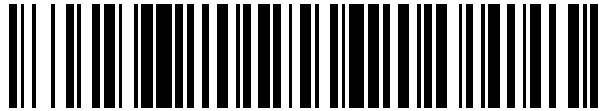


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 527 552**

51 Int. Cl.:

G10L 19/16 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.03.2006 E 06748775 (1)**

97 Fecha y número de publicación de la concesión europea: **05.11.2014 EP 1878010**

54 Título: **Verificación de metadatos de audio**

30 Prioridad:

13.04.2005 US 671361 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
26.01.2015

73 Titular/es:

**DOLBY LABORATORIES LICENSING
CORPORATION (100.0%)
100 Potrero Avenue San Francisco
California 94103, US**

72 Inventor/es:

CROCKETT, BRETT GRAHAM

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 527 552 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

Verificación de metadatos de audio

5 **Campo técnico**

La presente invención se refiere al procesamiento de señales de audio, en particular a la verificación y corrección de los metadatos usados en tales sistemas. La invención es particularmente útil en los sistemas de codificación de audio conocidos como Dolby Digital (AC-3), Dolby Digital Plus, y Dolby E. Dolby; Dolby Digital, Dolby Digital Plus y 10 Dolby E son marcas registradas de Dolby Laboratories Licensing Corporation. Los aspectos de la invención también se pueden usar con otros tipos de codificación de audio, tales como MPEG-4 AAC.

Técnica anterior

15 Los detalles de la codificación Dolby Digital se exponen en las siguientes referencias:

ATSC Standard A52/A: Digital Audio Compression Standard (AC-3), Revision A, Advanced Television Systems Committee, 20 de agosto de 2001. El documento A/52A se encuentra disponible en Internet en <http://www.atsc.org/standards.html>.

20 «Flexible Perceptual Coding for Audio Transmission and Storage», de Craig C. Todd, y cols., 96th Convention of the Audio Engineering Society, 26 de febrero de 1994, republicación 3796;

25 «Design and Implementation of AC-3 Coders», de Steve Vernon, IEEE Trans. Consumer Electronics, Vol. 41, n.º 3, agosto de 1995.

«The AC-3 Multichannel Coder» de Mark Davis, Audio Engineering Society Preprint 3774, 95th AES Convention, octubre de 1993.

30 «High Quality, Low-Rate Audio Transform Coding for Transmission and Multimedia Application», de Bosi y cols., Audio Engineering Society Preprint 3365, 93rd AES Convention, octubre de 1992.

Patentes de EE.UU. 5.583.962, 5.632.005, 5.633.981, 5.727.119 y 6.021.386.

35 Los detalles de la codificación Dolby Digital Plus se exponen en «Introduction to Dolby Digital Plus, an Enhancement to the Dolby Digital Coding System», AES Convention Paper 6196, 117th AES Convention, 28 de octubre de 2004.

40 Los detalles de la codificación Dolby E se exponen en «Efficient Bit Allocation, Quantization, and Coding in an Audio Distribution System», AES Preprint 5068, 107th AES Conference, agosto de 1999 y «Professional Audio Coder Optimized for Use with Video», AES Preprint 5033, 107th AES Conference, agosto de 1999.

45 Los detalles de la codificación MPEG-2 AAC se exponen en ISO/IEC 13818-7:1997(E) «Information technology - Generic coding of moving pictures and associated audio information -, Part 7: Advanced Audio Coding (AAC)», International Standards Organization (April 1997); «MP3 and AAC Explained» de Karlheinz Brandenburg, AES 17th International Conference on High Quality Audio Coding, agosto de 1999; y «ISO/IEC MPEG-2 Advanced Audio Coding» de Bosi, y cols., AES preprint 4382, 101st AES Convention, octubre de 1996.

50 Se ofrece una visión general de diversos codificadores perceptuales, incluyendo los codificadores Dolby, los codificadores MPEG y otros en «Overview of MPEG Audio: Current and Future Standards for Low-Bit-Rate Audio Coding», de Karlheinz Brandenburg y Marina Bosi, J. Audio Eng. Soc., Vol. 45, n.º 1/2, enero/febrero de 1997.

55 Aunque la invención no se limita al uso en AC-3, por motivos de comodidad, se describirá en el entorno de un sistema AC-3. El sistema AC-3 es un sistema de compresión de datos de audio digital que se usa para el suministro de audio en aplicaciones tales como la televisión digital, el video DVD y el audio DVD. Un flujo de bits AC-3 consta de dos componentes principales: el contenido de audio y los metadatos. El contenido de audio de uno a seis canales son los datos comprimidos usando la codificación perceptual de audio. Entre los distintos tipos de metadatos en AC-3 se encuentran diversos parámetros de metadatos de audio que se pretende que cambien, de manera específica, el sonido del programa suministrado a un entorno de escucha. Estos parámetros se describen más adelante.

65 El sistema AC-3 suministra un flujo de bits compuesto de audio de datos comprimidos en tramas de información binaria. Cada trama comprende contenido de audio y metadatos para 1536 muestras de audio digital. Para una velocidad de muestreo de 48 kHz, esto representa 32 milisegundos de audio digital o una velocidad de 31,25 tramas por segundo de audio. El número de bits contenidos en cada trama depende del número de canales que se están

suministrando y de la cantidad de compresión de datos que se aplica a los canales. Por ejemplo, los discos de video DVD normalmente suministran seis canales de audio a una velocidad de datos de 448.000 bits por segundo o un tamaño de trama de 1792 bytes (un byte son 8 bits).

5 Cada trama AC-3 está dividida en secciones. Estas incluyen: (1) la Información de Sincronización (SI, por sus siglas en inglés), que contiene una palabra de sincronización (SW, por sus siglas en inglés), y la primera de dos palabras de corrección de error (CRC1); (2) la Información de flujo de bits (BSI, por sus siglas en inglés), que contiene la mayoría de los metadatos; (3) seis bloques de audio (AB0 a AB5), que comprenden contenido de audio de datos comprimidos; (4) los bits de desperdicio (W), que contienen cualquiera de los bits no usados que quedan una vez
10 que se ha comprimido el contenido de audio; (5) la información auxiliar (AUX), que contiene más metadatos; y (6) la segunda de las dos palabras de corrección de error (CRC2). Estas se muestran en la figura 9, la cual se describe más adelante. La trama AC-3, que incluye la compresión perceptual de datos de audio y los metadatos que la acompañan, se describe en detalle en las referencias AC-3 citadas anteriormente, y más adelante en la descripción de la figura 9.

15 Según se mencionó anteriormente, en el sistema AC-3 existen varios parámetros de metadatos de audio que se pretende cambien, de manera específica, el sonido del programa suministrado a un entorno de escucha. Tres de estos parámetros de metadatos se refieren al nivel de señal de reproducción y al rango dinámico: DIALNORM, COMPR y DYNRNG. El parámetro DIALNORM afecta el nivel de señal de reproducción de audio, mientras que los
20 parámetros relacionados COMPR y DYNRNG, en algunas ocasiones denominados en lo sucesivo los parámetros de «compresión de rango dinámico», afectan el rango dinámico de la señal de reproducción de audio. Uno o ninguno, aunque no ambos, de los parámetros COMPR y DYNRNG se usa en la descodificación, dependiendo de un modo de descodificación. Normalmente, el parámetro DIALNORM es ajustado por el usuario (no se genera de manera automática), aunque existiría un valor predeterminado DIALNORM si el usuario no ajustase ningún valor. Por
25 ejemplo, el usuario, o el «creador de contenido», podrían realizar mediciones de la intensidad acústica con un proceso o dispositivo externo al codificador AC-3 y, posteriormente, transferir el resultado hacia el codificador. De esta manera, existe una dependencia del usuario para ajustar el valor de parámetro DIALNORM de manera correcta. Los parámetros COMPR y DYNRNG, si bien están relacionados con el parámetro DIALNORM, normalmente se calculan de manera automática durante la codificación en respuesta al valor de parámetro DIALNORM ajustado por
30 el usuario y uno de una serie de perfiles de compresión de rango dinámico (o sin perfil, lo que se traduce en la aplicación de DIALNORM aunque permite la reproducción del rango dinámico completo). Cada perfil de este tipo contiene información estándar del parámetro de compresión de rango dinámico de audio, que incluye constantes de tiempo de ataque y liberación y relaciones de compresión. Entre otros parámetros de metadatos que afectan al sonido en un entorno de escucha se incluyen los distintos parámetros de «mezclado descendente»: CLEV, CMIXLEV, SLEV, SURMIXLEV, MIXLEVEL y MIXLEVEL2. Tales metadatos de mezclado descendente proporcionan
35 instrucciones a un descodificador para el mezclado descendente de canales 5.1 originales a un número más pequeño de canales de reproducción, por ejemplo, uno o dos canales.

40 El parámetro DIALNORM permite la reproducción uniforme del diálogo hablado durante la descodificación de cualquier flujo de bits AC-3. El nivel subjetivo de diálogo normal hablado se usa como referencia. De esta manera, la ganancia del sistema de reproducción se convierte en una función tanto del nivel de presión de sonido de reproducción deseada del oyente para el diálogo como del valor DIALNORM. Aunque, en principio, el valor DIALNORM se podría aplicar en el dominio de tiempo subsiguiente a la descodificación (ya sea en el dominio digital o en el dominio analógico) para ajustar la ganancia de reproducción, normalmente los descodificadores AC-3 usan el
45 valor DIALNORM en el dominio digital dentro del descodificador para escalar la ganancia, lo que se traduce en el ajuste de la ganancia de reproducción.

50 Si bien existen herramientas útiles para realizar una medida cómoda y fácil del nivel de diálogo en el contenido de audio (por ejemplo, el medidor de intensidad acústica Dolby LM100) y AC-3 proporciona metadatos para transmitir el nivel de diálogo (usando el parámetro DIALNORM), no existe forma de verificar que el valor DIALNORM en un flujo de bits AC-3 se haya ajustado de manera correcta y que coincida con el verdadero valor de intensidad acústica de diálogo del audio sin descodificar por completo el audio comprimido en PCM y realizar una medida de la intensidad acústica mediante una tecnología de medida aprobada. Este proceso de descodificación completa se describe en la
55 solicitud de patente de EE.UU. con número de serie 10/884.177, presentada el 1 de julio del 2004 de Smithers y cols., titulada «Method for Correcting the Playback Loudness and Dynamic Range of AC-3 (Dolby Digital) Compressed Audio Information».

60 Existen diversas razones distintas por las cuales el parámetro DIALNORM en un flujo de bits AC-3 podría ser incorrecto. En primer lugar, según se mencionó anteriormente, cada codificador AC-3 tiene un valor DIALNORM predeterminado que se usa durante la generación del flujo de bits si el valor DIALNORM no ha sido ajustado por el creador de contenido. Este valor predeterminado, elegido normalmente como -27 dB, podría ser sustancialmente diferente que el nivel real de intensidad acústica de diálogo del audio. En segundo lugar, incluso si el creador de contenido midiera la intensidad acústica y ajustara el valor DIALNORM, en consecuencia, se podría haber usado un algoritmo o medidor de la intensidad acústica que no sea conforme con el método recomendado de medida de
65 intensidad acústica AC-3, lo que daría lugar a un valor DIALNORM incorrecto. En tercer lugar, incluso si se ha creado un flujo de bits AC-3 con el valor DIALNORM medido y ha sido ajustado de manera correcta por el creador de

5 contenido, este podría haber sido cambiado a un valor incorrecto durante la transmisión y/o almacenamiento del flujo de bits. Por ejemplo, no es infrecuente en aplicaciones de radiodifusión de televisión que los flujos de bits AC-3 sean descodificados, modificados y posteriormente codificados de nuevo usando una información de metadatos DIALNORM incorrecta. Por lo tanto, mientras que un valor DIALNORM siempre está contenido en un flujo de bits AC-3, este podría ser incorrecto o inexacto y, por lo tanto, podría tener un impacto negativo sobre la calidad de la experiencia de escucha.

10 De este modo, existe la necesidad de una manera de verificar que el valor DIALNORM en un flujo de bits AC-3 haya sido ajustado de manera correcta por un creador de contenido y que no haya sido cambiado durante la distribución y la transmisión. De manera preferente, esta verificación no debería alterar la sintaxis estándar del flujo de bits AC-3, de manera que el flujo de bits siga siendo compatible con los descodificadores AC-3 existentes (es decir, se mantiene la compatibilidad con versiones anteriores).

15 En la referencia de Tim Carroll: «Audio Metadata: You can get there from here», 11 de octubre de 2004, páginas 1-4, URL: http://tvtechnology.com/features/audio_notes/f-TC-metadata-08.21.02.shtml, se comenta un aspecto de audio que acompaña a una señal de vídeo en un sistema de televisión de EE.UU. En tal sistema, el audio se distribuye desde una ubicación y tiempo de origen hasta una estación de televisión, en la que, con independencia de su formato (p. ej., monofónico, estereofónico, canales 5.1, etc.), en los EE. UU., debe estar codificado con Dolby AC-3 para acompañar a las señales de vídeo para la transmisión. La referencia señala que la distribución se puede realizar en forma de audio codificado con un sistema de codificación Dolby diferente, Dolby E, que permite la transmisión de los metadatos junto con el audio, y que los metadatos Dolby E, denominados «un flujo de metadatos válido», se pueden introducir en el codificador Dolby AC-3 con el fin de evitar la introducción manual o la aplicación de los valores predeterminados.

25 Descripción de los dibujos

La figura 1 es un diagrama de bloques esquemático funcional de una disposición para la generación de un flujo de bits de acuerdo con los aspectos de la presente invención.

30 La figura 2 es una representación abstracta de un ejemplo de un formato para los datos de verificación de metadatos en un flujo de bits.

35 La figura 3 representa la naturaleza de un diagrama de flujo de decisión que muestra los detalles de una etapa de decisión en el diagrama de flujo de la figura 4.

La figura 4 representa la naturaleza de un diagrama de flujo de decisión útil para comprender los aspectos de la invención que se refieren a garantizar que un flujo de bits tenga los metadatos correctos y los correspondientes datos de verificación de metadatos.

40 La figura 5 es un diagrama de bloques esquemático funcional de una disposición para la práctica de los distintos subconjuntos de las etapas 404 a 413 de la figura 4.

