

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 777 600**

51 Int. Cl.:

G10L 19/008 (2013.01)

H04S 3/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.07.2016 PCT/US2016/043932**

87 Fecha y número de publicación internacional: **09.02.2017 WO17023601**

96 Fecha de presentación y número de la solicitud europea: **25.07.2016 E 16748414 (6)**

97 Fecha y número de publicación de la concesión europea: **11.12.2019 EP 3329487**

54 Título: **Control de rango dinámico basado en metadatos extendidos de audio codificado**

30 Prioridad:

31.07.2015 US 201562199819 P
22.07.2016 US 201615217632

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
05.08.2020

73 Titular/es:

APPLE INC. (100.0%)
One Apple Park Way
Cupertino, California 95014-2084, US

72 Inventor/es:

BAUMGARTE, FRANK

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 777 600 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

Control de rango dinámico basado en metadatos extendidos de audio codificado

5 **[0001]** Esta solicitud reivindica el beneficio de la fecha de presentación anterior de la solicitud provisional de patente de los EE. UU. n.º 62/199.819, presentada el 31 de julio de 2015.

Campo

10 **[0002]** Un modo de realización de la invención se refiere en general a la codificación y decodificación de una señal de audio, y al uso de metadatos asociados con la señal codificada durante la reproducción de la señal decodificada, para mejorar la calidad de reproducción en varios tipos de dispositivos de usuario final de electrónica de consumo. También se describen otros modos de realización.

15 **ANTECEDENTES**

[0003] El contenido de audio digital aparece en muchos casos, incluidos, por ejemplo, archivos de música y películas. En la mayoría de los casos, una señal de audio se codifica con fines de reducción de la velocidad de transferencia de datos o conversión de formato, de modo que la transferencia o entrega del archivo o flujo multimedia sea más práctica, consume menos ancho de banda y/o sea más rápida, lo que permite que muchas otras transferencias puedan ocurrir simultáneamente. El archivo o flujo multimedia se puede recibir en diferentes tipos de dispositivos de usuario final, donde la señal de audio codificada se decodifica antes de presentarse al consumidor a través de altavoces incorporados 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 contenido de audio digital (programas) tienen varios enfoques a su disposición, que pueden usarse para codificar y decodificar contenido de audio. Estos incluyen el Estándar de compresión de audio digital (AC-3, E-AC-3), Revisión B, Documento A/52B, 14 de junio de 2005 publicado por el Advanced Television Systems Committee, Inc. (el "Estándar ATSC"), Instituto Europeo de Estándares de telecomunicaciones, ETSI TS 101154 Digital Video Broadcasting (DVB) basado en el flujo de transporte MPEG-2 en ISO/IEC 13818-7, Advanced Audio Coding (AAC) ("MPEG-2 AAC Standard") e ISO/IEC 14496-3 ("MPEG-4 Audio"), publicado por la Organización Internacional de Normalización (ISO).

[0004] El contenido de audio puede decodificarse y luego procesarse (reproducirse) de forma diferente a como se masterizó originalmente. Por ejemplo, un ingeniero de masterización podría grabar una orquesta o un concierto de tal manera que al reproducirlo sonaría (para un oyente) como si el oyente estuviera sentado en la audiencia del concierto, es decir, frente a la banda u orquesta, escuchando el aplauso desde atrás. De forma alternativa, el ingeniero de masterización podría hacer una reproducción diferente (del mismo concierto), de modo que, por ejemplo, durante la reproducción, el oyente escucharía el concierto como si estuviera en el escenario (donde escucharía los instrumentos "a su alrededor", y los aplausos "al frente"). Esto también se conoce como crear una perspectiva diferente para el oyente en la sala de reproducción, o reproducir el contenido de audio para una "ubicación de audición" diferente o una sala de reproducción diferente.

[0005] El contenido de audio también puede reproducirse para diferentes entornos acústicos, por ejemplo, reproducción a través de auriculares, un teléfono inteligente con altavoz o los altavoces integrados de una tableta, un ordenador portátil o un ordenador de escritorio. En particular, las técnicas de reproducción de audio basadas en objetos ahora están disponibles, en las que un objeto de audio digital individual, que es una grabación de audio digital de, por ejemplo, una sola persona que habla, una explosión, aplausos o sonidos de fondo, puede reproducirse de manera diferente a través de cualquiera de uno o más canales de altavoces en un entorno acústico dado.

50 **[0006]** El rango dinámico en la reproducción de audio contextual se refiere a una relación entre los sonidos más altos y más débiles (niveles de sonoridad) calculados a partir del contenido de audio digital. El nivel de sonoridad se puede calcular utilizando cualquier modelo matemático adecuado, que estima cómo los humanos perciben (o escuchan) el sonido. El control de rango dinámico (DRC) se refiere a enfoques para controlar el rango dinámico, por ejemplo, comprimirlo o expandirlo, para cambiar la forma en que se escuchan las partes fuertes y las partes débiles del contenido de audio durante la reproducción. Los ingenieros de audio aplican DRC a una señal de audio digital, para optimizar una grabación de audio particular para un entorno acústico particular o para una perspectiva particular del oyente. Por ejemplo, una obra de música pop moderna puede tener su rango dinámico comprimido para que pueda reproducirse a un nivel más alto (sin recorte), mientras que una pieza de música clásica a menudo se graba con un mayor rango dinámico.

SUMARIO

65 **[0007]** Un modo de realización de la invención es un sistema de producción o distribución (por ejemplo, un sistema de servidor) que produce valores de ganancia de DRC que son parte de los metadatos de un archivo codificado de contenido de audio digital (o grabación de audio). Por ejemplo, los valores de ganancia de DRC

pueden ser positivos (refuerzo) o negativos (atenuación), y se deben aplicar a la grabación de audio durante la reproducción (por ejemplo, después de que un descodificador haya extraído la grabación de audio del archivo codificado) para ajustar una parte alta y/o una parte débil de la grabación durante la reproducción. El ajuste por DRC puede actualizarse, por ejemplo, en cada trama de la señal de audio digital. El ajuste por DRC puede ayudar a adaptar mejor un tipo particular de grabación de audio a un entorno acústico de reproducción particular o una perspectiva de escucha. Esto permite la reproducción de contenido de audio ajustado por DRC, donde el ajuste por DRC se especificó en la etapa de codificación. El archivo de contenido de audio puede ser, por ejemplo, un archivo de imagen en movimiento, por ejemplo, un archivo de película MPEG, un archivo de solo audio, por ejemplo, un archivo AAC o un archivo que tenga cualquier formato multimedia adecuado.

[0008] En un modo de realización, un procesador de Control de Rango Dinámico (DRC) produce una secuencia de valores de ganancia de DRC del codificador, aplicando una característica seleccionada de una serie de características de DRC, a un grupo de uno o más de los canales de audio u objetos de audio. Los valores de ganancia de DRC del codificador deben ser aplicados por un sistema de descodificación, para ajustar el grupo de canales de audio u objetos de audio al descodificarlos a partir de la grabación de audio digital codificada. Un multiplexor de flujo de bits combina a) la grabación de audio digital codificada con b) la secuencia de valores de ganancia de DRC del codificador, una indicación de la característica de DRC seleccionada y una indicación de una característica alternativa de DRC seleccionada de la pluralidad de características de DRC, esta última como metadatos asociados con la grabación codificada de audio digital. Esto permite que el sistema de codificación ordene o permita como opción de descodificador, un DRC alternativo (que se puede aplicar a la grabación descodificada durante la reproducción).

[0009] La construcción anterior permite al codificador proporcionar información de sonoridad sobre el efecto de haber aplicado la característica alternativa de DRC, además de identificar los escenarios en los que se debe aplicar la característica alternativa de DRC (en lugar de la característica de DRC "predeterminada" también seleccionada en el sistema de codificación). Se logra un ahorro significativo de la velocidad de bits, ya que los valores de ganancia del DRC alternativa pueden obtenerse mediante el sistema de descodificación basado en una única secuencia de ganancia de DRC que se recibe en los metadatos. Esto evita la necesidad de que el sistema de codificación transmita una secuencia de ganancia de DRC independiente para cada escenario de compresión. La secuencia de ganancia de DRC, especialmente cuando cambia por trama, puede considerarse como la porción de metadatos que consume más velocidad de bits.

