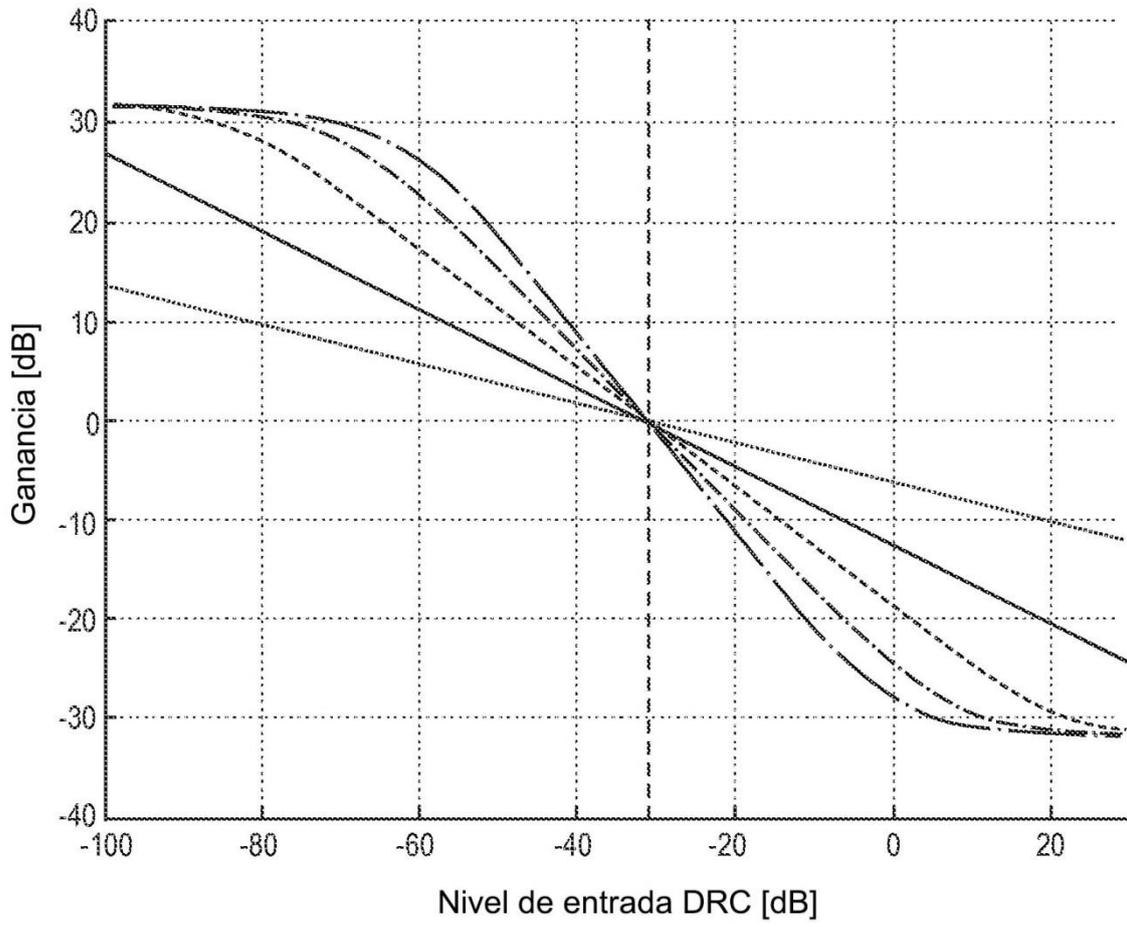


FIG. 2