La figura 6a es un diagrama de bloque esquemático funcional que muestra una disposición para la práctica del subconjunto de las etapas 408 a 410 de la figura 4.

45 La figura 6b es un diagrama de bloque esquemático funcional que muestra una disposición para la práctica del subconjunto de las etapas 408 y 4011 a 413 de la figura 4.

50 La figura 7 es un diagrama de bloque esquemático funcional que muestra una disposición para la práctica de la función o dispositivo de reempaque de flujo de bits de la figura 6a.

La figura 8 representa la naturaleza de un diagrama de flujo de decisión útil para comprender los aspectos de la invención que se refieren a la descodificación conforme a los datos de verificación.

55 La figura 9a es un diagrama esquemático que ilustra una trama de un flujo codificado de bits en serie AC-3. No está a escala.

La figura 9b es un diagrama esquemático que ilustra en mayor detalle la parte SÍ de un flujo codificado de bits en serie AC-3. No está a escala.

60 La figura 9c es un diagrama esquemático que ilustra en mayor detalle la parte de información de encabezamiento de flujo de bits (BSI) de un flujo codificado de bits en serie AC-3. No está a escala.

65 La figura 9d es un diagrama esquemático que ilustra en mayor detalle una parte de bloque de audio de un flujo codificado de bits en serie AC-3. No está a escala.

La figura 9e es un diagrama de bloques esquemático funcional de un codificador o una función de codificación AC-3.

La figura 10a es una gráfica hipotética que muestra el nivel DIALNORM y el rango dinámico de tres elementos de audio de ejemplo.

5 La figura 10b es una gráfica hipotética que muestra el nivel DIALNORM y el rango dinámico de tres elementos de audio de ejemplo durante la reproducción.

10 La figura 11a es un gráfico hipotético que muestra el efecto de los parámetros de control del rango dinámico en tres elementos de audio de ejemplo.

La figura 11b es un gráfico hipotético que muestra el efecto de los parámetros DIALNORM y de control del rango dinámico en tres elementos de audio de ejemplo durante la reproducción.

15 **Exposición de la invención**

Se puede considerar que la invención tiene una serie de aspectos, todos los cuales involucran información de verificación de metadatos de audio. Algunos de estos aspectos incluyen lo siguiente:

20 (1) un flujo codificado de bits de audio que tiene los metadatos correctos y la información que verifica la corrección de al menos una parte de los metadatos;

(2) un proceso o dispositivo que genera un flujo codificado de bits de audio que tiene los metadatos correctos y la información que verifica la corrección de al menos una parte de los metadatos;

25 (3) un proceso o dispositivo que asegura que un flujo codificado de bits de audio tenga los metadatos correctos y también que contenga la información que verifica la corrección de al menos una parte de los metadatos correctos; y

30 (4) un proceso o dispositivo que descodifica un flujo codificado de bits de audio independientemente de que todos sus metadatos sean correctos o no, genera y sustituye los metadatos corregidos y tiene en cuenta, si existe, la información que verifica la corrección de al menos una parte de los metadatos.

Otros aspectos de la invención se exponen en las reivindicaciones y en la descripción escrita y las figuras.

35 Se debe indicar que la información de verificación de metadatos de audio no sirve para la función de proporcionar la detección y/o corrección de error de bits. Normalmente, los flujos de bits, en los cuales se transporta la información de verificación, tienen algún tipo de detección y/o corrección de error de bits, por ejemplo, las palabras de código CRC en un flujo de bits AC-3. En aspectos de la presente invención, los metadatos podrían ser incorrectos, por ejemplo, debido a que inicialmente no se ajustaron correctamente o, incluso si inicialmente se hubiesen ajustado correctamente, a que hayan cambiado durante la transmisión o almacenamiento como resultado de la intervención humana o por otro motivo, no debido a los errores de bits en la transmisión o almacenamiento. En su lugar, la información de verificación de metadatos de audio no serviría para el propósito de la detección o corrección de error de bits debido a que sirve para el propósito de cambio de los metadatos, si no fueran correctos, a un valor correcto. La corrección de error de bits simplemente corregiría los errores de bits en los metadatos, dejándolos incorrectos, aunque sin errores de bits. En otras palabras, la información de verificación de metadatos de audio se refiere a la corrección de la información subyacente de los bits que representan los metadatos, no a la corrección de los bits en sí mismos.

50 Aunque en ejemplos de los aspectos de la invención descritos en este documento el flujo codificado de bits de audio es un flujo codificado de bits Dolby Digital (AC-3), los conjuntos de metadatos son los metadatos DIALNORM y de control de rango dinámico relacionados, y la información de verificación corresponde con los metadatos DIALNORM correctos, los aspectos de la invención se pueden aplicar a otros sistemas de codificación de audio y a otros metadatos en los flujos de bits de tales sistemas de codificación. Otros sistemas de codificación de audio en los cuales los aspectos de la invención podrían resultar útiles incluyen, por ejemplo, el sistema Dolby E y el sistema MPEG-4 AAC. Con respecto al sistema Dolby Digital, los metadatos podrían ser los metadatos de mezclado descendente, además o en lugar de los metadatos DIALNORM y los metadatos de control de rango dinámico relacionados (en cuyo caso la información de verificación también se refiere a o en su lugar se refiere a los metadatos de mezclado descendente).

60 La información de verificación se podría transportar en el flujo codificado de bits de audio, de tal manera que el flujo de bits sea compatible con versiones anteriores de los procesos y dispositivos existentes o heredados. En ejemplos de los aspectos de la invención descrita en este documento, la información de verificación se transporta en los bits de desperdicio AC-3 mencionados anteriormente, que de otra manera no podrían transportar información útil y normalmente son ignorados por los descodificadores AC-3 estándar. Otros sistemas de codificación de audio podrían tener bits de «desperdicio» o bits similares que pudieran estar disponibles (en algunas ocasiones denominados bits «nulos», bits de «relleno» o similares) y que normalmente son ignorados por los descodificadores

estándar, por ejemplo, los campos de datos adicionales tales como el elemento de flujo de datos en MPEG-4 AAC, una sección de datos definidos por el usuario. Sin embargo, tales bits en algunos sistemas de codificación podrían no ser destruidos por una operación de codificación/descodificación, lo cual es una característica útil de los aspectos de la presente invención cuando se incluyen en un sistema de codificación AC-3. Las técnicas para transportar datos en bits de «desperdicio» o similares en flujos de bits codificados se describen en la patente de EE.UU. n.º 6.807.528 B1, «Adding Data to a Compressed Data Frame», de Truman y cols.

La información de verificación también se podría transportar en el flujo codificado de bits de audio, de tal manera que esté «escondida». Por ejemplo, la información de verificación transportada en los bits de desperdicios podría estar cifrada. El ocultamiento de los datos de verificación tiene la ventaja de que alguien que cambie a propósito un valor DIALNORM en un flujo codificado de bits tendrá dificultad para cambiar, o no será capaz de cambiar, la información de verificación.

Aunque no todas las tramas AC-3 podrían tener suficientes bits de datos no usados para transmitir la información adicional, este no es un problema cuando el valor DIALNORM sea constante dentro de un programa completo; es suficiente que al menos algunas tramas AC-3 tengan suficientes bits de datos no usados para usarlos en los datos de verificación.

De manera alternativa, en lugar de transportar la información de verificación en bits de desperdicio que podrían estar codificados, esta se podría codificar de manera esteganográfica en el flujo de bits usando técnicas tales como las descritas en la solicitud de patente de EE.UU. con número de serie 10/344.388, presentada el 15 de agosto del 2001 (PCT), titulada «Modulating One or More Parameters of an Audio or Video Perceptual Coding System in Response to Supplemental Information», de Watson y cols., publicada el 5 de febrero del 2004 como US 2004/0024588 A1. La codificación esteganográfica tiene la ventaja de que preserva la compatibilidad con versiones anteriores y también oculta los datos. Sin embargo, la descodificación y nueva codificación del flujo de bits no podrían borrar o «limpiar» la información de verificación (como se comenta más adelante), lo cual es una desventaja.

Los datos de verificación, en su forma más simple, podrían ser una copia del valor DIALNORM correcto (junto con los datos adecuados de formación de tramas o sincronización e identificación). Debido a que los bits no usados en un flujo de bits AC-3 normalmente están ajustados en valores nulos o aleatorios, es altamente inexistente la probabilidad de que coincidan los bits de datos no usados en un flujo de bits AC-3 con el formato de datos de verificación DIALNORM. Asimismo, cuando solo se use un valor DIALNORM constante por programa AC-3 codificado, como es típico, la información de verificación DIALNORM contenida en los bits de datos no usados también es un valor constante fijo. En este caso, la verificación de múltiples instancias de datos de verificación DIALNORM en una serie de tramas AC-3 disminuye la probabilidad que los bits de datos no usados sean erróneos para los bits de datos de verificación.

Si la medida DIALNORM, la generación de metadatos y la inserción de datos de verificación se realizan en tiempo real, de manera continua, en un flujo de bits AC-3, no se podría presentar un valor DIALNORM constante en todo el programa. En este caso, la verificación se podría realizar analizando una serie de tramas AC-3 (que podrían contener diversos valores de metadatos DIALNORM) y comprobando que coincidan los campos de verificación DIALNORM colocados en los bits de datos no usados, cuando se encuentran disponibles, con los valores DIALNORM. Podría ser necesario un número mínimo de campos DIALNORM y de datos de verificación DIALNORM coincidentes con el fin de reducir la probabilidad que los datos aleatorios en los campos de datos no usados coincidan con los valores del parámetro DIALNORM.

Un aspecto adicional de esta invención es, de manera preferente, permitir que solo procesos o dispositivos aprobados escriban los datos de verificación DIALNORM en un flujo de bits AC-3. De esta manera, se garantiza la validez de los datos de verificación. Por lo tanto, aunque la corrección del valor de parámetro de metadatos DIALNORM no esté garantizada por razones tales como las mencionadas anteriormente, los datos de verificación DIALNORM se pueden usar con confianza en su exactitud. Además, el problema de la invalidación de un parámetro DIALNORM correcto se soluciona debido a que los datos de verificación DIALNORM se colocan en bits de datos no usados de otra manera de un flujo de bits AC-3. Si un flujo de bits AC-3, que contiene datos de verificación válidos se descodificase y se codificase de nuevo, entonces sería altamente improbable que los bits de datos no usados resultantes que reemplazan a los datos de verificación DIALNORM como resultado de la nueva codificación siguiesen siendo correctos, incluso en el caso de que se usasen los mismos valores de metadatos AC-3 DIALNORM. Esto garantiza que cualquier procesamiento adicional de un flujo de bits AC-3 verificado «limpie» los datos de verificación (a menos que se use para el nuevo procesamiento un codificador AC-3 aprobado con capacidades de medida y verificación de intensidad acústica, como se explica más adelante).

Estos y otros aspectos de la invención resultarán más claros con la lectura y comprensión de las siguientes maneras de llevar a cabo la invención.

Mejor modo de llevar a cabo la invención

Generación de un flujo de bits AC-3 que contiene datos DIALNORM y de verificación de coincidencia correctos

Este aspecto de la invención se refiere a la creación de un flujo de bits AC-3 que tiene un valor de parámetro DIALNORM correcto y que tiene datos de verificación de DIALNORM coincidentes.

5 La figura 1 muestra una disposición 100 que comprende dos elementos, una función de codificación AC-3 modificada o un codificador AC-3 modificado («codificación AC-3 modificada») 102 y una función de medida de nivel de diálogo o medidor de nivel de diálogo («nivel de medida de diálogo») 104. El audio PCM 101 se aplica en el codificador AC-3 modificado 102 y el nivel de medida de diálogo 104. El codificador AC-3 modificado podría ser el mismo que la función de codificación o el codificador AC-3 estándar, excepto porque también es capaz de aceptar
10 datos de verificación DIALNORM y de introducirlos en el flujo de bits AC-3 de alguna manera adecuada, como se comentó anteriormente. La codificación AC-3 modificada proporciona una salida de flujo de bits AC-3 compatible con versiones anteriores que incluye los datos de verificación DIALNORM. El nivel de medida de diálogo 104 analiza la PCM de entrada, después calcula el valor DIALNORM correcto y, finalmente, lo envía (por medio de 103) a la codificación AC-3 modificada 102.

15 Durante la codificación AC-3 normal, el número de bits no utilizados disponibles está directamente relacionado con la complejidad del audio (es decir, con el grado de dificultad de la codificación del audio a una velocidad de bits deseada). Debido a que el número de bits disponible por trama de audio AC-3 es fijo, cuanto más difícil es la codificación del audio, mayor cantidad de bits se usan para conseguir un nivel de calidad y, por lo tanto, menor
20 cantidad de bits se usan en el proceso de codificación y se encuentran disponibles para transportar los datos de verificación DIALNORM. Por consiguiente, las señales de audio que son más sencillas de codificar tendrán más bits de datos no usados disponibles para el almacenamiento de los datos de verificación DIALNORM. Por lo tanto, una modificación opcional, aunque útil, del codificador AC-3 modificado consiste en la capacidad para especificar un número mínimo de bits de datos que el proceso de codificación no usará durante la codificación. Dado el pequeño
25 número de bits requeridos para transmitir los datos de verificación DIALNORM (como se describe más adelante), conservar a propósito algunos bits de datos no usados podría tener poco o ningún efecto sobre la calidad de la señal de audio codificada.

Nivel de medida de diálogo 104

30 Se podría realizar una medida del nivel de intensidad acústica del diálogo aislando en primer lugar segmentos del contenido de audio que comprendan, de manera predominante, conversación o voz. Este método se describe en la solicitud de patente de EE.UU. con número de serie 10/233.073, de Vinton y cols., titulada «Controlling Loudness of Speech in Signals That Contain Speech and Other Types of Audio Information», publicada el 4 de marzo como
35 documento US 2004/0044525 A1. Sin embargo, se podrían usar otros métodos. A continuación, los segmentos de audio que predominantemente son de conversación o voz se pasan a un algoritmo de medida de intensidad acústica. En AC-3, este algoritmo es una medida estándar de la potencia ponderada A. También se podrían usar otras medidas de intensidad acústica, entre las que incluyen las mediciones estándar de potencia ponderada B o C, o las basadas en modelos psicoacústicos de intensidad acústica. La medida de potencia se calcula con relación a una onda sinusoidal digital de audio de escala completa (0 dB FS).