[0010] En otro modo de realización, los metadatos se definen como que tienen un formato en el que el sistema de producción o distribución (sistema de codificación) puede incluir dos o más secuencias de valores de ganancia de DRC del codificador. Además, los metadatos se definen para permitir que se incluyan instrucciones, que son instrucciones para un sistema de descodificación del sistema de codificación, en el que los metadatos pueden contener instrucciones en las que el sistema de codificación puede especificar que cualquiera de las secuencias de valores de ganancia de DRC del codificador (presentes en los metadatos) se pueden aplicar para ajustar por DRC cualquier subbanda de la grabación de audio digital descodificada. Por ejemplo, los metadatos pueden especificar que cada una de las secuencias de los valores de ganancia de DRC del codificador (que están en los metadatos) se aplicará a una subbanda diferente de la grabación de audio digital descodificada. En otras palabras, los metadatos pueden permitir una asignación arbitraria de las dos o más secuencias de ganancia de DRC que pueden incluirse dentro de los metadatos, a las subbandas arbitrariamente seleccionadas de las subbandas en las que el sistema de descodificación realiza la compresión por subbanda. Una vez más, se logra un ahorro en la velocidad de bits porque, por ejemplo, el sistema de descodificación puede usar la misma secuencia de ganancia de DRC para comprimir múltiples subbandas.

[0011] En otro modo de realización más, además de la capacidad de asignar arbitrariamente una única secuencia de ganancia de DRC a dos o más subbandas, los metadatos también admiten el formato que permite que el sistema de producción o distribución especifique en los metadatos que una primera subbanda se debe ajustar escalando una de las secuencias de ganancia de DRC de acuerdo con un factor de escala, mientras escala la secuencia de ganancia de DRC de acuerdo con otro factor de escala y aplicando esta última a una subbanda diferente. Esto da como resultado que el sistema de descodificación, de acuerdo con las instrucciones en los metadatos, escala una secuencia de ganancia de DRC especificada por un primer factor de escala (antes de aplicar esa secuencia escalada a una primera subbanda), y escala la secuencia de ganancia de DRC especificada por un segundo factor de escala (antes de aplicar esa secuencia escalada a una subbanda diferente), todo como se especifica en los metadatos.

[0012] El resumen anterior no incluye una lista exhaustiva de todos los aspectos de la presente invención. Se contempla que la invención incluye todos los sistemas y procedimientos que pueden practicarse a partir de todas las combinaciones adecuadas de los diversos aspectos resumidos anteriormente, así como los divulgados en la descripción detallada a continuación y particularmente señalados en las reivindicaciones presentadas con la solicitud. Tales combinaciones tienen ventajas particulares que no se enumeran específicamente en el resumen anterior.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0013] 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, en aras de la concisión y la reducción del número total de figuras, una figura dada puede usarse para ilustrar las características de más de un modo de realización de la invención, y no todos los elementos mostrados en una figura pueden ser necesarios para un modo de realización dado.

La Figura 1 es un diagrama de bloques que se utiliza para ilustrar aspectos de un sistema de codificación de audio digital.

La Figura 2 muestra varios ejemplos de características de control de rango dinámico (DRC).

La Figura 3 es un diagrama de bloques que se utiliza para ilustrar aspectos de un sistema de descodificación de audio digital y, en particular, uno en el que el procesamiento de datos se realiza durante la reproducción de la señal de audio descodificada.

La Figura 4 es un diagrama de bloques que describe aspectos de un bloque de aplicación de DRC en el dominio de frecuencia de múltiples bandas de ejemplo.

La Figura 5 se usa para ilustrar un ejemplo de DRC multibanda realizado en el dominio del tiempo como parte de un descodificador de audio.

La Figura 6 muestra algunos campos de ejemplo en los metadatos que se relacionan con DRC.

DESCRIPCIÓN DETALLADA

[0014] Aquí se describen e ilustran diversos modos de realización de la invención en las figuras, que incluyen ejemplos de componentes relevantes de un sistema para producir una grabación de audio digital codificada, y un sistema descodificador para aplicar DRC para ajustar la grabación descodificada, durante la reproducción. Debe observarse la presencia de numerosos detalles sobre los metadatos, incluido su formato y su uso en el sistema descodificador, algunos de los cuales pueden no ser necesarios cuando se ponen en práctica ciertos modos de realización de la invención. Muchos de los detalles se consideran ejemplos del lenguaje utilizado en las siguientes reivindicaciones.

[0015] En otros casos, no se han mostrado en detalle circuitos, estructuras y técnicas bien conocidas para no entorpecer 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 bits de acuerdo con los estándares MPEG; sin embargo, los enfoques para incorporar valores de ganancia de DRC e información relacionada en los metadatos de un archivo de contenido de audio codificado 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).

[0016] **La Figura 1** es un diagrama de bloques que se utiliza para ilustrar aspectos de un sistema de codificación de audio digital. La grabación de audio original o la señal de audio en la **Figura 1** puede tener la forma de un flujo de bits o archivo (donde estos términos se usan aquí indistintamente) de un fragmento de contenido de programa de sonido, como un trabajo musical o un trabajo audiovisual, por ejemplo, la pista de sonido de una película que tiene varios canales de audio; de forma alternativa, o además de los canales de audio, la grabación puede incluir una serie de objetos de audio, por ejemplo, el contenido del programa de sonido de instrumentos musicales individuales, voces y efectos de sonido. El procesamiento de la etapa del codificador puede ser realizado, por ejemplo, por un ordenador (o red de ordenadores) de un productor o distribuidor de contenido de programas de sonido, tal como un productor de actuaciones musicales o películas; el procesamiento de la etapa de descodificación (ver la **Figura 3** a continuación) puede ser realizado, por ejemplo, por un ordenador (o red de ordenadores) de, por ejemplo, un sistema de audio doméstico, una base de altavoces, un sistema de audio en un vehículo de un consumidor. El diagrama de bloques se usa para describir no solo un aparato codificador de audio digital, sino también un procedimiento para codificar una señal de audio.

[0017] El sistema de codificación tiene un codificador 2 que codifica una grabación de audio digital (o también conocida aquí como señal de audio digital), que tiene varios canales de audio u objetos de audio originales (indicados en las figuras aquí por la barra diagonal sobre las líneas que representa el flujo de señal), en un formato digital diferente. El nuevo formato puede ser más adecuado para el almacenamiento de un archivo codificado (por ejemplo, en un dispositivo portátil de almacenamiento de datos, como un disco compacto o un disco de vídeo digital), o para transmitir un flujo de bits al ordenador de un consumidor (por ejemplo, a través de Internet). El codificador 2 también puede realizar una reducción de velocidad de bits con pérdida o sin pérdida (compresión de datos), sobre los canales de audio u objetos de audio originales, por ejemplo, de acuerdo con los estándares MPEG, o la compresión de datos sin pérdida como Apple Lossless Audio Codec (ALAC).

[0018] El procesamiento de la etapa de codificación también puede tener un multiplexor (mux) 8 que combina o ensambla la grabación de audio digital codificada con una o más secuencias de valores de ganancia de DRC, esta última como metadatos asociados con la grabación de audio digital codificada. El resultado de la combinación puede ser un flujo de bits o un archivo codificado (denominado genéricamente de ahora en adelante "un flujo de bits") que contiene la grabación codificada y sus metadatos asociados. Cabe señalar que los metadatos pueden estar incorporados con la grabación codificada en el flujo de bits, o pueden proporcionarse en un archivo separado o canal lateral, genéricamente denominado aquí como un canal de datos auxiliar 7 (con el que está asociada la grabación codificada). Los metadatos asociados con la grabación de audio digital codificada pueden transportarse en varios campos de extensión de ISO/IEC 23003-4:2015 - Tecnología de la información - Tecnologías de audio MPEG - Parte 4: Control de rango dinámico ("MPEG-D DRC").