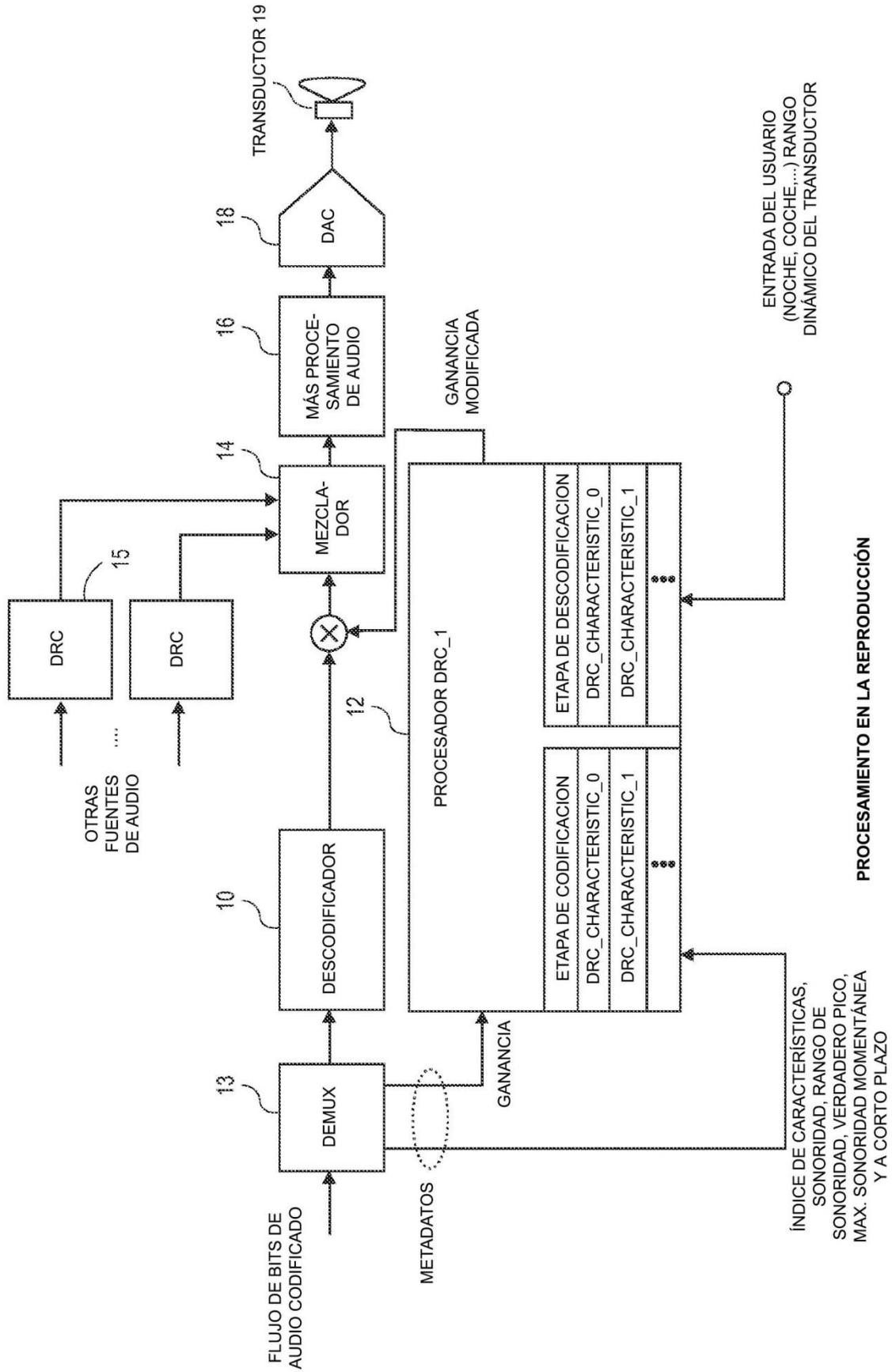


FIG. 3

PROCESAMIENTO EN LA REPRODUCCIÓN