40 El aislamiento de los segmentos de conversación o voz no es esencial; sin embargo, mejora la exactitud de la medida y proporciona resultados más satisfactorios desde la perspectiva del oyente. Debido a que no todo el contenido de audio comprende conversación o voz, la medida de intensidad acústica de la totalidad del contenido de audio podría proporcionar una aproximación suficiente del nivel de diálogo del audio, en el caso de que hubiera estado presente la conversación o la voz.

45 Si el método está operando en un flujo de bits continuo, en lugar de un flujo de bits de longitud finita, esta medida se podría actualizar de manera continua y podría representar el nivel de diálogo, por ejemplo, solo para los últimos pocos segundos. Si el método está operando sobre la base de un flujo de bits de longitud finita previamente almacenado (tal como un archivo de audio almacenado en un disco duro), entonces se podría analizar el programa completo y calcular un valor DIALNORM único.

Codificación AC-3 modificada 102

55 La PCM de audio de entrada se codifica usando la codificación AC-3 modificada que usa el valor o los valores calculados para ajustar el DIALNORM de flujo de bits y los parámetros de metadatos de compresión de rango dinámico relacionados. El codificador AC-3 modificado podría ser el mismo que el codificador AC-3 normal, excepto porque una función o dispositivo adicional de medida de intensidad acústica 104, según se describió anteriormente,
60 mide, de manera explícita y correcta, el valor de parámetro DIALNORM y lo proporciona al codificador para su inclusión en el flujo de bits. El codificador modificado también crea e introduce datos de verificación DIALNORM en los bits de datos no usados de otra manera del AC-3, debido a que, en este ejemplo, un proceso de medida DIALNORM aprobado ha proporcionado una medida objetiva. De manera preferente, el flujo de bits AC-3 producido por el codificador AC-3 modificado 102 es conforme a los estándares de un flujo de bits AC-3 definido en el
65 documento A/52A citado anteriormente, haciéndolo compatible con versiones anteriores con los descodificadores AC-3 existentes.

Si el codificador AC-3 modificado también tuviera la capacidad de especificar un número mínimo de bits de datos no usados, esto se podría implementar cambiando el valor del número total de bits por trama AC-3 disponible para la codificación de audio. Por ejemplo, si el número de bits disponibles para el proceso de codificación AC-3 es normalmente $N_TOTAL_ENCODE_BITS$ y se desea tener al menos $N_TOTAL_VERIFICATION_BITS$, entonces el nuevo número total de bits disponibles de codificación será $(N_TOTAL_ENCODE_BITS - N_TOTAL_VERIFICATION_BITS)$ y el proceso de codificación de audio continúa como es habitual.

Formato de datos de verificación DIALNORM

Con el fin de que los datos de verificación DIALNORM se puedan identificar y leer con facilidad a partir de un flujo de bits AC-3 sin descodificación, sería útil que los datos tuviesen un formato predefinido. La figura 2 expone un formato adecuado para el almacenamiento de los datos de verificación DIALNORM de una manera alineada por byte que simplifica la ubicación y la lectura de los datos de un flujo de bits AC-3 no descodificado (ya sea en un flujo de bits AC-3 en tiempo real o en un flujo de bits AC-3 que se almacene como un archivo digital). El formato no es crítico y se podrían usar otros formatos. Según se muestra en la figura 2, el formato de ejemplo para los datos de verificación DIALNORM consta de varios bytes consecutivos. El primer byte es un byte de encabezamiento de verificación DIALNORM predefinido. Este byte de encabezamiento podría tener cualquier valor. Sin embargo, es preferible un valor diferente de cero (similar, aunque no el mismo que AC-3 SYNCWORD), debido a que los bits de datos no usados se podrían haber inicializado con valores de cero en otros flujos de bits AC-3. Después del encabezamiento de verificación DIALNORM en este ejemplo, los bytes de datos se usan para transmitir los datos de verificación DIALNORM y los datos opcionales adicionales. Debido a que el valor AC-3 DIALNORM estándar comprende 5 bits, un byte de datos proporciona 3 bits de datos adicionales y 2 bytes de datos proporcionan 11 bits de datos adicionales. El uso de dos bytes de datos de verificación DIALNORM se muestra en el ejemplo de la figura 2. Estos bytes se podrían usar para almacenar información, tal como el tipo o la versión de algoritmo de intensidad acústica utilizado u otra información. El byte final que se muestra en la figura 2 es un byte de datos de verificación de redundancia de ciclo (CRC) que se calcula usando un encabezamiento de verificación DIALNORM y bytes de datos. Este byte resulta útil porque reduce en gran medida la probabilidad de que los bits de datos no usados en un flujo de bits AC-3 (que contiene una secuencia de bytes que tienen un byte de encabezamiento de verificación DIALNORM válido, 2 bytes de datos intermedios y un byte CRC) pasen una verificación CRC en todos los cuatro bytes.

Como se comentó anteriormente, si un codificador AC-3 modificado reserva suficientes bits de datos no utilizados para contener los datos de verificación DIALNORM, dada la estructura expuesta en la figura 2, esto solamente requiere 4 bytes o 32 bits para cada trama de datos AC-3 de 1792 bytes, lo cual corresponde solamente al 0,2 % de los datos totales.

Cómo asegurar que un flujo de bits AC-3 contenga datos DIALNORM y de verificación de coincidencia correctos

Otro aspecto de la invención consiste en asegurar que el valor DIALNORM en un flujo de bits AC-3 sea correcto y que el flujo de bits contenga la información de verificación DIALNORM coincidente. Este aspecto de la invención se expone en el diagrama de flujo de ejemplo de la figura 4. Como se explica más adelante, se podría usar ya sea la totalidad del proceso de la figura 4 o subconjuntos del proceso de la figura 4. Estos procesos o dispositivos que usan etapas de los procesos podrían ser útiles, por ejemplo, en la transmisión o almacenamiento de un flujo de bits, de manera subsiguiente a la creación de un flujo de bits por un creador de contenido y antes de la descodificación final del flujo de bits para un oyente. Se entenderá que las etapas de la figura 4 o los subconjuntos de la misma podrían representar partes de uno o más procesos o podrían ser funciones realizadas en uno o más dispositivos.

Las etapas de la figura 4 se podrían realizar sobre la base de un flujo de bits que represente el audio de un elemento de audio de longitud finita. Por ejemplo, un elemento de audio consta de un programa o anuncios de televisión que están almacenados de manera digital en un servidor de archivos o en otro tipo de soporte. Como se usa en este documento, el término «elemento de audio» es un fragmento continuo de información de audio; por ejemplo, un anuncio de televisión de 30 segundos o una película completa (imagen animada). Sin embargo, las etapas de la figura 4 también se podrían usar para medir y actualizar un flujo de bits continuo en tiempo real de tramas AC-3, por ejemplo un flujo de bits AC-3 continuo que represente el audio de una estación o canal de televisión.

Prueba de existencia de datos de verificación DIALNORM (Etapas 401 de la figura 4 y de la figura 3)

Según se muestra en la figura 4, la primera etapa realizada (etapa 401) es la determinación de si existen datos de verificación AC-3 DIALNORM en un flujo de bits AC-3 codificado. La figura muestra un diagrama de flujo de ejemplo para realizar tal verificación. Según se muestra en la figura 3, la entrada es un flujo de bits de audio AC-3, el cual se puede procesar sobre una base de trama por trama. Debido a que se conoce la ubicación de los bits de datos no usados dentro de una trama AC-3, no es necesario realizar una búsqueda exhaustiva de una trama o flujo de bits completo, sino que la búsqueda podría comenzar en el inicio de la sección o secciones de bits no usados. Aunque los datos de verificación DIALNORM constan de bytes de datos consecutivos, estos datos podrían estar o no alineados por byte con otros datos de tramas AC-3. Por lo tanto, la primera etapa (etapa 301) en el proceso de la figura 3, «LEER DATOS DEL FLUJO DE BITS», podría requerir la lectura de los datos bit por bit y construir bytes de

datos consecutivos a partir de cada bit leído.

En la etapa 303 del ejemplo de la figura 3, cada byte de datos leído de la trama AC-3 se compara con el byte de encabezamiento de verificación DIALNORM predefinido. Si no coincide un byte, entonces se leen más datos y se realiza otra comparación de bytes. Si el valor del byte coincide con el valor del encabezamiento, entonces se leen los bytes de datos consecutivos seguidos del byte coincidente. Si el byte que coincide con el byte de encabezamiento de verificación, está cerca del final de la trama AC-3, según se determina en la etapa 302, y no están disponibles tres bytes de datos siguientes, entonces se detiene la búsqueda de datos. Los cuatro bytes (que incluyen el byte de encabezamiento de verificación coincidente según se expone en la figura 2) se utilizan para calcular una verificación CRC en la etapa 304. Si la verificación CRC es correcta (etapa 305), entonces existen los datos de verificación DIALNORM y la información de verificación DIALNORM se podría recuperar de los bytes de datos como se describe más adelante.

Prueba de si los datos de verificación DIALNORM coinciden con los datos AC-3 DIALNORM (402)

Según se muestra en la etapa 401 de la figura 4, cuando existan datos de verificación DIALNORM (salida SÍ de la etapa 401), la siguiente etapa, etapa 402, consiste en determinar si los datos de verificación coinciden con el valor AC-3 DIALNORM. Según se muestra en la figura 9, y como se comenta más adelante, se conocen la ubicación y el formato de los datos DIALNORM normales y se pueden leer con facilidad a partir del flujo de bits AC-3. La prueba para determinar si coinciden los valores DIALNORM normal y DIALNORM de verificación es una comparación numérica simple. Si los valores coinciden, entonces el valor normal DIALNORM es correcto y no es necesario realizar un análisis o procesamiento adicional. Se puede considerar que los valores «coinciden» lo suficiente si el valor absoluto de la diferencia entre los dos valores es menor que un umbral. Si así se desea, este umbral se puede ajustar como igual a cero, aunque en las implementaciones preferidas se elige un umbral para lograr una relación de equilibrio entre la exactitud de los parámetros de metadatos, el coste de los recursos de cálculo necesarios para implementar la presente invención y la posibilidad de que la diferencia entre el valor DIALNORM y el valor de verificación degrade la calidad de la información de audio durante la reproducción. Un valor umbral de tres (3 dB) podría ser adecuado para muchas aplicaciones. La salida de flujo de bits AC-3 se podría almacenar, transmitir o descodificar.

Corrección de datos AC-3 DIALNORM normales con datos de verificación DIALNORM (etapa 403)

Cuando los datos de verificación DIALNORM se extraen del flujo de bits AC-3 y no coinciden con los metadatos AC-3 DIALNORM normales (salida NO de la etapa 402), entonces los metadatos DIALNORM se actualizan (es decir, se ajustan o se igualan) con la verificación del valor DIALNORM (403). Debido a que se ha determinado que el valor DIALNORM normal es incorrecto y es necesario actualizarlo, es posible que los metadatos de compresión de rango dinámico AC-3 relacionados también sean incorrectos. Por lo tanto, es necesario analizar la información de compresión de rango dinámico y, si es correcta, solamente se actualiza el parámetro de metadatos DIALNORM. Si la información de compresión de rango dinámico es incorrecta, entonces también es necesario actualizarla. Los detalles de este análisis y actualización se explican más adelante en conexión con las etapas 409 a 413 de las figuras 4, figuras 6a, 6b y 9c).

Datos de verificación no presentes en el flujo de bits AC-3

Metadatos DIALNORM correctos (etapas 404-407)

Según se muestra en el ejemplo de la figura 4, si los datos de verificación DIALNORM no están contenidos dentro del flujo de bits AC-3 (salida NO en la etapa 401), entonces el flujo de bits AC-3 se descodifica en PCM sin aplicar el parámetro DIALNORM y los parámetros de control de rango dinámico relacionados (debido a que estos parámetros de metadatos podrían ser incorrectos), de manera que el contenido de audio descodificado se encuentra al mismo nivel que esa entrada al codificador que se usó para crear el flujo de bits) (etapa 404). La intensidad acústica del diálogo se mide entonces con el fin de determinar el nivel DIALNORM correcto (etapa 405). Esta medida se podría conseguir mediante el dispositivo o función del nivel de medida de diálogo 104 descrito anteriormente. Después de la medida del nivel de diálogo en la etapa 405, el valor medido se compara con los metadatos AC-3 DIALNORM en la etapa 406 (más adelante se describen los detalles de esta comparación). Como se muestra después de la salida SÍ de la etapa 406, si los metadatos normales DIALNORM son correctos, entonces la única acción requerida es formatear y almacenar los datos de verificación DIALNORM en el flujo de bits AC-3 (etapa 407).

Como se comentó anteriormente, el número de bits de datos no usados disponibles depende de la complejidad del audio y algunas tramas AC-3 podrían no tener suficientes bits de datos no usados para almacenar los datos de verificación DIALNORM en cada trama. Son posibles dos opciones: almacenar los datos de verificación solo en tramas AC-3 con suficientes bits de datos no usados (en cuyo caso los datos de verificación DIALNORM se podrían introducir en el flujo de bits AC-3 de entrada original en lugar de un flujo de bits AC-3 que resulta de la nueva codificación del flujo de bits AC-3 descodificado en la etapa 404) o codificar nuevamente el audio que resulta de la descodificación AC-3 de la etapa 404, reservando un número suficiente de bits de datos no usados para garantizar que los datos de verificación quepan en cada trama.

Una alternativa a la descodificación AC-3 de la etapa 404 y la medida de la intensidad acústica de la etapa 405 consiste en la obtención de una aproximación de la intensidad acústica a través de una técnica que no requiere una descodificación completa del flujo de bits AC-3. Esta técnica, que descodifica parcialmente un flujo de bits tal como un flujo de bits AC-3 con el fin de obtener una estimación somera del espectro de audio codificado sobre la base de la magnitud de los exponentes de sub-banda, se describe en la solicitud de patente provisional de EE.UU. de Brett Graham Crockett, Michael John Smithers, Alan Jeffrey Seefeldt, referencia de agente DOL157, presentada el mismo día que la presente solicitud.