[0019] La etapa de codificación también tiene un procesador de DRC 4 que produce las secuencias de valores de ganancia de DRC del codificador. Se produce una secuencia de ganancia de DRC predeterminada aplicando una característica o perfil de DRC seleccionada de una serie de características o perfiles de DRC (donde hay al menos dos, o N, que pueden almacenarse en el procesador de DRC 4) a un grupo de uno o más de los canales de audio u objetos de audio que forman parte de la señal de audio digital. Esto puede repetirse para que se produzcan múltiples secuencias de ganancia de DRC, correspondientes a múltiples grupos de canales u objetos de audio. Una característica o perfil de DRC puede almacenarse en la memoria como parte del procesador de DRC 4 y también como parte del procesador DRC_1 12 en el sistema de decodificación; consulte la **Figura 3**. En la **Figura 2** se dan ejemplos de características de DRC, donde el nivel de entrada a lo largo del eje x se refiere a un valor de sonoridad a corto plazo (también denominado aquí nivel de entrada de DRC), mientras que se dan valores de rango de ganancia de DRC a lo largo de eje y.

[0020] La característica de DRC predeterminada puede ser seleccionada por un usuario, a través de la entrada del usuario (por ejemplo, una interfaz gráfica de usuario). El usuario puede ser un ingeniero de mezcla o de sonido que evalúa el tipo de contenido en el canal u objeto relevante, incluyendo, por ejemplo, escuchar el canal u objeto a través del equipo de reproducción (no se muestra), y hace la selección basándose en la experiencia, el tipo de contenido y cómo sonaría el canal u objeto cuando su rango dinámico se haya modificado (de acuerdo con la característica predeterminada) en un entorno acústico o en un escenario de dispositivo de reproducción particular (por ejemplo, auriculares frente a altavoces integrados de un ordenador portátil o de escritorio frente a altavoces independientes). Esto se puede hacer para modificar, por ejemplo, una banda sonora de una película para reproducirla a través de un sistema de audio que puede tener un rango menos dinámico que el sistema de audio de una sala de cine pública.

[0021] Para un nivel de entrada de DRC dado, la característica produce un valor de ganancia correspondiente que es positivo (efecto expansivo) o negativo (efecto de compresión) y que debe aplicarse a la señal de audio de entrada, mediante un bloque de aplicación de DRC 3 - véase la **Figura 1**. En otras palabras, se dice que el bloque de DRC 3 está configurado con una característica de DRC seleccionada para que calcule cualquier nivel de entrada necesario a partir de la señal de audio de entrada, obtiene una ganancia de salida aplicando el nivel de entrada a la característica, y aplica la ganancia de salida a la señal de audio de entrada para realizar el ajuste del rango dinámico. Los valores de ganancia en el gráfico de la **Figura 2** también se denominan aquí valores de ganancia de DRC, que en este ejemplo particular se dan en el formato logarítmico (dB). El nivel de la señal de audio de entrada que se aplica a la característica (nivel de entrada de DRC) puede calcularse durante un intervalo de tiempo predeterminado de la señal de audio de entrada, también denominado aquí como trama, por ejemplo, del orden de menos de 5 milisegundos, por ejemplo, menos de 1 milisegundo. Por lo tanto, una secuencia de ganancia de DRC puede proporcionar valores de ganancia de DRC actualizados por trama. Tenga en cuenta que la señal de audio digital que se está codificando puede estar en un formato modulado por impulsos codificados (PCM) o en un formato basado en paquetes en el que las tramas o fragmentos de la señal de audio están disponibles secuencialmente donde puede estar cada trama o fragmento, por ejemplo, entre 20 y 100 milisegundos de largo, de modo que se aplican varios valores de ganancia de DRC en secuencia a cada trama o fragmento de audio. Estos números, por supuesto, son solo ejemplos, de modo que debe entenderse que los conceptos aplicados aquí no se limitan a la longitud de la trama definida para cada valor de ganancia en una secuencia de ganancia de DRC o para procesar digitalmente una señal de audio.

[0022] Los valores de ganancia producidos al aplicar la señal de audio de entrada a una característica de DRC predeterminada seleccionada (por el procesador de DRC 4 en el sistema de codificación) se deben aplicar para ajustar un grupo de uno o más canales u objetos de audio, al decodificar este último a partir de la grabación de audio digital codificada (en el sistema de decodificación). Eso puede ser parte del procesamiento durante la reproducción como se describe más adelante en la **Figura 3**. Para lograr este objetivo, la etapa de codificación también tiene algunos medios para proporcionar, como metadatos asociados con la grabación de audio digital codificada, la secuencia de valores de ganancia de DRC del codificador al sistema de decodificación. Esto se describió anteriormente, por ejemplo, como el multiplexor 8 por sí mismo, o en combinación con el canal auxiliar de datos 7.

[0023] En un modo de realización, los metadatos también incluyen una indicación de la característica de DRC predeterminada, así como una indicación de una característica alternativa de DRC que se ha seleccionado entre las características de DRC_characteristic_0, 1, ... N. Como se describe a continuación, esto permite que la fuerza de compresión del control de rango dinámico que se aplica en el sistema de descodificación se modifique según lo dicte la entrada del usuario en la etapa de codificación. Las técnicas que permiten que esto tenga lugar son eficientes en la velocidad de bits, ya que las nuevas opciones de control de rango dinámico se dan al sistema de descodificación sin requerir que los metadatos tengan secuencias de ganancia de DRC adicionales (más allá de una sola secuencia de ganancia de DRC predeterminado). Por lo tanto, una modificación relativamente general está disponible para el sistema de descodificación para realizar una correlación de ganancia de la secuencia de ganancia de DRC predeterminado utilizando el conocimiento de la característica alternativa de DRC que se ha especificado en los metadatos. Los metadatos ahora se mejoran definiendo campos adicionales en los que se puede indicar la característica alternativa de DRC, además de, por ejemplo, identificar el escenario o condición particular en el que el sistema de descodificación debe aplicar el control de rango dinámico de acuerdo con la característica alternativa de DRC (en lugar de la característica de DRC predeterminada). Esta correlación de ganancia de la secuencia de ganancia de DRC predeterminado se describe a continuación en relación con la **Figura 3**.

[0024] Aun haciendo referencia a la **Figura 1**, en un modo de realización, los parámetros de sonoridad, o también referidos aquí como información de sonoridad, pueden ser calculados por el procesador de DRC 4 y, en particular, por un bloque de medición de sonoridad 6 (calculador de sonoridad), y donde estos también pueden ser incluido en los metadatos. Estos parámetros de sonoridad dan una medida de la sonoridad de la versión alternativa ajustada por DRC de la grabación de audio digital, que es útil para que el sistema de descodificación evalúe, cuando se le da la opción, si aplicar o no DRC, entre el DRC predeterminado y el alternativo. La entrada al bloque de medición de audio 6 recibe la versión alternativa ajustada por DRC de la señal de audio de entrada, que es proporcionada por un bloque de aplicación de DRC 3, donde este último se ha configurado de acuerdo con la característica alternativa de DRC (que puede haber sido seleccionada a través de la entrada del usuario).

[0025] Se puede tomar cualquiera de varios enfoques para proporcionar la "indicación" de la característica de DRC predeterminada o alternativa (dentro de los metadatos). Como se muestra en la **Figura 1**, el ejemplo particular allí utiliza un índice, que es una referencia o puntero, a una curva o gráfico predeterminado de nivel de entrada o sonoridad frente a la ganancia de DRC de salida. La curva o gráfico pueden almacenarse en el sistema de descodificación como DRC_characteristic_0, 1, ... N en la memoria del DRC_1_processor 12. El sistema de descodificación recuperará la característica de DRC especificada por el índice recibido en los metadatos. De forma alternativa, los metadatos pueden indicar una característica de DRC al contener una serie de constantes o parámetros o coeficientes que, cuando son insertados por el sistema de descodificación en una función matemática predefinida, producen una curva de sonoridad respecto a ganancia de DRC particular. En otro modo de realización, la indicación de una característica de DRC puede ser una tabla de consulta de todos los niveles de entrada o valores de sonoridad y los valores de ganancia de DRC correspondientes que definen una curva de ganancia de DRC. Por último, la indicación de una característica de DRC puede ser un número reducido de valores de sonoridad y valores de ganancia de DRC correspondientes a partir de los cuales el sistema de descodificación interpola la curva de ganancia de DRC o un valor de ganancia de DRC particular para un nivel de sonoridad de entrada no especificado (que no está especificado en los metadatos). Para una eficiencia de velocidad de bits, las indicaciones de las características de DRC deben ser meramente índices de curvas o gráficos predeterminados de sonoridad frente a ganancia de DRC (que se almacenan en el sistema de descodificación).