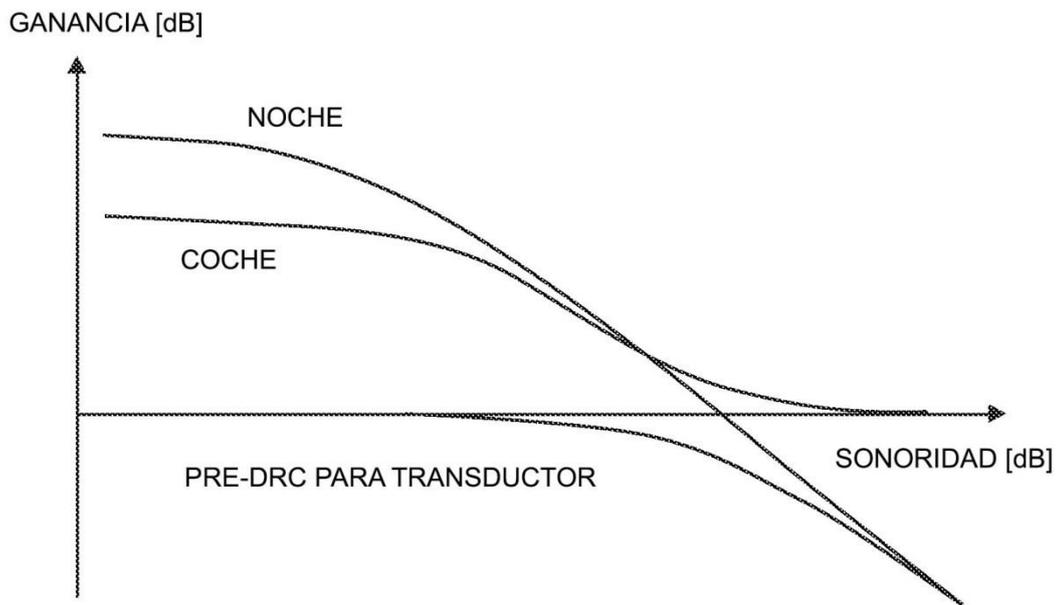


FIG. 4