La figura 5 es un ejemplo de una disposición 500 para la práctica de los distintos subconjuntos de las etapas 404 a 413 de la figura 4. Según se muestra en la figura 5, las tramas AC-3 501 se descodifican a través de una función de descodificación o descodificador AC-3 modificado («descodificador AC-3») 502 en audio digital 503. Durante la descodificación de las tramas AC-3 por el descodificador AC-3 502, el parámetro DIALNORM y la información de compresión de rango dinámico, aunque se recuperan para otro uso posible, según se describe más adelante, se ignoran a los efectos de descodificación de audio, de manera que el audio descodificado 503 se encuentre al mismo nivel y tenga el mismo rango dinámico que la entrada al codificador que se usó para crear el flujo de bits. Una función de medida de nivel de diálogo o medidor de nivel de diálogo («nivel de medida de diálogo») 504 recibe el audio descodificado 503 y calcula el nivel del diálogo 505. La medida de diálogo 504 podría ser la misma función o dispositivo que la medida de diálogo 104 descrita anteriormente en conexión con la figura 1. El descodificador AC-3 502 podría realizar la etapa 404 según se describió anteriormente, y el nivel de medida de diálogo 504 podría efectuar la etapa 405, según se describió anteriormente. Una función de actualización de flujo de bits o actualizador («flujo de bits de actualización») 506 compara el nivel de diálogo con el parámetro DIALNORM presente en cada trama. Más adelante se describen detalles adicionales de la comparación. Además, en función de las decisiones de las etapas 406 y 408, este también realiza las etapas 407, 408-410 (véase la figura 6a y su descripción más adelante) o las etapas 408 y 411-413 (véase la figura 6b y su descripción más adelante). Cuando se realiza la etapa 407, introduce la información de verificación DIALNORM en el flujo de bits AC-3 de entrada, dejando la información DIALNORM original y la información de control de rango dinámico relacionada. Durante la realización de la etapa 407, el flujo de bits de actualización 506 también busca las tramas AC-3 para detectar bits de datos no usados. Las tramas AC-3 con un número suficiente de bits de datos no usados se modifican, de manera que los bits no usados se actualizan para contener los datos de verificación DIALNORM. De manera alternativa, el audio descodificado producido por el descodificador AC-3 502 se podría codificar nuevamente, reservando un número suficiente de bits de datos no usados para garantizar que los datos de verificación quepan en cada trama (en este caso, el flujo de bits de actualización 506 incluye un codificador AC-3 modificado tal como el codificador AC-3 modificado 102 de la figura 1).

De manera más específica, durante la realización de la etapa 406, el flujo de bits de actualización 506 compara el nivel medido del diálogo con el nivel de diálogo indicado por el parámetro DIALNORM. El parámetro DIALNORM tiene un rango desde -31 dB hasta -1 dB inclusive, en incrementos de 1 dB. Si el nivel medido del diálogo está dentro de este rango y es diferente del valor de DIALNORM del flujo de bits, el parámetro DIALNORM se actualiza de manera condicional (es «condicional» tras la determinación de si están disponibles suficientes bits no usados como para transportar la información de verificación) el nivel medido, redondeado, por ejemplo, al 1 dB más cercano. Se puede considerar que el nivel medido del diálogo es diferente del valor DIALNORM en el flujo de bits si el valor absoluto de la diferencia entre los dos valores es menor que un umbral. Si así se desea, este umbral se puede ajustar como igual a cero, aunque en las implementaciones preferidas se elige un umbral para lograr una relación de equilibrio entre la exactitud de los parámetros de metadatos, el coste de los recursos de cálculo necesarios para implementar la presente invención y la posibilidad de que la diferencia entre el valor DIALNORM y el nivel de diálogo medido degrade la calidad de la información de audio durante la reproducción. Un valor umbral de tres (3 dB) podría ser adecuado para muchas aplicaciones. Además, para actualizar el flujo de bits de manera que tenga el parámetro DIALNORM correcto, el flujo de bits de actualización 506 también busca los bits de datos no usados en cada trama AC-3. Si una trama contiene un número suficiente de bits de datos no usados, estos se reemplazan por los datos de verificación DIALNORM, indicando que se ha realizado un proceso de medida de la intensidad acústica exacto y aprobado y que el valor DIALNORM embebido en el flujo de bits AC-3 es correcto.

Datos de verificación no presentes en el flujo de bits AC-3

Metadatos DIALNORM incorrectos

Intensidad acústica dentro del rango del parámetro DIALNORM (etapas 408-410)

Según se muestra en el ejemplo de la figura 4, si no existe la información de verificación (salida NO de la etapa 401), y si el valor AC-3 DIALNORM existente es incorrecto (salida NO de la etapa 406), entonces se debe determinar si el nivel de intensidad acústica medido se encuentra dentro del rango válido del parámetro DIALNORM (etapa 408). El parámetro DIALNORM no tiene alcance suficiente para transmitir el nivel cuando el nivel medido del diálogo 505 se encuentra fuera del rango válido del parámetro DIALNORM como se permite en el flujo de bits AC-3. Es decir, el nivel medido es menor que -31 dB o mayor que -1 dB. Si el parámetro DIALNORM tuviera un alcance suficiente para transmitir el nivel (salida SÍ de la etapa 408), entonces las etapas 409 y 410 se realizan según se explica a

continuación, mostrándose detalles adicionales de las mismas en conexión con las figuras 6a y 7.

La figura 6a muestra cómo se determina la nueva información de compresión de rango dinámico (etapa 409) y cómo se actualiza el flujo de bits y se empaqueta nuevamente (etapa 410) cuando se cambia el valor de DIALNORM y se introducen los datos de verificación DIALNORM. Según se indicó anteriormente, el ejemplo de la figura 6a es una variación del flujo de bits de actualización 507 de la figura 5 que resulta útil para realizar las etapas 408-410. Los elementos de la figura 6a se pueden representar de la siguiente manera.

Extracción de DIALNORM 602

El valor del parámetro DIALNORM se extrae del flujo de bits AC-3, como se indicó en las figuras 5 y 6a, el flujo de bits no descodificado 501 se aplica al dispositivo o función de extracción de DIALNORM 602 («Extracción de DIALNORM» 602).

Determinación de perfiles de compresión de rango dinámico 604

Según se muestra en la figura 6a, un dispositivo o función de determinación del perfil de compresión de rango dinámico («Determinación del perfil de compresión de rango dinámico» 604) recibe el valor de parámetro DIALNORM extraído del flujo de bits no descodificado y la salida del descodificador AC-3 (502 de la figura 5) y determina un perfil de compresión de rango dinámico. Los metadatos de compresión de rango dinámico en una trama AC-3 representan los cambios de ganancia que se pueden aplicar al contenido de audio durante la descodificación. Estos metadatos existen como dos parámetros diferentes. El parámetro COMPR en la información de flujo de bits (BSI) tiene un rango de -48,14 a +47,88 dB y es una escala que se aplica en una trama de audio completa. El parámetro DYNRNG, uno en cada bloque de audio (AB), tiene un rango de -24,06 a +23,94 dB y proporciona un medio de escala de manera independiente en cada bloque. Uno o ninguno, aunque no ambos, de estos parámetros se utilizan en el descodificador, dependiendo del modo de descodificación.

Según se mencionó anteriormente, los parámetros COMPR y DYNRNG se calculan durante el proceso de codificación usando el parámetro DIALNORM y ninguno o uno de una serie de perfiles de compresión de rango dinámico. Cada perfil contiene información estándar del parámetro de compresión de rango dinámico de audio, que incluye constantes de tiempo de ataque y liberación y relaciones de compresión.

Debido a que se cambia el parámetro DIALNORM, es posible que los valores de los parámetros COMPR y DYNRNG en el flujo de bits ya no sean correctos. Los parámetros COMPR y DYNRNG se podrían dejar inalterados en el flujo de bits, aunque el audio en la reproducción podría presentar graves y molestas fluctuaciones de ganancia y/o podría conducir a una sobrecarga del descodificador (o corte de señales digitales). Un mejor proceso consiste en actualizar los parámetros COMPR y DYNRNG. Esto se consigue de mejor manera conociendo el perfil de compresión de rango dinámico usado para calcular sus valores originales. Debido a que la información acerca del perfil no está presente en el flujo de bits, se podría elegir un perfil arbitrario (que incluye la desactivación de toda la compresión de rango dinámico), o el perfil se podría inferir a partir de los valores COMPR and DYNRNG originales en el flujo de bits. Al inferir el perfil, podría coincidir de manera más estrecha la intención original del creador de contenido con respecto a la compresión de rango dinámico.

En la determinación del perfil de compresión de rango dinámico 604, el audio descodificado 503 y el valor DIALNORM original 603 se usan conjuntamente para calcular múltiples conjuntos de valores COMPR y DYNRNG, un conjunto para cada perfil que se sabe que existe en los codificadores AC-3. El número de índice del perfil cuyo conjunto de valores COMPR y DYNRNG coincide de manera más estrecha con los valores COMPR y DYNRNG en el flujo de bits original se envía como salida 606.

Si este método está operando sobre un flujo continuo de tramas, el índice de perfil se podría actualizar de manera continua. Por ejemplo, podría representar el perfil más probable durante varios segundos previos de tramas.

Es posible que el perfil de compresión de rango dinámico estimado no sea el mismo que el perfil usado originalmente. Por lo tanto, podría ser deseable actualizar la información DIALNORM y de compresión de rango dinámico solo si la diferencia absoluta entre el nivel medido del diálogo y el valor DIALNORM fuera más grande que un umbral, según se mencionó anteriormente.

Cálculo de la nueva información de rango dinámico 607

Una función o dispositivo («cálculo de la nueva información de rango dinámico») 607 calcula la nueva información de rango dinámico. La medida del verdadero nivel del diálogo 505 (figura 5) se redondea y se convierte en el nuevo valor DIALNORM. Se ha observado que un redondeo al 1 dB más cercano resulta útil, aunque esto no es crítico. Usando el índice de perfil 606, el audio descodificado 503 (al que no se le ha aplicado el antiguo DIALNORM y la compresión de rango dinámico) y el nuevo valor DIALNORM (redondeado en 505), se calcula un nuevo conjunto de valores COMPR y DYNRNG 608.

Reempaque de flujo de bits 609

Un reempacador de flujo de bits o función de reempaque («Reempaque de flujo de bits») 609 recibe el flujo de bits AC-3 no descodificado 501, los valores COMPR y DYNRNG 608 y el nivel de diálogo medido 505. Al igual que antes, la medida del verdadero nivel del diálogo 505 se redondea, por ejemplo, al 1 dB más cercano, aunque esto no es crítico, y se convierte en el nuevo valor DIALNORM. El nuevo valor DIALNORM y los nuevos valores COMPR y DYNRNG se actualizan en el flujo de bits AC-3 no descodificado 501. Además, si existieran suficientes bits de datos no utilizados, como se determina en el reempaque de flujo de bits 609, entonces los datos de verificación DIALNORM se usan para reemplazar algunos o la totalidad de los bits de datos no usados. El nuevo flujo de bits AC-3 actualizado se envía como salida como un nuevo flujo de bits 610.

Los detalles del reempaque de flujo de bits 609 se exponen en el ejemplo de la figura 7, que se puede representar de la siguiente manera.

15 *Determinación del espacio disponible 701*

Una función o dispositivo («Determinación del espacio disponible») 701 identifica todos los bits de datos no usados que se pueden usar para actualizar los valores COMPR y DYNRNG y para incluir los nuevos datos de verificación DIALNORM. Cada uno de ambos de los parámetros COMPR y DYNRNG requiere 8 bits en el flujo de bits AC-3. Cada instancia de estos parámetros tiene una marca condicional de «existe». El parámetro COMPR tiene una marca COMPR que, si está ajustada en 1, indica que en el flujo de bits sigue un parámetro COMPR. De manera similar, cada parámetro DYNRNG tiene una marca DYNRNGE que, si está ajustada en 1, indica que en el flujo de bits sigue un parámetro DYNRNG. Si la marca DYNRNGE del primer bloque de una trama está ajustado en 0, entonces el descodificador supone un valor DYNRNG inicial de 0 dB. Si la marca DYNRNGE de cualquiera de los bloques 1 a 5 de una trama está ajustado en 0, entonces el descodificador reutiliza el valor DYNRNG del bloque previo. Esta presencia condicional de los parámetros COMPR y DYNRNG en cada trama significa que el número total de bits usados por COMPR y DYNRNG podría variar.

Debido a que el número total de bits requeridos para los nuevos valores COMPR y DYNRNG podría ser más grande que el número total de bits usados por los valores COMPR y DYNRNG (debido a que la existencia y los valores de COMPR y DYNRNG dependen del valor de DIALNORM), es necesario determinar si existen algunos bits no usados en la trama. Estos bits no usados se pueden usar para los nuevos datos de verificación DIALNORM, así como también para mover la información dentro de la trama AC-3 para hacer espacio para los bits adicionales requeridos para los nuevos valores COMPR y DYNRNG.

35 *Reducción de la información de compresión de rango dinámico 703*

Una función o dispositivo («Reducción D.R.C. de la Información») 703 recibe la identificación de los bits de datos no usados 702 y calcula la nueva información de rango dinámico 608 con el fin de disminuir el número de bits requeridos por los nuevos valores COMPR y DYNRNG si el número total de bits para estos valores es mayor que la suma de los bits no usados más el número total de bits usados por los valores anteriores COMPR y DYNRNG. La salida de la función o dispositivo 703 son los nuevos valores COMPR y DYNRNG, puesto que se podrían haber ajustado a la vista de tales requerimientos de bits.

45 Existe una restricción para cada trama AC-3. La figura 9a muestra dos límites de trama, el límite de trama de $5/8^{\circ}$ y el límite entre el bloque de audio 1 y bloque de audio 2 (AB1-AB2). La restricción es que, cuando una trama está codificada, el límite AB1-AB2 no puede estar más dentro del flujo de bits que el límite de trama de $5/8^{\circ}$. Si el número de bits requeridos para los nuevos valores DYNRNG en los bloques de audio 0 y 1 es mayor que la suma de los bits de datos no usados omitidos en los bloques de audio 0 y 1 más el número de bits usados por los valores DYNRNG antiguos en los bloques de audio 0 y 1, entonces después de hacer espacio para los bits adicionales se podría empujar el límite AB1-AB2 más allá del límite de trama de $5/8^{\circ}$. Si esto sucede, entonces es necesario reducir el número de bits requerido para los nuevos valores DYNRNG en los bloques 0 y 1. Esto se puede realizar de diversas maneras.