[0026] Habiendo descrito cómo se pueden llenar los metadatos en el sistema de codificación, el uso de los metadatos mientras se procesa para la reproducción se describe ahora usando el ejemplo de la **Figura 3**. La **Figura 3** es un diagrama de bloques que se usa para ilustrar aspectos de un sistema de descodificación y en particular uno en el que el procesamiento de datos se realiza durante la reproducción de la señal de audio descodificada. Este es un sistema para producir una grabación de audio digital descodificada en la que se recibe un flujo de bits en el que se ha codificado una grabación de audio digital (ver **Figura 1**). Las operaciones de procesamiento de señal digital descritas aquí para los componentes que se muestran en la **Figura 3** pueden implementarse mediante hardware dedicado (circuitos), o pueden implementarse mediante una combinación de circuitos de hardware y uno o más procesadores programados en los que la memoria ha almacenado instrucciones que cuando son ejecutadas por uno o más procesadores (genéricamente denominado aquí "procesador") realiza las operaciones descritas aquí. En particular, un demultiplexor (demux) 13 recibe el flujo de bits de audio codificado y extrae el audio codificado, multicanal o de múltiples objetos, que se alimenta a un descodificador 10, mientras que los metadatos extraídos se proporcionan a un procesador DRC_1 12. En un modo de realización, los metadatos incluyen una secuencia de valores de ganancia de DRC del codificador (ganancias de DRC, como se muestra en la **Figura 3**) que pueden ser los valores de ganancia de DRC predeterminados mencionados anteriormente en la **Figura 1**. Los metadatos también incluyen una indicación de una característica de DRC seleccionada (característica de DRC predeterminada) que se utilizó para obtener la secuencia de valores de ganancia de DRC predeterminados por el sistema codificador (al aplicar la grabación de audio digital original a la característica de DRC seleccionada o predeterminada). Además, también se recibe una indicación de una característica alternativa de DRC en los metadatos. Debe entenderse que algunos o todos los metadatos pueden

estar en un canal separado que el flujo de bits de audio codificado, por ejemplo, el canal auxiliar de datos 7 - ver la **Figura 1**.

[0027] El descodificador 10 descodificará la grabación de audio digital (por ejemplo, deshará o realizará el inverso de las operaciones realizadas por el codificador 2 de la **Figura 1**), y luego la reproducción de la grabación descodificada se realiza comenzando con un bloque multiplicador 11 que aplica los valores de ganancia de DRC predeterminados a la señal de audio descodificada o un conjunto de ganancias de DRC de nuevo correlacionadas, para producir una grabación de audio ajustada por rango dinámico (ajustada por DRC). Las señales de audio ajustadas por DRC pueden someterse a un procesamiento de audio adicional 16 (por ejemplo, mezcla descendente) antes de convertirse a una forma analógica (por un convertidor digital a analógico, DAC, 18) y luego alimentarse a una entrada de controlador de altavoz de un transductor electroacústico 19.

[0028] La secuencia alternativa de valores de ganancia de DRC, también conocida como las ganancias de DRC de nuevo correlacionadas en la **Figura 3**, puede ser calculada por el procesador DRC_1 12 que realiza el siguiente proceso. Primero, se produce una inversa de la característica de DRC predeterminada, utilizando la indicación de la característica de DRC predeterminada que se recibe en los metadatos. Por ejemplo, los metadatos pueden incluir el índice de la característica de DRC predeterminada. Este índice puede usarse para buscar la característica de DRC predeterminada que puede almacenarse en el procesador DRC_1 12 como se muestra (como uno de DRC_characteristic_0, 1, ... N). La inversa puede obtenerse, por ejemplo, invirtiendo las variables de entrada y salida de una función matemática (curva de ganancia de DRC) que representa la característica de DRC, y aplicando la secuencia de valores de ganancia de DRC codificada recibidos en los metadatos a la "salida" de la función matemática (o como entrada a una inversa calculada de la función matemática) para producir una secuencia correspondiente de valores de sonoridad, por trama de DRC.

[0029] El proceso continúa con la obtención de una característica alternativa de DRC, utilizando la indicación recibida en los metadatos. Por ejemplo, DRC_characteristic_3 puede ser el valor predeterminado, mientras que la alternativa es DRC_characteristic_5. La secuencia de valores de sonoridad que se calculó utilizando la inversa de la característica predeterminada, DRC_characteristic_3, ahora se aplica como entrada a la característica alternativa, DRC_characteristic_5, para producir una secuencia de valores de ganancia de DRC a los que se hace referencia en la **Figura 3** como ganancias de DRC de nuevo correlacionadas o "ganancias alternativas de la DRC". Las ganancias de DRC de nuevo correlacionadas se aplican luego por el bloque multiplicador 11 a la grabación de audio digital descodificada (procedente de la salida del descodificador 10) para producir una versión alternativa ajustada por DRC de la grabación de audio descodificada.

[0030] El sistema de descodificación de la **Figura 3**, por lo tanto, tiene la opción de aplicar (a la salida del descodificador 10) los valores de ganancia de DRC predeterminados que se reciben en los metadatos o producir (y luego aplicar) ganancias de nuevo correlacionadas utilizando el procedimiento descrito anteriormente que se basa en la indicación de la característica alternativa de DRC (donde se recibió la indicación en los metadatos). En un modo de realización, la elección entre esos dos ajustes de control de rango dinámico puede estar de acuerdo con las instrucciones recibidas en los metadatos. De forma alternativa, la elección puede hacerse únicamente por el sistema de descodificación, en base a la entrada del usuario y/o al conocimiento predeterminado del rango dinámico de un transductor 19 que se está utilizando para la reproducción. En términos más generales, la sensibilidad del sistema de reproducción, incluidas las ganancias aplicadas durante el procesamiento de audio adicional 16, y la sensibilidad del convertidor digital a analógico (DAC) 18 también pueden tenerse en cuenta al decidir entre el DRC predeterminado o alternativo.

[0031] Otro modo de realización también se representa en la **Figura 3**, donde también puede haber un mezclador 14 que sirve para combinar señales de audio de otras fuentes de audio que pueden haber realizado ajustes de control de rango dinámico separados o independientes (como se muestra en los bloques de aplicación de DRC independientes 3).

[0032] La **Figura 1** y la **Figura 3**, como se describió anteriormente, representan un modo de realización de la invención en el que se implementa una característica de correlación de ganancia de DRC más útil usando los metadatos, al incorporar los índices de las características de DRC predeterminada y alternativa (junto con parámetros opcionales de sonoridad relacionados con el DRC alternativo) en los metadatos. La **Figura 1** y la **Figura 3** también representan otros modos de realización de la invención en los que se puede realizar un DRC multibanda (por el bloque multiplicador 11 o por ciertos elementos internos del descodificador 10) sobre la señal de audio descodificada, como se especifica en los metadatos. (por el sistema de codificación). Primero, existe la capacidad de modificar los valores de ganancia de DRC predeterminados, especificando la escala individual, por subbanda, de los valores de ganancia de DRC predeterminados (por el sistema de codificación y mediante instrucciones en los metadatos). La misma secuencia de ganancia de DRC predeterminada ahora puede ser reutilizada por el sistema de descodificación y aplicada a múltiples subbandas. Por lo tanto, haciendo referencia a la **Figura 1**, el procesador de DRC 4 ahora produce, además de una secuencia de ganancia de DRC predeterminada, una definición de subbanda y una asignación de secuencia de ganancia de DRC a subbanda. La definición de subbanda puede ser completamente convencional, por ejemplo, definiendo varias frecuencias de cruce para al menos dos subbandas dentro del espectro de audio general. Además, los metadatos ahora

especifican que una de las múltiples secuencias de valores de ganancia de DRC del codificador (por ejemplo, las secuencias de ganancia de DRC predeterminadas) que están en los metadatos se deben aplicar al rango dinámico: ajustar dos o más subbandas de un canal de audio u objeto de audio que se va a descodificar (a partir de la grabación de audio digital codificada producida por el codificador 2). Los metadatos pueden especificar, además,