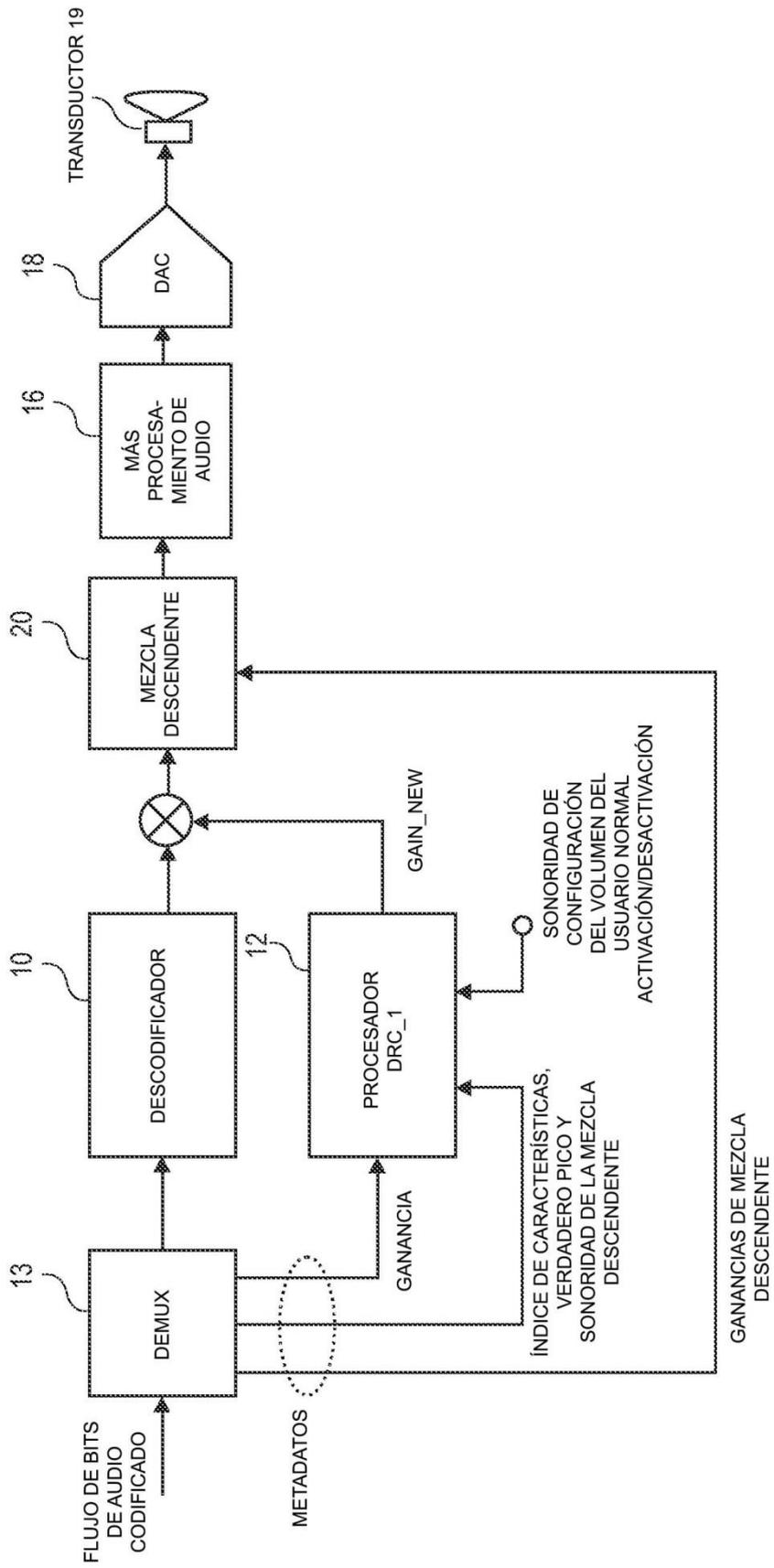


FIG. 5

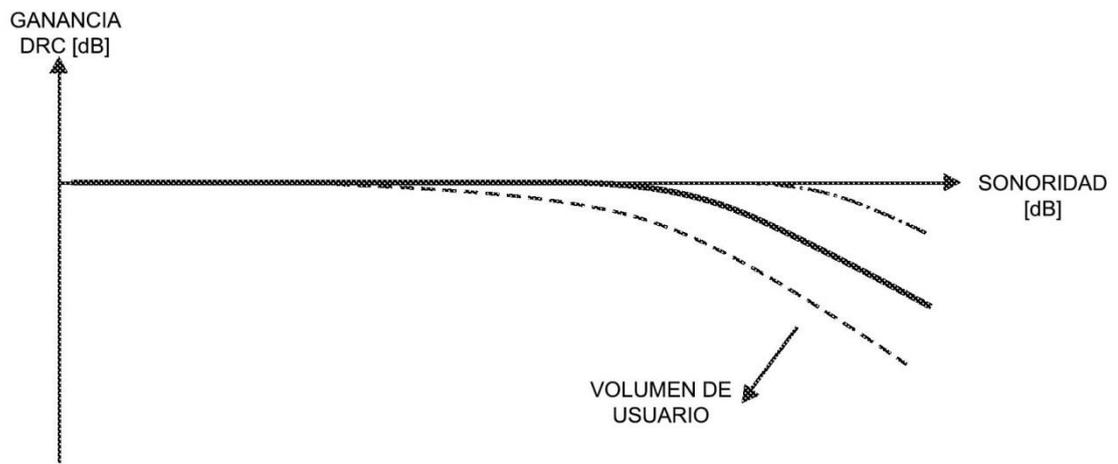