55 Un método adecuado consiste en analizar en primer lugar los nuevos valores DYNRNG y las marcas DYNRNGE para los bloques de audio 0 y 1. Si solo la nueva marca DYNRNGE en el bloque 0 está ajustada en 1, entonces esta marca se ajusta en 0 y el nuevo valor DYNRNG del bloque 0 y del bloque 1 se ajusta igual a cero. Si solo la nueva marca DYNRNGE en el bloque 1 está ajustada en 1, entonces la marca se ajusta en 0 y el nuevo valor DYNRNG del bloque 1 se ajusta igual que el del bloque 0. Si las nuevas marcas DYNRNGE de los bloques 0 y 1 están ajustadas en 1, entonces se realizan dos comparaciones. Si la diferencia absoluta entre el nuevo valor del DYNRNG para el bloque 0 y 0 dB es menor que la diferencia absoluta entre los nuevos valores del DYNRNG para los bloques 0 y 1, entonces la nueva marca DYNRNGE para el bloque 0 se ajusta en 0 y el nuevo valor del DYNRNG para el bloque 0 se ajusta en 0. De lo contrario, la nueva marca DYNRNGE para el bloque 1 se ajusta en 0 y los nuevos valores DYNRNG para los bloques 0 y 1 se ajustan en el valor mínimo del DYNRNG de los bloques 0 y 1. Esto reduce el número de bits requeridos para las nuevas palabras DYNRNG en 8 bits. Si una de las marcas DYNRNGE del bloque 0 o 1 se ajusta en 1 y se requiere una reducción de bits adicional, entonces se repetiría el proceso anterior. Una vez

que se completa cualquier reducción de bits, el nuevo valor DYNRNG para el bloque 1 se compara con el nuevo valor DYNRNG para el bloque 2. Si estos valores son iguales, la nueva marca DYNRNGE para el bloque 2 se ajusta en 0. Si los nuevos valores DYNRNG no son iguales, la nueva marca DYNRNGE para el bloque 2 se ajusta en 1.

5 Considerando la trama completa de seis bloques: si el número total de bits requeridos para los nuevos valores COMPR y DYNRNG es mayor que la suma de los bits no usados más el número total de bits usados por los valores COMPR y DYNRNG antiguos, entonces es necesario reducir el número de bits requeridos por los nuevos parámetros. Esto se puede realizar de diversas maneras.

10 Un método adecuado consiste en buscar en los nuevos valores DYNRNG y las marcas DYNRNGE a través de los seis bloques de audio en una trama y agrupar los bloques en regiones en donde cada región representa un bloque con una marca DYNRNGE ajustada en 1, o el primer bloque si la marca DYNRNGE del bloque 0 está ajustado en 0, más cualquiera de los siguientes bloques con las marcas DYNRNGE ajustadas en 0. En consecuencia, el número de regiones podría ser tan bajo como 1, en donde o bien todos los bloques no tiene la marca DYNRNGE ajustada en 1 o bien el primer bloque solo tiene una marca ajustada en 1, o bien el número de regiones podría ser tan alto como
15 seis, en donde cada bloque tiene una marca DYNRNGE ajustada en 1. El valor del DYNRNG para cada región se compara con el valor de DYNRNG en cada región adyacente. Seguidamente, el par de regiones adyacente con los valores de DYNRNG más cercanos se combina en una región ajustando, en primer lugar, los valores DYNRNG en ambas regiones al valor mínimo de cualquier región y, en segundo lugar, ajustando la marca DYNRNGE de la
20 segunda región en 0. Esto reduce en 8 bits el número total de bits requeridos por la nueva información COMPR y DYNRNG. Este proceso se repite hasta que el número total de bits requeridos por los nuevos valores COMPR y DYNRNG es menor o igual a la suma de los bits no usados más el número total de bits requeridos por los valores COMPR y DYNRNG antiguos.

25 Como se indicó anteriormente, es posible que todos los bits de datos no usados en una trama AC-3 sean usados por los parámetros DYNRNG y COMPR actualizados, con lo cual no se dejan bits no usados para los datos de verificación DIALNORM. Como se comentó anteriormente, esto es algo esperado y no reduce la utilidad de la introducción de los datos de verificación DIALNORM en las tramas en las que existan suficientes bits de datos no usados.

30 *Actualización de DIALNORM, información de compresión de rango dinámico y de verificación DIALNORM 705*

Un dispositivo o función («Actualización DIALNORM D.R.C e información de verificación DIALNORM») 705 recibe el flujo de bits AC-3 no descodificado 501. Los nuevos valores COMPR y DYNRNG 704, que se podrían haber
35 ajustado a la vista de los requerimientos de bits, y el nivel de diálogo medido 505, y actualiza el parámetro DIALNORM de flujo de bits, los parámetros de rango dinámico e introduce la información de verificación DIALNORM en el flujo de bits.

Debido a que una trama siempre tiene un parámetro DIALNORM, el nuevo valor DIALNORM se puede escribir en su
40 ubicación predeterminada en el BSI. Sin embargo, la actualización de los parámetros COMPR y DYNRNG implica posiblemente el movimiento de partes de la trama AC-3 a otros lugares con el fin de dejar espacio para los nuevos valores. Si el número total de bits requerido para los nuevos valores COMPR y DYNRNG es más grande que el número total de bits usado por los valores anteriores, es necesario reducir las longitudes de algunos de los campos SKIPD y posiblemente los bits de desperdicio (W). No obstante, si el número total de nuevos bits requerido es
45 menor, entonces se incrementa la longitud de los bits de desperdicio (W). Si existe un número suficiente de datos no usados después de estas actualizaciones de parámetro, entonces los datos de verificación DIALNORM se colocan en las ubicaciones de bits de datos no usados.

Para actualizar el parámetro COMPR, si la marca COMPR antigua está ajustada en 1, el valor anterior COMPR se
50 puede sobrescribir con el nuevo valor COMPR. No obstante, si la marca COMPR anterior está ajustada en 0 y el valor COMPR recién calculado se ajusta en 1, es necesario desplazar 8 bits todos los datos binarios después del parámetro COMPR con el fin de dejar espacio para el nuevo valor COMPR. Seguidamente, la marca COMPR en la trama se ajusta en 1 y el nuevo valor COMPR se escribe en los 8 bits de espacio recién creados. Si la marca COMPR antigua está ajustado en 1 y la nueva marca COMPR se ajusta en 0, entonces el COMPR en la trama se
55 ajusta en 0 y es necesario desplazar 8 bits todos los datos binarios después del parámetro COMPR, debido a que el parámetro COMPR ya no existe en la trama.

Para actualizar los parámetros DYNRNG en cada bloque de audio, si la marca DYNRNGE antigua se ajusta en 1, el
60 valor DYNRNG antiguo se puede sobrescribir con el nuevo valor DYNRNG. No obstante, si la marca DYNRNGE anterior está ajustada en 0 y el nuevo valor DYNRNG se ajusta en 1, es necesario desplazar todos los datos binarios después del parámetro DYNRNG con el fin de dejar espacio para el nuevo valor DYNRNG. Seguidamente, la marca DYNRNGE en la trama se ajusta en 1 y el nuevo valor DYNRNG se puede escribir en los 8 bits de espacio recién creados. Si la marca DYNRNGE antigua está ajustado en 1 y la nueva marca DYNRNGE se ajusta en 0, entonces la
65 marca DYNRNGE en la trama se ajusta en 0 y es necesario desplazar 8 bits todos los datos binarios después del parámetro DYNRNG, debido a que el parámetro DYNRNGE ya no existe en el flujo de bits.

- El parámetro SKIPL indica la longitud del campo SKIPD en bytes. Para reducir la longitud del campo SKIPD, es necesario desplazar un múltiplo de 8 bits los datos binarios a la derecha del campo SKIPD. A continuación, el parámetro SKIPL se actualiza para reflejar la nueva longitud del campo SKIPD. De manera ocasional, un campo SKIPD podría contener información opcional que no está definida de manera oficial en el estándar AC-3 (véase, por ejemplo, el documento A52/A citado anteriormente). Si el primer bit en el campo SKIPD es igual a 1, entonces la información que lleva los datos sigue en el campo SKIPD, de lo contrario todos los bits en el campo SKIPD se ajustan en 0. Si la información está presente y es necesario recortar el campo SKIPD, entonces este solo se puede acortar hasta esta información. Esto permite mantener la información dentro de la trama AC-3.
- Después de la modificación y la actualización de los bits de datos no usados, los datos de verificación DIALNORM se pueden introducir en los bits no usados. Según se comentó anteriormente, estos datos pueden tomar varias formas, que incluyen el duplicado del parámetro DIALNORM de las tramas con suficiente información de datos de sincronización y de identificación. Esto permite que un proceso de descodificador de verificación DIALNORM realice la búsqueda de los bits de datos no usados, identifique si existen los datos de verificación DIALNORM y lo compare con el parámetro estándar DIALNORM incrustado en el flujo de bits AC-3.

Actualización de los CRC 707

- El flujo de bits AC-3 actualizado, que incluye la información de verificación DIALNORM, se aplica a un dispositivo o función de generación de palabra de corrección de error («Actualización de CRC») 707. Cuando han cambiado los datos en una trama AC-3, se deben recalcular las dos palabras de detección de error CRC1 y CRC2. Si solo han cambiado los datos de límite de trama hasta $5/8^{\circ}$, entonces solo es necesario recalcular CRC1. Del mismo modo, si solo han cambiado los datos del límite de trama de $5/8^{\circ}$ hasta el fin de la trama, entonces solo es necesario recalcular CRC2.

Datos de verificación no presentes en el flujo de bits AC-3

Metadatos DIALNORM incorrectos, rango DIALNORM insuficiente para transmitir el nivel

- Según se muestra en el ejemplo de la figura 4, si no existe la información de verificación (salida NO de la etapa 401), y si el valor AC-3 DIALNORM existente es incorrecto (salida NO de la etapa 406), entonces se debe determinar si el nivel de intensidad acústica medido se encuentra dentro del rango válido del DIALNORM (etapa 408). Según se mencionó anteriormente, el parámetro DIALNORM no tiene alcance suficiente para transmitir el nivel cuando el nivel medido del diálogo 505 se encuentra fuera del rango válido del parámetro DIALNORM como se permite en el flujo de bits AC-3. Es decir, el nivel medido es menor que -31 dB o mayor que -1 dB. En este caso, la salida de la etapa 408 es NO. Una forma de corregir esta situación es actualizar el parámetro DIALNORM en la trama con el valor válido más cercano, según se describió anteriormente. Sin embargo, esto podría conducir a algún error entre el valor DIALNORM y el nivel medido del diálogo. Una alternativa adecuada que minimiza este error es la realización de las etapas 411, 412 y 413 de la figura 4, como se describe a continuación haciendo referencia al ejemplo de la figura 6b. Según se indicó anteriormente, la figura 6b es una variación del flujo de bits de actualización 507 de la figura 5 que resulta útil para realizar las etapas 411, 412 y 413. Los elementos de la figura 6b que difieren de aquellos de la figura 6a se pueden representar de la siguiente manera. Los elementos comunes a las figuras 6a y 6b mantienen los mismos números de referencia respectivos.

Ajuste de la ganancia 611

- El audio descodificado 503 se aplica al cambiador ajustable de ganancia o función de cambio de ganancia («Ajuste de la ganancia») 611. Se podría aplicar al audio un cambio de ganancia adecuado con el fin de reducir el error entre el nivel medido de diálogo y el valor DIALNORM (etapa 411). Por ejemplo, si el nivel medido de diálogo es -36 dB, el DIALNORM se podría ajustar en el valor válido más cercano, -31 dB, intensificando de esta manera el audio en 5 dB, de -36 dB a -31 dB.

Codificación AC-3 modificada 629

- A continuación, el audio al que se le ha ajustado la ganancia se codifica nuevamente usando la codificación de audio AC-3 modificada aplicándola, junto con el nuevo DIALNORM y la información de compresión de rango dinámico 608 (etapa 412), a un codificador AC-3 modificado o función de codificación («Codificación AC-3 modificada») 629. La codificación AC-3 modificada se caracteriza como «modificada» debido a que tiene en cuenta la aplicación de las capacidades de los datos de verificación DIALNORM e introduce estos datos en los bits de datos no usados después del proceso de codificación y antes del empaquetado final del flujo de bits. Esta nueva codificación mantiene todos los BSI originales (excepto para DIALNORM, la información de compresión de rango dinámico y la verificación DIALNORM) y los metadatos AUX de la trama original de flujo de bits 501 e incluye el cálculo de las nuevas palabras de detección de error.
- Otras funciones y dispositivos de la figura 6b podrían ser los mismos que los correspondientes funciones y dispositivos de la figura 6a, según se mencionó anteriormente.

La puesta en práctica de las etapas 411, 412 y 413 podría conducir a alguna pérdida de la calidad de sonido debido a la descodificación y nueva codificación del contenido de audio. Por lo tanto, podría ser deseable que solo se vuelva a codificar el contenido si el error absoluto entre el nivel medido de diálogo y el valor DIALNORM más cercano es mayor que un umbral. Un valor umbral de tres (3 dB) podría ser adecuado para muchas aplicaciones.

Subconjuntos de la figura 4

Según se mencionó anteriormente, se podrían usar ya sea la totalidad de los procesos de la figura 4 o subconjuntos del proceso de la figura 4 o dispositivos que utilicen las etapas de los procesos.

Un subconjunto útil y de bajo coste del proceso de la figura 4 consiste en usar las etapas 401 a 403. Si existe información de verificación en el flujo de bits (401 es SÍ), las etapas 402 y 403 funcionan según se describió anteriormente, ya sea para dejar el flujo de bits AC-3 sin cambios o para ajustar el valor DIALNORM en el valor de verificación. Si no existe información de verificación en el flujo de bits (la salida 401 es NO), el DIALNORM se podría dejar inalterado o ajustarse en un valor predeterminado.

Otro subconjunto útil del proceso de la figura 4 consiste en usar las etapas 401 y 404 a 407. Esto resulta útil cuando no existe información de verificación y se desea añadir información de verificación cuando el DIALNORM existente es correcto. Si existe información de verificación en el flujo de bits (401 es SÍ), el flujo de bits se podría dejar inalterado. Si no existe información de verificación (la salida de 401 es NO), las etapas 404, 405 y 406 determinarían si el DIALNORM existente es correcto (la salida de 406 es SÍ) o no (la salida de 406 es NO). Si el DIALNORM existente es correcto, se podría añadir información de verificación al flujo de bits. Si el DIALNORM existente no es correcto, el DIALNORM se podría ajustar en un valor predeterminado.