5 1) un primer valor de escalado que se aplicará para escalar una secuencia especificada de las secuencias de valores de ganancia de DRC, antes de aplicar la secuencia escalada a una primera subbanda del canal de audio u objeto de audio descodificado, y 2) un segundo valor de escalado diferente que se aplicará para escalar la secuencia especificada de las secuencias de valores de ganancia de DRC del codificador antes de aplicar la secuencia escalada a una segunda subbanda del canal de audio u objeto de audio descodificado. Como se ve en

10 la **Figura 6**, se muestran algunos campos de ejemplo en los metadatos relacionados con DRC multibanda. En particular, una estructura de datos denominada índice de frecuencia de cruce puede definir las frecuencias de cruce de dos o más subbandas. Las frecuencias de cruce se indican junto con el recuento de banda de la estructura de datos, que indica el número de subbandas. Una estructura de datos adicional, multibandDRCscaling(p , $band1$, $band2$, ..., $scalar1$, $scalar2$, ...) especifica cuál ($p= 1, 2, \dots, K$) de las secuencias de ganancia de DRC múltiple ($K \geq 2$) se debe aplicar para ajustar dos o más de las subbandas $band1$, $band2$, ... que se han definido (son conocidas por el sistema de descodificación), y los diferentes valores de escala $scalar1$, $scalar2$, ... (escala de atenuación o amplificación) que se aplicarán a la misma secuencia de ganancia de DRC p antes de aplicar la secuencia de DRC escalada a las dos o más subbandas, respectivamente.

20 **[0033]** El ejemplo de la **Figura 6** también ilustra el modo de realización en el que los metadatos incluyen un conjunto de ganancia de DRC codificado, que es una estructura de datos que tiene una o más secuencias de ganancia de DRC (o secuencias de valores de ganancia de DRC del codificador), y donde puede haber múltiples conjuntos de ganancia en los metadatos (como se indica en la estructura de datos GainSetCount).

25 **[0034]** En un modo de realización, los metadatos especifican que una de las secuencias de ganancia de DRC (en los metadatos) se aplicará para ajustar dos o más subbandas específicas de un canal de audio u objeto de audio (que se ha descodificado de la grabación de audio digital codificada). Los metadatos pueden especificar de forma alternativa que la secuencia de valores de ganancia de DRC del codificador se aplique a todas las subbandas del canal u objeto de audio descodificado. En algunos modos de realización, los metadatos no se refieren a ninguna agrupación de canales u objetos, de modo que el procesador en el sistema de descodificación no realiza ninguna agrupación de canales de audio u objetos de audio de la grabación de audio descodificada, cuando realiza el DRC multibanda sobre la grabación de audio descodificada. Por ejemplo, puede haber solo dos canales de audio que son descodificados, y se debe aplicar el mismo DRC de subbanda a ambos canales, a menos que se especifiquen diferentes valores de escala en los metadatos para diferentes subbandas.

35 **[0035]** La aplicación de los valores de ganancia de DRC a una señal de audio descodificada (por un procesador programado o una combinación de procesador programado y lógica cableada, en el sistema de descodificación) puede estar en el dominio de la frecuencia o en el dominio del tiempo. La **Figura 4** muestra un ejemplo de una implementación en el dominio de la frecuencia, en la que un filtro de cruce multibanda 17 recibe como entrada un canal u objeto de audio único descodificado. El filtro 17 dividirá su señal de entrada en dos o más bandas constituyentes. El filtro 17 puede programarse para definir las bandas o frecuencias de cruce, como se especifica en los metadatos. Las señales de subbanda resultantes a , b , ... n se alimentan a continuación en paralelo a un número de multiplicadores 11a, 11b, ... 11n, respectivamente, que sirven para atenuar o amplificar las señales de subbanda de acuerdo con sus ganancias de DRC asociadas, respectivamente. Estos últimos pueden ser los valores predeterminados que se especifican en los metadatos (seleccionados por el sistema de codificación) o pueden ser valores "modificados". Un valor de ganancia de DRC modificado puede ser una ganancia de DRC predeterminada que se ha escalado según lo especificado en los metadatos, o puede ser el resultado de correlacionar una ganancia de DRC predeterminada a través de una característica alternativa de DRC según el procedimiento descrito anteriormente. Las salidas de los multiplicadores 11a, 11b, ... luego se suman mediante una unidad sumadora 20 para producir un canal u objeto de audio único ajustado por DRC, que luego se alimenta al mezclador 14.

40 **[0036]** La **Figura 5** muestra un ejemplo de una implementación en el dominio del tiempo de la aplicación de valores de ganancia de DRC. Este enfoque puede ser particularmente deseable cuando el descodificador 10 (ver **Figura 3**) ya tiene el canal u objeto de audio descodificado en forma de subbanda (donde el sistema de codificación también tiene conocimiento de las definiciones de estas bandas y, por lo tanto, puede especificarlas en el metadatos). El descodificador 10 también puede tener un banco de filtros de síntesis que se usa para combinar la forma de subbanda de la señal de audio descodificada en una sola secuencia de muestra de tiempo o flujo de bits modulado por impulsos codificados. Este banco de filtros tiene un doble propósito para el ajuste por DRC, al proporcionar a sus n entradas escalares n ganancias de DRC (en forma lineal en lugar de forma logarítmica o decibelios). El banco de filtros de síntesis aplica los valores de ganancia en sus n entradas escalares a las n señales de subbanda, respectivamente, antes de combinarlas en una única secuencia en el dominio del tiempo. Como en la solución del dominio de la frecuencia, las ganancias de DRC pueden ser los valores predeterminados en los metadatos que han sido seleccionados por el sistema de codificación, o pueden ser los valores modificados

55 analizados anteriormente.

60

65

5 **[0037]** Se debe comprender que los modos de realización descritos aquí son meramente ilustrativos y no restrictivos de la amplia invención y no restrictivos, y que la invención no se limita a las construcciones y disposiciones específicas mostradas y descritas, ya que a los expertos en la técnica se les pueden ocurrir otras diversas modificaciones. Por ejemplo, aunque cada una de las etapas de codificación y descodificación puede describirse en un modo de realización como operando por separado, por ejemplo, en una máquina productora de contenido de audio y en una máquina consumidora de contenido de audio que se comunican por Internet, la codificación y descodificación también podrían realizarse dentro de la misma máquina (por ejemplo, como parte de un proceso de transcodificación). Por lo tanto, la descripción debe considerarse como ilustrativa, no limitativa.

10

REIVINDICACIONES

1. Un sistema para producir una grabación de audio digital codificada que tiene una pluralidad de canales de audio u objetos de audio, que comprende:

5 un codificador de audio (2) para codificar una grabación de audio digital que tiene una pluralidad de canales de audio u objetos de audio;

10 un procesador de control de rango dinámico, DRC, (4) para producir una secuencia de valores de ganancia de DRC del codificador aplicando una característica seleccionada de una pluralidad de características de DRC a un grupo de uno o más de la pluralidad de canales de audio u objetos de audio, en el que los valores de ganancia de DRC del codificador se deben aplicar para ajustar el grupo de canales de audio u objetos de audio al descodificarlos a partir de la grabación de audio digital codificada; y

15 medios para proporcionar como metadatos asociados con la grabación de audio digital codificada i) la secuencia de valores de ganancia de DRC del codificador, ii) una indicación de la característica de DRC seleccionada, y iii) una indicación de una característica alternativa de DRC seleccionada de la pluralidad de características de DRC.