FIG. 6

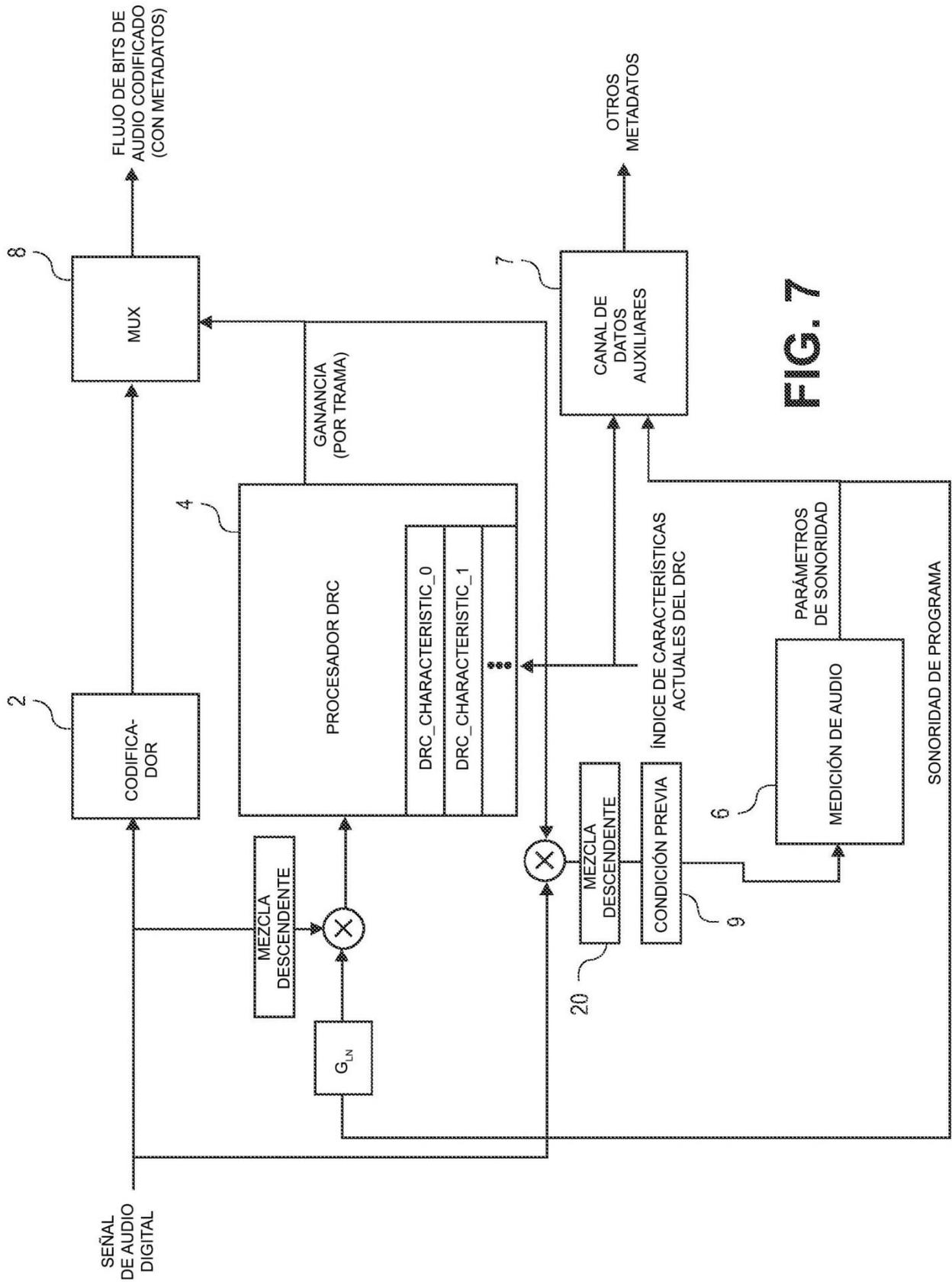


FIG. 7