Otro subconjunto útil, aunque algo más costoso, del proceso o dispositivos de la figura 4 consiste en usar las etapas 401 a 406. La operación es justamente como se describió cuando existe la información de verificación (la salida de 401 es SÍ), aunque cuando no existe información de verificación (la salida de 401 es NO), las etapas 404, 405 y 406 determinarían si el DIALNORM existente es correcto (la salida de 406 es SÍ) o no (la salida de 406 es NO). Si el DIALNORM existente es correcto, el flujo de bits se podría dejar inalterado. Si el DIALNORM existente no es correcto, el DIALNORM se podría ajustar en un valor predeterminado.

Descodificación conforme a los datos de verificación

Otro aspecto de la presente invención consiste en la descodificación adecuada de un flujo de bits AC-3 tenga o no DIALNORM y datos de verificación de coincidencia correctos, aunque usando estos datos de verificación cuando estén presentes. Esto se podría denominar descodificación «conforme a los datos de verificación». Este aspecto de la invención se expone en el diagrama de flujo de ejemplo de la figura 8. Como se explica más adelante, se podría usar ya sea la totalidad del proceso de la figura 8 o subconjuntos del proceso de la figura 8. Tales procesos o dispositivos que usan las etapas de los procesos podrían ser útiles, por ejemplo, en la descodificación de un flujo de bits. Las etapas de la figura 8 que corresponden, de manera general, con las etapas de la figura 4, usan números de referencia correspondientes (p. ej., «801» y «401»). Se entenderá que las etapas de la figura 8 o los subconjuntos de la misma podrían representar partes de uno o más procesos o podrían ser funciones realizadas en uno o más dispositivos.

Prueba de existencia de datos de verificación DIALNORM (Etapa 801)

Según se muestra en la figura 8, la primera etapa realizada (etapa 801) es la determinación de si existen datos de verificación AC-3 DIALNORM en el flujo de bits AC-3. Esta etapa se podría realizar de la misma manera que la etapa 401 de la figura 4, descrita anteriormente (incluyendo los detalles de la misma mostrados en la figura 3).

Prueba de si los datos de verificación DIALNORM coinciden con los datos AC-3 DIALNORM (802)

Según se muestra en la etapa 801 de la figura 8, cuando existan datos de verificación DIALNORM (salida SÍ de la etapa 801), la siguiente etapa, etapa 802, consiste en determinar si los datos de verificación coinciden con el valor AC-3 DIALNORM. Esta etapa se podría realizar de la misma manera que la etapa 402 de la figura 4, descrita anteriormente. Si los valores coinciden, entonces el valor DIALNORM normal es correcto y el flujo de bits AC-3 aplicado al proceso (entrada de la etapa 801) se podría descodificar usando su DIALNORM existente y los metadatos de rango dinámico relacionados según se indica en la etapa 814, proporcionando de esta manera un flujo de bits de audio AC-3 descodificado. La «coincidencia» o no de los valores se puede determinar sabiendo si están dentro de un umbral, como se explicó anteriormente en relación con la descripción de la etapa 402.

Corrección de datos AC-3 DIALNORM normales con datos de verificación DIALNORM (803)

Cuando los datos de verificación DIALNORM se extraen del flujo de bits AC-3 y no coinciden con los metadatos AC-3 DIALNORM normales (salida NO de la etapa 802), entonces los metadatos DIALNORM se actualizan con la

verificación del valor DIALNORM (803). Debido a que se ha determinado que el valor DIALNORM normal es incorrecto y es necesario actualizarlo, es posible que los metadatos de compresión de rango dinámico AC-3 relacionados también sean incorrectos. Por lo tanto, es necesario analizar la información de compresión de rango dinámico y, si es correcta, solamente se actualiza el parámetro de metadatos DIALNORM. Si la información de compresión de rango dinámico es incorrecta, entonces también es necesario actualizarla. Los detalles de este análisis y actualización se explican en este documento en conexión con las etapas 409 a 413 de las figuras 4 y 6b).

Datos de verificación no presentes en el flujo de bits AC-3

10 Metadatos DIALNORM correctos

Según se muestra en el ejemplo de la figura 8, si los datos de verificación DIALNORM no están contenidos dentro del flujo de bits AC-3 (salida NO en la etapa 801), entonces el flujo de bits AC-3 se puede descodificar en audio (p. ej., en audio codificado PCM) (etapa 804) sin aplicar al audio el parámetro DIALNORM y los parámetros de control de rango dinámico relacionados (debido a que estos parámetros de metadatos podrían ser incorrectos), de manera que el contenido de audio descodificado se encuentra al mismo nivel que la entrada al codificador que se usó para crear el flujo de bits). A continuación, se mide el nivel DIALNORM del audio descodificado (etapa 805). Tal medida del DIALNORM podría ser la misma que se realizó a través del nivel de medida de diálogo 104 descrito anteriormente. Después de la medida del nivel de diálogo en la etapa 405, el valor medido se compara, en la etapa 806, con el valor de metadatos AC-3 DIALNORM del flujo de bits AC-3 de entrada. Como se muestra después de la salida SÍ de la etapa 806, si ese valor DIALNORM es correcto (salida SÍ de la etapa 806), el valor DIALNORM original y la información de compresión de rango dinámico original relacionada del flujo de bits AC-3 de entrada se aplicarían al audio descodificado producido por el descodificador AC-3 de la etapa 804 con el fin de proporcionar un flujo de bits de audio AC-3 descodificado al cual se han aplicado los valores DIALNORM y de parámetro de compresión de rango dinámico correctos.

Datos de verificación no presentes en el flujo de bits AC-3

30 Metadatos DIALNORM incorrectos

Quando los metadatos DIALNORM existentes son incorrectos (salida NO de la etapa 806), es necesario ajustar el valor DIALNORM al valor DIALNORM medido de la etapa 805 y determinar la nueva información de compresión de rango dinámico de este valor de parámetro DIALNORM medido. Esto se podría conseguir en la etapa 815, etapa que podría ser la misma que la etapa 412. Seguidamente, el valor DIALNORM medido y la información de compresión de rango dinámico determinados por la etapa 815 se podrían aplicar, en la etapa 816, al audio digital o analógico descodificado proporcionado por la etapa 804.

Subconjuntos de la figura 8

40 Según se mencionó anteriormente, se podrían usar ya sea la totalidad de los procesos de la figura 8 o subconjuntos del proceso de la figura 8 o dispositivos que utilicen las etapas de los procesos.

45 Un subconjunto útil y de bajo coste del proceso de la figura 8 consiste en usar las etapas 801 a 803 y 814. Si existe información de verificación en el flujo de bits (801 es SÍ), las etapas 802, 803 y 814 funcionan según se describió anteriormente para descodificar el flujo de bits AC-3. Si no existe información de verificación en el flujo de bits (la salida 801 es NO), el flujo de bits se podría descodificar usando su valor DIALNORM existente y los valores del parámetro de compresión de rango dinámico relacionados o mediante el uso de un valor DIALNORM predeterminado y los valores del parámetro de compresión de rango dinámico relacionados.

50 Otro subconjunto útil, aunque algo más costoso, del proceso o dispositivos de la figura 8 consiste en usar todas las etapas excepto la 815. Esto evita el cálculo requerido para determinar la información de compresión de rango dinámico relacionada con el DIALNORM medido. La operación es justamente como se describió en conexión con la figura 8, excepto que cuando la etapa 806 determina que el DIALNORM existente no es correcto, el flujo de bits AC-3 de entrada se podría descodificar ajustando el DIALNORM y el valor del parámetro de rango dinámico relacionado en un valor predeterminado.

Antecedentes adicionales

Asignación de bits AC-3 y bits de datos no usados

60 Un diagrama de bloques simplificado del codificador AC-3 se muestra en la figura 9e. Las muestras de audio PCM se introducen en la función de transformada del dominio de la frecuencia 902. Una transformada discreta del coseno modificada de 512 puntos (MDCT) con una superposición del 50 % se usa para presentar los datos de entrada con el fin de evitar los artefactos de límite del procesamiento por bloques. En el caso de señales transitorias, se consigue un rendimiento temporal mejorado (ruido previo transitorio reducido) usando una técnica de cambio de bloque en la cual se calculan dos transformadas de 256 puntos en lugar de la transformada de 512 puntos. Los coeficientes de

transformada de la función 902 se aplican en un proceso de punto flotante de bloque 904 que segmenta cada coeficiente de transformada en pares de exponente y mantisa. Las mantisas del coeficiente de transformada son cuantificadas en la función de cuantificación de mantisa 906 con un número variable de bits asignado a través de la función de asignación de bits 908 que opera sobre la base de un modelo paramétrico de asignación de bits en respuesta a los exponentes de punto flotante de bloque.

El modelo de asignación de bits AC-3 usa los principios del enmascaramiento psicoacústico para seleccionar el número de bits asignados a cada mantisa en una banda de frecuencia determinada. Dependiendo del grado de enmascaramiento, algunas mantisas podrían recibir muy pocos bits o incluso ningún bit en absoluto. Esto reduce el número de bits requeridos para representar la fuente de audio, a expensas de un mayor ruido (aunque inaudible).

A diferencia de algunos otros sistemas de codificación, el AC-3 no pasa los resultados de asignación de bits al decodificador en el flujo de bits de audio comprimido. En su lugar, se adopta un proceso paramétrico en el cual el codificador de audio construye su modelo de enmascaramiento sobre la base de los exponentes del coeficiente de transformada y algunos parámetros clave que dependen de la señal. Estos parámetros se pasan de la función de asignación de bits 908 a la función de empaquetado de flujo de bits 910 para pasar al decodificador por medio del flujo de bits, requiriendo menos bits de los que serían necesarios para transmitir los valores de asignación de bits sin procesar. La función de empaquetado de flujo de bits 910 que genera el flujo codificado de bits de audio también recibe los exponentes y las mantisas cuantificadas para su inclusión en el flujo de bits. En el decodificador, la asignación de bits se reconstruye sobre la base de los exponentes recibidos y los parámetros de asignación de bits. Esta disposición constituye una asignación de bits adaptativa hacia atrás/hacia adelante híbrida.

La eficiencia de la codificación de AC-3 mejora a medida que se incrementa el número de canales de origen debido a dos características principales: un examen global de bits y un acoplamiento de alta frecuencia. La técnica de examen global de bits permite que el asignador de bits asigne los bits disponibles entre los canales de audio según sea necesario. Si uno o más canales estuvieran inactivos en un instante de tiempo específico, los canales restantes recibirían más bits de los necesarios.

En el sistema de compresión de audio AC-3, el proceso de asignación de bits usa una búsqueda finita. En cada iteración de la búsqueda, se varía el parámetro de señal a ruido (SNR) para controlar la asignación de los bits. Esto también afecta los valores de otros parámetros. Al final de la búsqueda, si el número de bits usados excede del número de bits asignados, se usa la última asignación legal. A menudo, esta asignación no es capaz de usar todos los bits disponibles, con lo que se dejan bits no usados o desperdiciados.

Como se comentó anteriormente, un flujo codificado de bits de audio en serie AC-3 está compuesto por una secuencia de tramas construida como se muestra, de manera general, en la figura 9a. Cada trama AC-3 representa un rango constante de tiempo de 1536 muestras PCM a través de todos los canales codificados y contiene seis bloques codificados de audio (AB0 a AB5), cada uno de los cuales representa 256 nuevas muestras de audio. Cada trama AC-3 tiene un tamaño fijo (uno de varios tamaños en números de bits comprendidos en el rango de entre 64 y 1920 bits) que depende de la velocidad de muestreo PCM (32, 44,1 o 48 kHz) y la velocidad de bits de audio codificados (con valores discretos comprendidos en el rango de entre 32 y 640 kbps). El encabezamiento de información de sincronización (SI) en el comienzo de cada trama contiene la información necesaria para adquirir y mantener la sincronización. El encabezamiento de información de flujo de bits (BSI) sigue el campo SI y contiene los parámetros que describen el servicio de audio codificado. Los campos SI y BSI describen la configuración de flujo de bits, incluyendo la velocidad de muestreo, la velocidad de datos, el número de canales de audio codificado y otros diversos elementos de nivel de sistemas. Después de los bloques de audio codificados (AB0 a AB5) se encuentra un campo de datos auxiliares (AUX). Al final de cada trama se encuentra una casilla de verificación de error que incluye una palabra CRC (palabra de código de corrección de redundancia cíclica) para la detección del error. Además, otra palabra CRC está situada en el encabezamiento SI.

Aunque el ancho de los elementos del flujo de bits de la figura 9a sugiere, de manera general, un número típico de bits en cada elemento, la figura no se encuentran a escala. El número de bits asignados y usados en los bloques de audio y en el campo AUX es variable. El bloque AB0 se muestra más ancho que los demás bloques debido a que cada trama es esencialmente independiente de otras tramas y los bloques AB1 a AB5 podrían compartir la información transportada por el bloque AB0 sin repetir la información, permitiendo que los bloques AB1 a AB5 transporten una menor cantidad de bits que el bloque AB0. Además de la posibilidad de compartir, los bloques de audio también tienen una longitud variable debido al número variable de bits que se pueden asignar a los datos de mantisa cuantificados en cada bloque.

Como se explica en la patente de EE.UU. 6.807.528 citada anteriormente, existen bits no usados en una trama cada vez que la función de asignación de bits del codificador no utilice todos los bits disponibles para la codificación de la señal de audio. Esto sucede si la asignación final de bits no llega a usar todos los bits disponibles o si el audio de entrada no requiere todos los bits disponibles. Debido a que estos bits no usados se tienen que colocar en algún lugar de la trama con el fin de que la trama tenga un tamaño fijo obligatorio, el codificador inserta bits ficticios o nulos en el flujo de bits con el fin de llenar la longitud de la trama. Estos bits nulos se introducen en un «campo omitido» en uno o más de los bloques de audio (como se muestra en la figura 9d) así como también en el campo AUX. Cada

campo omitido acepta bits nulos en bytes de 8 bits, mientras que el campo AUX acepta hasta siete bits nulos con el fin de proporcionar un «ajuste fino» de la longitud de trama y garantizar que la palabra final CRC esté presente en los últimos 16 bits de la trama. En la práctica, los bits nulos son bits aleatorios. Estos bits nulos son bits desperdiciados que no llevan información útil. Un aspecto de la presente invención consiste en el uso de los valores de algunos o de la totalidad de tales bits nulos para transportar los bits que llevan información relacionados con algunos de los parámetros AC-3 contenidos dentro del flujo de bits (de manera particular, el parámetro DIALNORM que se muestra en la figura 9c).