20 **2.** El sistema según la reivindicación 1, en el que los metadatos especifican un escenario o condición en los que un sistema de descodificación debe aplicar DRC de acuerdo con la característica alternativa de DRC en lugar de la característica seleccionada de DRC.

25 **3.** El sistema según la reivindicación 1, en el que el procesador de DRC debe recibir la grabación de audio digital como entrada, y aplicar la entrada a un bloque de aplicación de DRC que se ha configurado de acuerdo con la característica alternativa de DRC, para producir una versión alternativa ajustada por DRC de la grabación de audio digital,

30 en el que el sistema comprende, además, una calculadora de sonoridad para calcular la información de sonoridad que proporciona una medida de sonoridad de la versión alternativa ajustada por DRC de la grabación de audio digital,

35 y en el que los medios para proporcionar como metadatos asociados con la grabación de audio digital codificada incluyen la información de sonoridad, para la versión alternativa ajustada por DRC, como parte de los metadatos.

4. El sistema según la reivindicación 1, en el que en los metadatos, la indicación de la característica alternativa de DRC comprende uno de

40 a) un índice o referencia a una curva o gráfico de sonoridad respecto a la ganancia de DRC que se almacena en un sistema de descodificación,

b) una pluralidad de constantes o parámetros que, cuando se insertan mediante el sistema de descodificación en una función matemática predefinida, definen una curva de sonoridad respecto a la ganancia de DRC,

45 c) una tabla de consulta de sonoridad y los valores de ganancia de DRC correspondientes, o

d) una pluralidad de sonoridad y valores de ganancia de DRC correspondientes a partir de los cuales el sistema de descodificación interpola un valor de ganancia de DRC para un nivel de sonoridad de entrada.

50 **5.** El sistema según la reivindicación 1, en el que el procesador de DRC debe producir un conjunto de ganancia de DRC del codificador que tiene una pluralidad de secuencias de valores de ganancia de DRC del codificador, y en el que los medios para proporcionar como metadatos asociados con la grabación de audio digital codificada también incluyen el conjunto de ganancia de DRC codificada como parte de los metadatos,

55 y en el que los metadatos especifican que una de la pluralidad de secuencias de valores de ganancia de DRC del codificador se debe aplicar para ajustar una pluralidad de subbandas de un canal de audio u objeto de audio que se ha descodificado a partir de la grabación de audio digital codificada.

60 **6.** El sistema según la reivindicación 5 en el que los metadatos especifican que dicha una de las secuencias de valores de ganancia de DRC del codificador se aplicará a todas las subbandas de la grabación de audio digital descodificada.

65 **7.** El sistema según la reivindicación 5 en el que los metadatos especifican que 1) una primera subbanda de la grabación de audio digital descodificada debe ajustarse por DRC mediante una de las secuencias de valores de

ganancia de DRC del codificador, y 2) una segunda subbanda debe ser ajustada por DRC por otra de la pluralidad de secuencias de valores de ganancia de DRC del codificador.

8. El sistema según la reivindicación 5, en el que los metadatos especifican 1) un primer valor de escala que se aplicará para escalar la secuencia especificada de las secuencias de valores de ganancia de DRC antes de aplicar la secuencia escalada a una primera subbanda del canal de audio o audio objeto descodificado, y 2) un segundo valor de escala diferente que se aplicará para escalar la secuencia especificada de las secuencias de los valores de ganancia de DRC del codificador antes de aplicar la secuencia escalada a una segunda subbanda del canal de audio u objeto de audio descodificado.

9. Un sistema para producir una grabación de audio digital descodificada, que comprende:

un procesador (12); y

la memoria que tiene almacenada en ella instrucciones que, cuando son ejecutadas por el procesador, hacen que el procesador:

reciba un flujo de bits en el que se ha codificado una grabación de audio digital y metadatos asociados con la grabación de audio digital, en el que los metadatos incluyen una secuencia de valores de ganancia de DRC del codificador, una indicación de una característica de DRC seleccionada, en el que la secuencia de valores de ganancia de DRC del codificador se obtuvo basándose en la aplicación de la grabación digital de audio a la característica de DRC seleccionada, y una indicación de una característica alternativa de DRC,

descodifique la grabación digital de audio y realice la reproducción de la grabación descodificada produciendo una grabación de audio ajustada por DRC alternativo para la reproducción, mediante

a) producir una inversa de la característica de DRC seleccionada utilizando la indicación, recibida en los metadatos, de la característica de DRC seleccionada, y aplicando la secuencia de valores de ganancia de DRC del codificador, recibida en los metadatos, como entrada a dicha inversa para producir una secuencia de valores de sonoridad,

b) usar la indicación, recibida en los metadatos, de la característica alternativa de DRC, para obtener la característica alternativa de DRC, y aplicar la secuencia de valores de sonoridad como entrada a la característica alternativa de DRC para producir una secuencia alternativa de valores de ganancia de DRC, y

c) aplicar la secuencia alternativa de valores de ganancia de DRC a la grabación de audio digital descodificada para producir una versión alternativa ajustada por DRC de la grabación de audio digital.

10. El sistema según la reivindicación 9, en el que los metadatos incluyen un conjunto de ganancia de DRC del codificador, teniendo el conjunto de ganancia de DRC del codificador una pluralidad de secuencias de valores de ganancia de DRC del codificador,

y en el que los metadatos contienen instrucciones en las que un sistema de codificación puede especificar que cualquiera de la pluralidad de secuencias de valores de ganancia de DRC del codificador se puede aplicar a cualquier subbanda de la grabación de audio digital descodificada.

11. El sistema según la reivindicación 9, en el que los metadatos incluyen un conjunto de ganancia de DRC del codificador, teniendo el conjunto de ganancia de DRC del codificador una pluralidad de secuencias de valores de ganancia de DRC del codificador,

y en el que los metadatos contienen instrucciones para que el procesador aplique una secuencia especificada de las secuencias de valores de ganancia de DRC del codificador a una pluralidad de subbandas de la grabación de audio digital descodificada cuando se realiza un DRC multibanda.

12. Un procedimiento para producir una grabación de audio digital codificada, que comprende:

codificar una grabación de audio digital que tiene una pluralidad de canales de audio u objetos de audio;

producir una secuencia de valores de ganancia de DRC del codificador aplicando una característica seleccionada de una pluralidad de características de DRC a un grupo de uno o más de los canales de audio u objetos de audio, en el que los valores de ganancia de DRC del codificador se deben aplicar para ajustar el grupo de canales de audio u objetos de audio al descodificarlos a partir de la grabación de audio digital codificada; y

proporcionar como metadatos asociados con la grabación de audio digital codificada (i) la secuencia de valores de ganancia de DRC del codificador, (ii) una indicación de la característica de DRC seleccionada y (iii) una indicación de una característica alternativa de DRC seleccionada a partir de una pluralidad de características de DRC.

5

13. El procedimiento según la reivindicación 12, que comprende, además:

producir una versión alternativa ajustada por DRC de la grabación de audio digital de acuerdo con la característica alternativa de DRC;

10

calcular la información de sonoridad que proporciona una medida de la sonoridad de la versión alternativa ajustada por DRC de la grabación de audio digital; y

15

proporcionar como parte de dichos metadatos asociados con la grabación de audio digital codificada, la información de sonoridad para la versión alternativa ajustada por DRC.

14. El procedimiento según la reivindicación 12 o 13, que comprende, además

20

proporcionar como parte de dichos metadatos asociados con la grabación de audio digital codificada, una instrucción de que la misma secuencia de valores de ganancia de DRC del codificador debe ser aplicada por un sistema de decodificación para ajustar una pluralidad de subbandas de un canal de audio u objeto de audio que ha sido decodificado a partir de la grabación de audio digital codificada.

25

15. El procedimiento según la reivindicación 14, que comprende, además,

30

proporcionar como parte de dichos metadatos asociados con la grabación de audio digital codificada, 1) un primer valor de escala e instrucciones para aplicar el primer valor de escala para escalar la secuencia especificada de las secuencias de valores de ganancia de DRC antes de aplicar la secuencia escalada a una primera subbanda del canal de audio u objeto de audio descodificado, y 2) un segundo valor de escala diferente e instrucciones para aplicar el segundo valor de escala para escalar la secuencia especificada de las secuencias de valores de ganancia de DRC del codificador antes de aplicar la secuencia escalada a una segunda subbanda del canal de audio u objeto de audio descodificado.