El decodificador omite o ignora los bits nulos de los campos omitidos y del campo AUX. Aunque el decodificador AC-3 es capaz de identificar los bit nulos e ignorarlos, a priori se desconoce el número de bits nulos y su ubicación en el flujo de bits (su número y ubicación varía de trama a trama, es decir, los campos omitidos son de tamaño variable y varían sus posiciones de inicio en los bloques AB1 a AB5 y, de manera similar, el campo AUX es de tamaño variable y varía su posición de inicio) y no es posible discernir su número y ubicación mediante una mera inspección del flujo de bits AC-3 (los bits nulos son aleatorios y no se pueden distinguir de otros datos del flujo de bits).

Cada bloque de audio (AB0 a AB5) comienza con los «datos fijos» constituidos de los elementos de flujo de bits cuyos tamaños de palabra (longitudes de bit) se conocen a priori (es decir, estos elementos de datos fijos tienen un número de bits previamente asignado y no son bits asignados por la asignación de bits). Los datos fijos son una colección de parámetros y marcas que incluyen marcas de cambio de bloque, información de acoplamiento, exponentes y parámetros de asignación de bits. Después de los datos fijos se encuentran los datos de «campo omitido», que tienen un tamaño mínimo de 1 bit, si el campo omitido no contiene bits nulos, y un tamaño máximo de 522 bits si contiene bits nulos. Una palabra de un bit, el contenido mínimo de un campo omitido, indica si el campo omitido incluye bits nulos. Si es así, a continuación una palabra de 9 bits indica el número de bytes de bits nulos. A continuación van los bytes nulos. Después del salto se encuentran los datos de mantisa. El tamaño de los datos de mantisa es variable y viene determinado por la asignación de bits.

El hecho de que un bloque particular de audio contenga un campo omitido con bits nulos se determina a través de las siguientes reglas: 1) el tamaño combinado de los campos SYNCINFO (a saber, el SYNCWORD, la primera palabra CRC, la palabra de código de frecuencia de muestreo y la palabra de código de tamaño de trama), los campos BSI, el bloque de audio 0 (AB0) y el bloque de audio 1 (AB1) nunca excede de 5/8 de la trama, y 2) el tamaño combinado de los datos de mantisa de bloque de audio 5 (AB5), el campo de datos AUX y el campo de verificación de error nunca excede los 3/8 finales de la trama. La configuración de 5/8 y 3/8 se usa para reducir el tiempo de espera (la primera palabra CRC se aplica en los primeros 5/8 de la trama, lo que permite una decodificación más rápida). En principio, no siendo para la configuración de 5/8 y 3/8, todos los bits nulos se podrían introducir en el campo AUX sin la necesidad de uno o más campos omitidos.

El campo de datos AUX tiene dos funciones. Una función del campo de datos AUX, mencionada anteriormente, consiste en proporcionar un ajuste fino de la longitud de trama y garantizar que los últimos 16 bits del Industrial se usen para la segunda palabra CRC. Hasta 7 bits nulos se introducen en el campo AUX. Una segunda función del campo AUX, que es opcional e independiente de la primera función, consiste en transportar la información adicional («dato auxiliar») a expensas de usar los bits que de otra manera se podrían asignar a las mantisas en los bloques de audio. El último bit del campo de datos AUX indica si existe cualquier dato auxiliar opcional. Si el bit indica que existe, la palabra precedente de 14 bits indica la longitud de los datos auxiliares y en los siguientes bits precedentes están los datos auxiliares. Los bits nulos, si existen, a su vez preceden a los datos auxiliares en el campo AUX. Si el campo AUX no tiene datos auxiliares, los bits nulos, si existen, preceden al bit único al final del campo de datos AUX que indica si existen datos auxiliares. Por lo tanto, existan o no datos auxiliares, podrían existir o no bits nulos en el campo AUX. No existen bits nulos en el campo AUX si no existen bits no usados (es posible que los bits no usados existan en una trama determinada, aunque la probabilidad de que esto ocurra en muchas tramas consecutivas sería extremadamente baja) o si el número de bits nulos es divisible entre ocho y, por lo tanto, todos los bits nulos se transportan en uno o más de los campos omitidos.

En la disposición de codificación AC-3 estándar, los bits nulos en el campo AUX y/o el campo AUX y uno o más campos omitidos, son bits no usados o desperdiciados (es decir, no llevan información útil). De acuerdo con los aspectos de la presente invención, según se comentó anteriormente, algunos o la totalidad de tales bits no usados se reemplazan por bits de verificación de metadatos que llevan información, al tiempo que se preserva la compatibilidad total con los codificadores y decodificadores AC-3 existentes y se evita cualquier degradación de las señales de audio codificadas.

De manera preferente, los nuevos bits que llevan información se conforman en un formato o sintaxis conocido o predeterminado, de manera que se pueden recuperar a través de un proceso de decodificación de verificación de parámetro de metadatos (por ejemplo, DIALNORM). El reemplazo de los bits desperdiciados por bits de verificación de metadatos (DIALNORM) se puede conseguir una vez que cualquier codificador válido AC-3 cree un flujo de bits AC-3. Por ejemplo, un codificador convencional AC-3 no modificado se podría usar para generar el flujo de bits AC-3 estándar. El flujo de bits resultante se analiza con el fin de identificar las ubicaciones de algunos o de la totalidad de los bits no usados en cada trama. Seguidamente, algunos o la totalidad de los bits no usados identificados se

reemplazan por bits que llevan información (bits de datos de verificación DIALNORM) que se encuentran incrustados en ubicaciones anteriormente ocupadas por bits no usados. Debido a que algunos de los datos se cambian (se cambian algunos o la totalidad de los bits nulos), se recalcula la suma de verificación para la trama completa, y la segunda palabra CRC, que se aplica a la trama completa, se reemplaza por una nueva palabra CRC, y, si se cambian los datos en los primeros 3/8 de la trama, la suma de verificación para esta parte de la trama se recalcula y la primera palabra CRC, que se aplica a los primeros 3/8 de la trama, también se reemplaza por una nueva palabra CRC.

De manera alternativa, en lugar de reemplazar algunos o la totalidad de los bits no usados en un flujo de bits AC-3 por bits que llevan información después de una codificación estándar, un codificador AC-3 modificado que incluye capacidades adicionales de verificación y análisis de metadatos podría introducir bits que llevan información en algunas o en la totalidad de las posiciones de bits no usados de una trama en lugar de bits nulos aleatorios durante el proceso de codificación.

Si el flujo de bits AC-3 se modifica antes o después del proceso de codificación, el flujo de bits modificado resultante parece el mismo en un descodificador convencional AC-3. Un descodificador AC-3 no modificado que recibe el flujo de bits modificado ignora los bits que llevan información de la misma manera que ignora u omite los bits nulos en las mismas ubicaciones de bits. Los bits que llevan información que reemplazan los bits no usados se pueden recuperar, ya sea en un descodificador AC-3 modificado o en un proceso especial de análisis de metadatos AC-3 que identifica las ubicaciones de los bits no usados en una trama, detecta los datos en las ubicaciones de bits no usados y presenta los resultados del análisis de verificación de metadatos realizado en el flujo de bits AC-3. En cualquier caso, la recuperación y el análisis de los datos de verificación que reemplazan los bits no usados en un flujo de bits AC-3 no altera el resto del flujo de bits. Por lo tanto, los aspectos de la presente invención pueden preservar la calidad de audio de dos maneras: no se usan los bits que de lo contrario se usarían para audio y se puede evitar la necesidad de descodificar y codificar nuevamente el flujo de bits (aunque esto podría ser necesario y útil según se describió anteriormente).

Nivel de diálogo AC-3 y parámetros de metadatos de compresión

Según se mencionó anteriormente, en los metadatos de trama AC-3 se encuentra incluido un parámetro que indica el nivel de intensidad acústica de la conversación o diálogo contenido en el audio comprimido. Este parámetro se denomina DIALNORM y el propósito de este parámetro es que antes de codificar un elemento de audio o de comprimir sus datos, se mide el nivel predominante del diálogo o conversación del elemento. Entonces, esta medida se usa para ajustar el parámetro DIALNORM en cada trama del flujo de bits que contiene el elemento de audio comprimido. Durante la reproducción del flujo de bits, el descodificador AC-3 usa el parámetro DIALNORM para modificar el nivel de reproducción o la intensidad acústica del elemento, de manera que la intensidad acústica percibida del diálogo se encuentre en un nivel uniforme.

La figura 10a muestra un ejemplo que contiene tres elementos de audio diferentes. El nivel digital es el nivel del contenido de audio de datos comprimidos con relación a una onda sinusoidal digital de escala completa (0 dB FS). Se muestra el nivel máximo y mínimo para cada elemento junto con el nivel de diálogo predominante. El parámetro DIALNORM para cada elemento es el nivel de diálogo, redondeado a unidades de 1 dB. La figura 10b muestra, durante la reproducción, cómo el descodificador escala el nivel de cada elemento, de manera que el nivel o intensidad acústica del diálogo para cada elemento sea el mismo, o muy similar. Para el sistema AC-3, el nivel de referencia al cual se escala el diálogo de cada elemento es de -31 dB FS. A continuación, este nivel digital de referencia se puede calibrar en un sistema de reproducción hasta el nivel de presión sonora deseado.

El uso del parámetro DIALNORM en AC-3 proporciona a los oyentes una experiencia de escucha más uniforme y predecible, al reducir las importantes diferencias de intensidad acústica que existen entre los diferentes elementos de audio creados por distintas personas en diferentes entornos de escucha y para distintos propósitos. Sin embargo, el parámetro DIALNORM podría ser incorrecto por todas las razones comentadas anteriormente.

Compresión de rango dinámico

También se encuentran incluidos en los metadatos de trama AC-3 los parámetros que, si se aplicasen en el audio durante la reproducción, servirían para reducir el rango dinámico del contenido del audio. Es decir, hacen más silenciosas las partes más intensas del audio y más intensas las partes silenciosas del audio. Estos parámetros de compresión de rango dinámico se denominan COMPR y DYNRNG y se calculan de manera automática durante el proceso de codificación de un flujo de bits AC-3. Véase la figura 9.

La capacidad para reducir el rango dinámico del audio resulta útil en una diversidad de situaciones. Por ejemplo, cuando se ve una película por la noche, a menudo es necesario escucharla a un volumen de reproducción reducido, de manera que no perturbe el sueño de los miembros de la familia o de los ocupantes de los edificios adyacentes. Debido a que las películas tienden a tener un rango dinámico muy grande, el volumen de reproducción reducido origina que la película sea demasiado silenciosa para ser audible. El uso de la compresión de rango dinámico ayuda a incrementar las partes silenciosas, haciéndolas audibles, y reduce las partes de mayor intensidad acústica

haciéndolas menos molestas.

Los parámetros de compresión de rango dinámico se calculan con referencia al nivel de diálogo, según se indica mediante el parámetro DIALNORM. Esto garantiza que el nivel de diálogo promedio no se altere y que solo se alteren las partes más fuertes o más suaves del elemento de audio.

La figura 12 muestra un ejemplo que contiene tres elementos de audio diferentes. La figura 12a muestra el nivel de diálogo promedio y el rango dinámico de los elementos de audio no procesados. La figura 12b muestra cómo, durante la reproducción, la aplicación de la compresión de rango dinámico y el parámetro DIALNORM originan un nivel promedio de diálogo uniforme y una señal de salida de rango dinámico reducida en los tres elementos.

Debido a que los parámetros de compresión de rango dinámico se calculan con relación al nivel de diálogo, su uso depende de que los creadores de contenido midan y ajusten el parámetro DIALNORM de manera correcta. Si existe un error entre el nivel de diálogo indicado por el parámetro DIALNORM y el verdadero nivel de diálogo en el contenido de audio, entonces es probable que el diálogo presente cambios no deseables y audibles de la ganancia dinámica debido a la compresión.

DIALNORM2, COMPR2 y DYNRNG2

En la mayoría de las circunstancias, el sistema AC-3 usa un nivel único de diálogo y un único conjunto de parámetros de información de rango dinámico para todos los canales. Sin embargo, el AC-3 incluye un modo de funcionamiento que permite que dos canales funcionen de manera independiente; es decir, cada canal tiene un nivel de diálogo independiente, así como la información de rango dinámico. De esta manera, el segundo de los dos canales usa los parámetros DIALNORM2, COMPR2 y DYNRNG2. (Véase la figura 9.) Debido a que los parámetros DIALNORM2, COMPR2 y DYNRNG2 se interpretan y usan exactamente de la misma manera que los parámetros DIALNORM, COMPR2 y DYNRNG, en este documento solo se describe el funcionamiento del último.

Implementación

La invención se podría implementar en hardware o software, o una combinación de ambos (por ejemplo, matrices lógicas programables). A menos que se especifique de otra manera, los algoritmos o procesos incluidos como parte de la invención no están inherentemente relacionados con ningún ordenador particular u otro aparato. En particular, diversas máquinas de uso general se podrían usar con programas escritos de acuerdo con las explicaciones de este documento, o podría ser más cómoda la construcción de aparatos más especializados (p. ej., circuitos integrados) para realizar las etapas requeridas del método. Por lo tanto, la invención se podría implementar en uno o más programas informáticos que se ejecuten en uno o más sistemas informáticos programables, cada uno de los cuales comprende al menos un procesador, al menos un sistema de almacenamiento de datos (que incluyen una memoria volátil y una memoria no volátil y/o elementos de almacenamiento), al menos un dispositivo o puerto de entrada y por lo menos un dispositivo o puerto de salida. El código de programa se aplica en los datos de entrada para realizar las funciones descritas en este documento y generar la información de salida. La información de salida se aplica a uno o más dispositivos de salida, de una manera conocida.