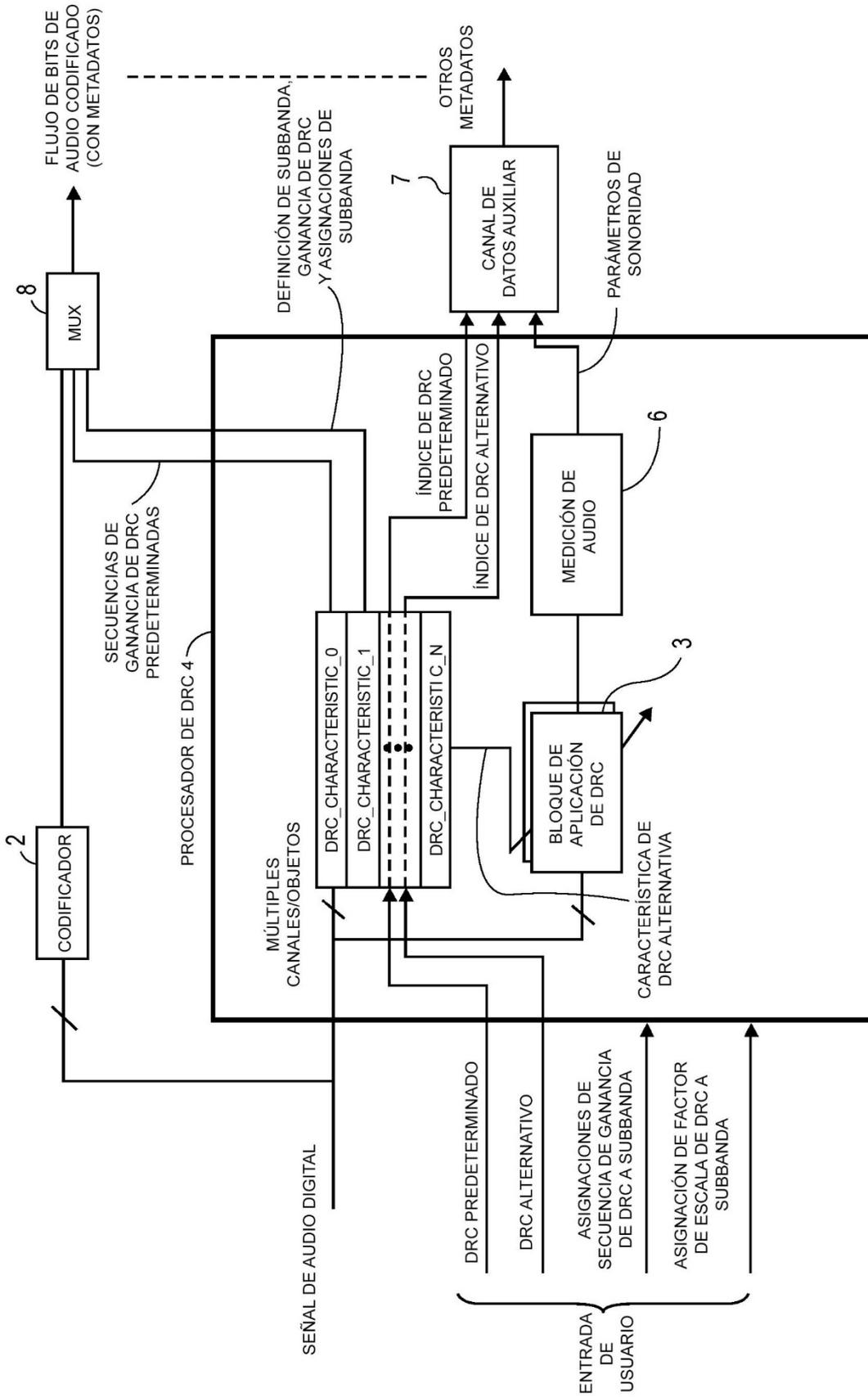


FIG. 1

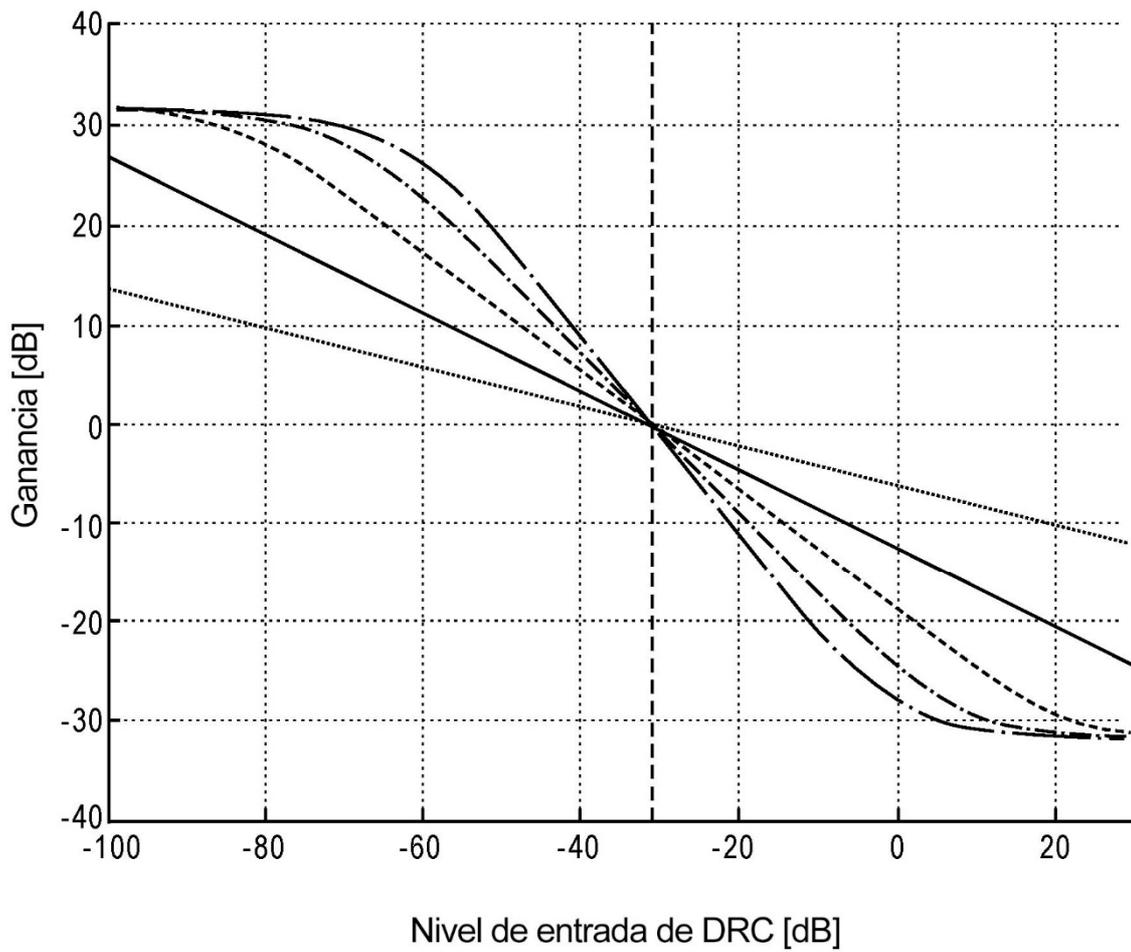
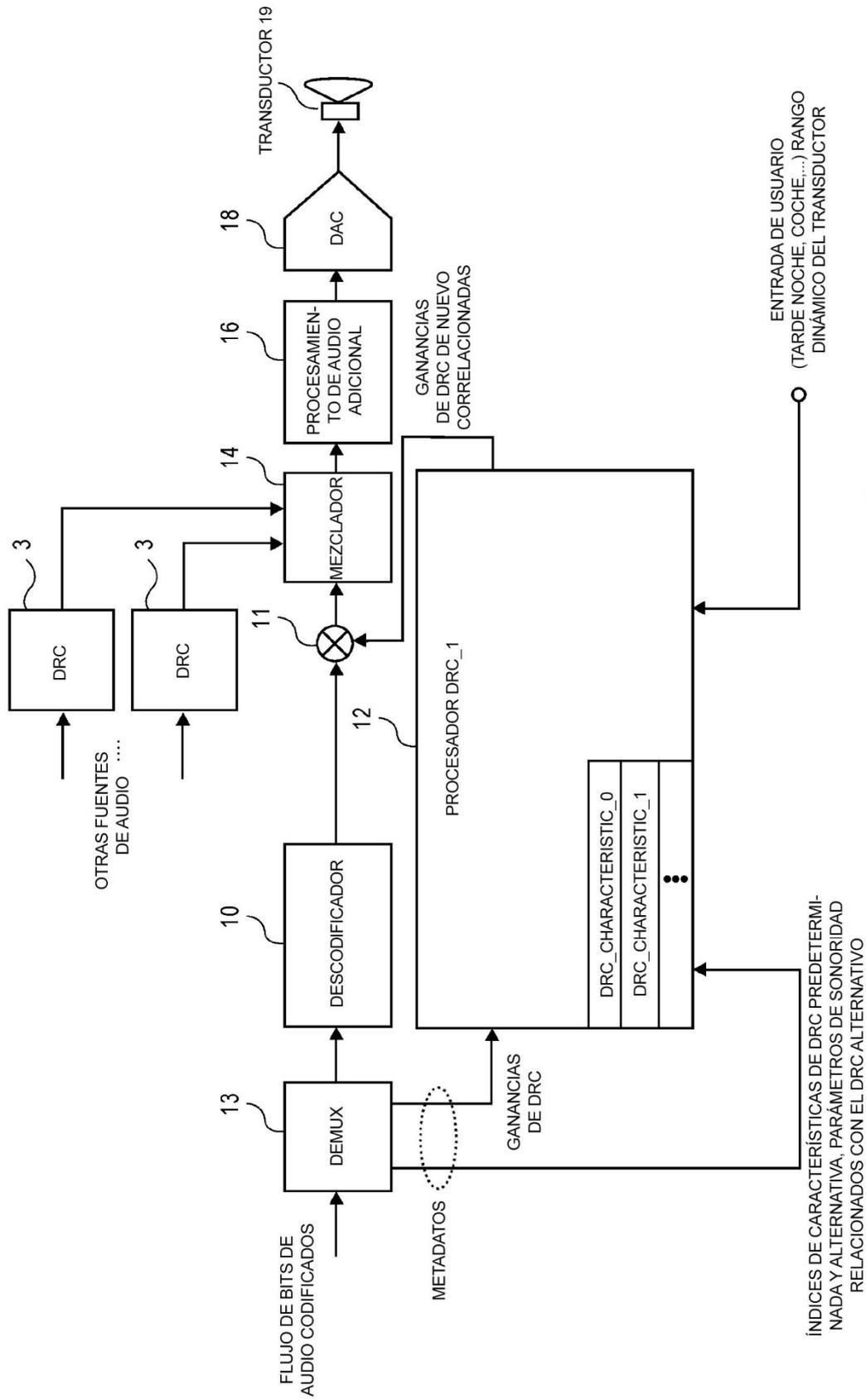


FIG. 2



PROCESAMIENTO EN LA REPRODUCCIÓN

FIG. 3

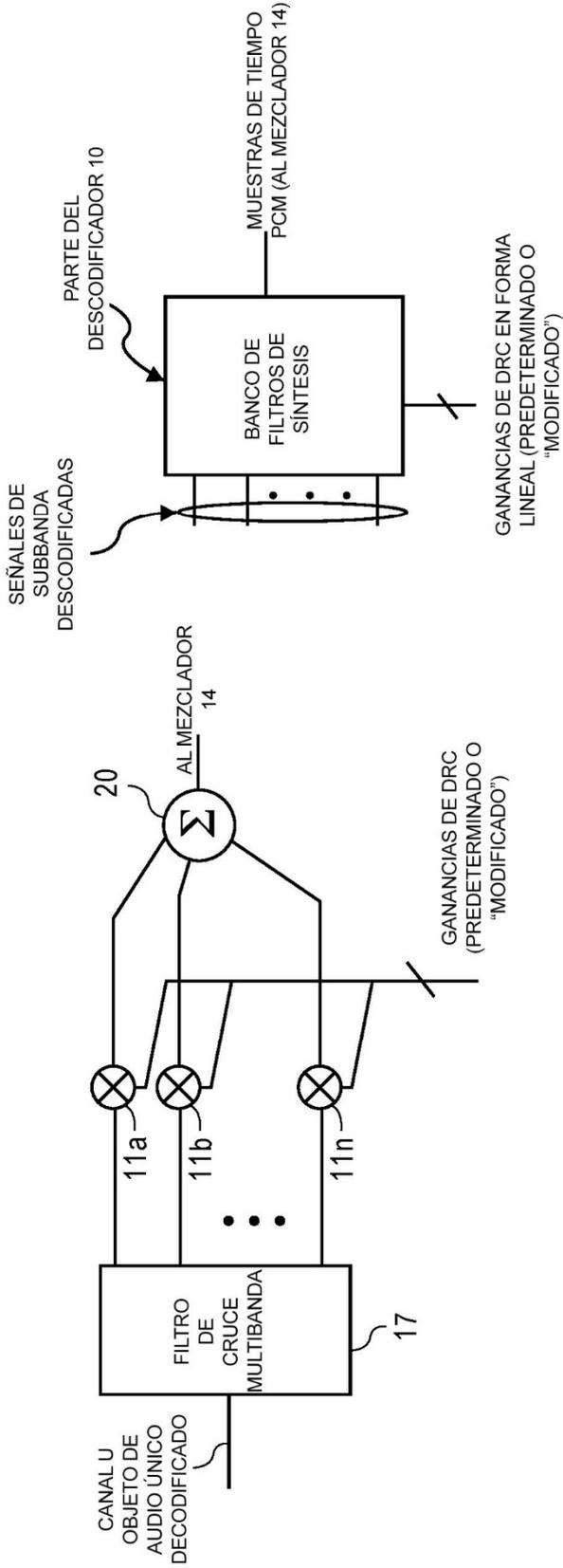


FIG. 4

FIG. 5

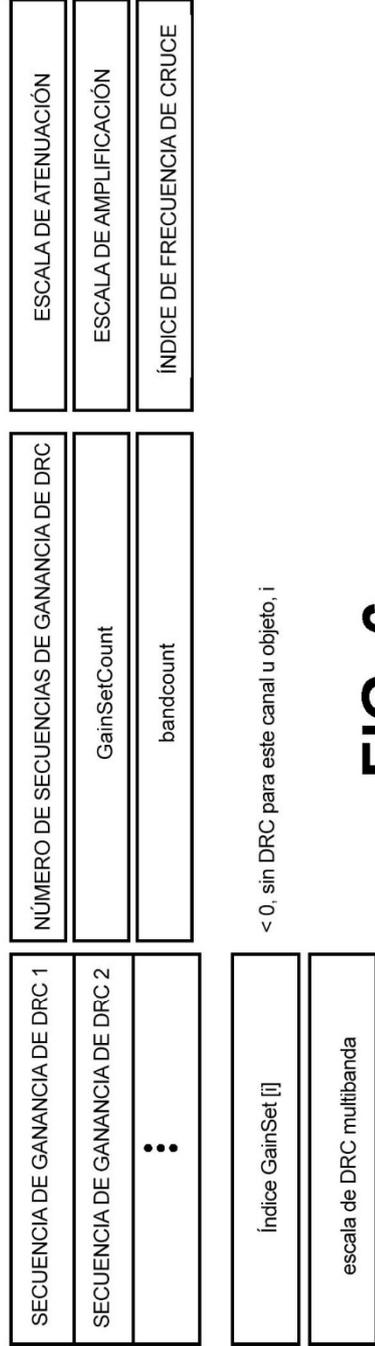


FIG. 6