Cada programa se podría implementar en cualquier lenguaje informático deseado (que incluye lenguajes de programación de máquina, de ensamblaje, de proceso de alto nivel, lógicos o lenguajes orientados a objetos) para comunicarse con un sistema informático. En cualquier caso, el lenguaje podría ser un lenguaje compilado o interpretado.

Se apreciará que algunas etapas o funciones mostradas en las figuras de ejemplo realizan múltiples subetapas y también se podrían mostrar como múltiples etapas o funciones en lugar de una etapa o función. Se apreciará que diversos dispositivos, funciones, etapas y procesos mostrados y descritos en varios ejemplos en este documento se podrían mostrar de manera combinada o separada de maneras diferentes a las mostradas en las distintas figuras. Por ejemplo, cuando sean implementadas por secuencias de instrucción de software informático, diversas funciones y etapas de las figuras de ejemplo se podrían implementar mediante las secuencias de múltiple lectura de instrucción de software que se ejecutan en el hardware de procesamiento de señal digital, en tal caso, los distintos dispositivos y funciones de los ejemplos mostrados en las figuras podrían corresponder a partes de las instrucciones de software.

De manera preferente, cada programa informático se almacena o descarga en medios o dispositivos de almacenamiento (p. ej., una memoria o soportes de estado sólido, o soportes magnéticos u ópticos) que pueden ser leídos por un ordenador programable de uso general o especial, para la configuración y operación del ordenador cuando los soportes o dispositivos de almacenamiento sean leídos por el sistema informático con el fin de realizar los procesos descritos en este documento. También se podría considerar que el sistema inventivo se implemente como un soporte de almacenamiento legible por ordenador, configurado con un programa informático, en donde el soporte de almacenamiento configurado de este modo provoca que el sistema informático funcione de una manera específica y predefinida para realizar las funciones descritas en este documento.

Se ha descrito una serie de realizaciones de la presente invención.

Sin embargo, se habrá de entender que se pueden contemplar varias modificaciones sin alejarse del alcance de la invención. Por ejemplo, algunas de las etapas descritas en este documento podrían ser independientes del orden y, por consiguiente, se podrían realizar en un orden diferente al orden descrito.

5

REIVINDICACIONES

1. Un flujo digital de bits, que comprende bits de datos que representan audio (101), metadatos para el audio (101) e información de verificación de metadatos, incluyendo dicha información de verificación de metadatos una copia, o una copia de datos comprimidos, de dichos metadatos, pudiendo ser usada dicha información de verificación para detectar si los metadatos y la copia de los mismos están o no dentro de una diferencia de umbral entre sí, y si no lo están, reemplazar los metadatos por la copia.

5
2. Un flujo digital de bits de acuerdo con la reivindicación 1, en el que la información de verificación de metadatos está cifrada.

10
3. Un flujo digital de bits de acuerdo con cualquiera de las reivindicaciones 1-2, en el que los bits que representan la información de verificación de metadatos sustituyen la totalidad o algunos de la pluralidad de bits del flujo de bits que normalmente no llevan información.

15
4. Un flujo digital de bits de acuerdo con cualquiera de las reivindicaciones 1-3, en el que la información de verificación de metadatos está codificada de manera esteganográfica en el flujo de bits.

20
5. Un flujo digital de bits de acuerdo con cualquiera de las reivindicaciones 1-4, en el que el audio (101) es un audio de datos comprimidos.

25
6. Un codificador (100) que genera un flujo digital de bits de acuerdo con una cualquiera de las reivindicaciones 1-5.

30
7. Un descodificador que recibe un flujo digital de bits de acuerdo con una cualquiera de las reivindicaciones 1-5, en donde el descodificador descodifica los bits de datos que representan audio (101) usando dichos metadatos y dicha información de verificación de metadatos.

35
8. Un descodificador de acuerdo con la reivindicación 7, en donde el descodificador, en la descodificación de los bits de datos que representan el audio, cambia los metadatos usando dicha información de verificación de metadatos y usa tales metadatos cambiados en la descodificación del audio (101).

40
9. Un proceso para generar un flujo digital de bits en respuesta al audio (101), comprendiendo el proceso:

45

generar metadatos para el audio (101),

generar información de verificación de metadatos, incluyendo dicha información de verificación de metadatos una copia, o una copia de datos comprimidos, de dichos metadatos, pudiendo ser usada dicha información de verificación para detectar si los metadatos y la copia de los mismos se encuentran o no dentro de una diferencia de umbral entre sí, y si no lo estuvieran, reemplazar los metadatos por la copia, y

ensamblar un flujo digital de bits que incluye los bits de datos que representan el audio, los metadatos y la información de verificación de metadatos.
10. El proceso de acuerdo con la reivindicación 9, en el que dicha generación de metadatos genera metadatos basados en una medida del audio (101).

50
11. El proceso de acuerdo con la reivindicación 10, en el que dicha medida del audio (101) es una medida de la intensidad acústica del audio.

55
12. Un proceso de acuerdo con cualquiera de las reivindicaciones 9-11, en el que la información de verificación de metadatos, que se puede usar para detectar y cambiar los metadatos, incluye una copia o una copia de datos comprimidos de una versión correcta de tales metadatos.

60
13. Un proceso de acuerdo con cualquiera de las reivindicaciones 9-12, en el que la información de verificación está cifrada.

65
14. Un proceso de acuerdo con cualquiera de las reivindicaciones 9-13, en el que los bits que representan la información de verificación de metadatos sustituyen la totalidad o algunos de la pluralidad de bits del flujo de bits que normalmente no llevan información.

70
15. Un proceso de acuerdo con cualquiera de las reivindicaciones 9-13, en el que la información de verificación está codificada de manera esteganográfica en el flujo de bits.

75
16. Un proceso de tratamiento de un flujo de bits de audio digital que incluye bits de datos que representan audio (101), metadatos de audio (101) e información de verificación de metadatos de audio, incluyendo la información de verificación de metadatos una copia, o una copia de datos comprimidos, de dichos metadatos de audio, pudiendo

80

ser usada dicha información de verificación para detectar si los metadatos y una copia tal de los mismos se encuentran o no dentro de una diferencia de umbral entre sí, y si no lo estuvieran, reemplazar los metadatos por la copia, que comprende:

- 5 cambiar (403) los metadatos, y
cambiar (403) la información de verificación de manera que los metadatos y la copia, o la copia de datos comprimidos, de los metadatos se encuentren dentro de una diferencia de umbral entre sí.
- 10 17. Un programa informático, almacenado en un soporte legible por ordenador, diseñado para hacer que un ordenador realice los métodos de una cualquiera de las reivindicaciones 9 a 16.

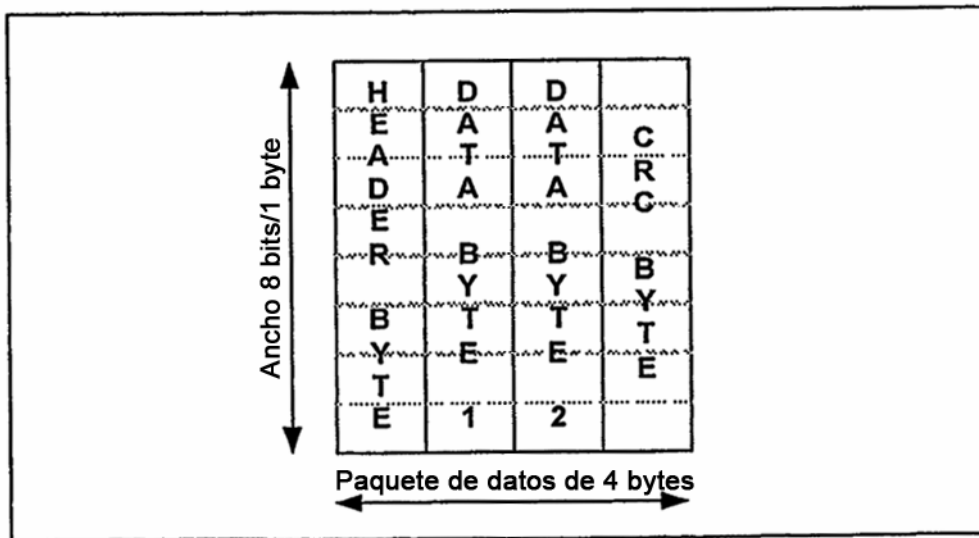
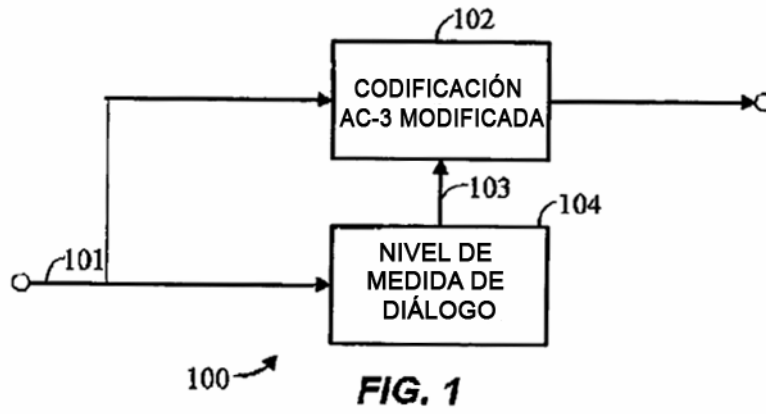


FIG. 2

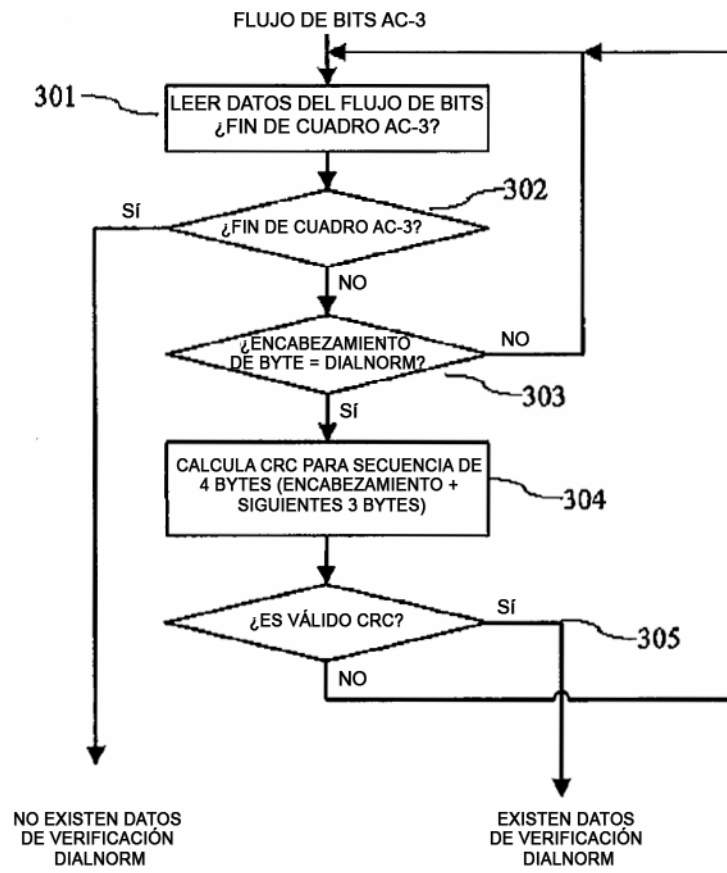


FIG. 3

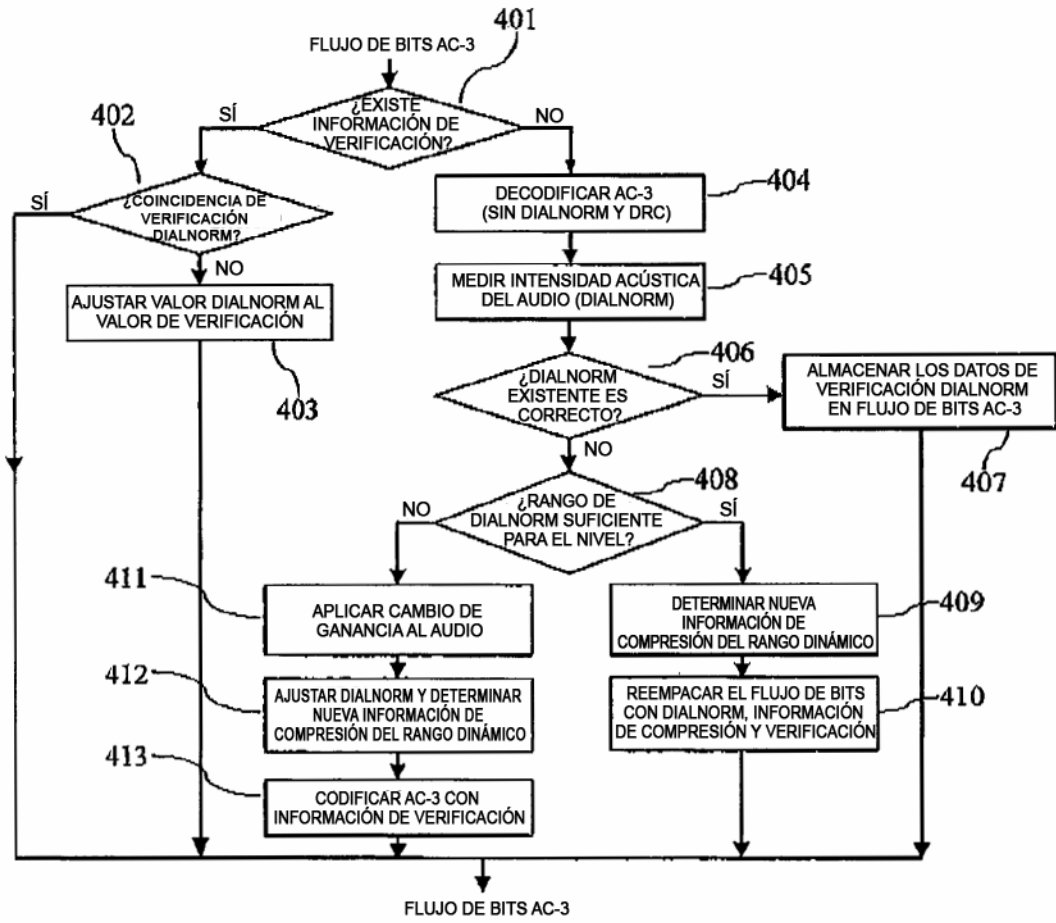
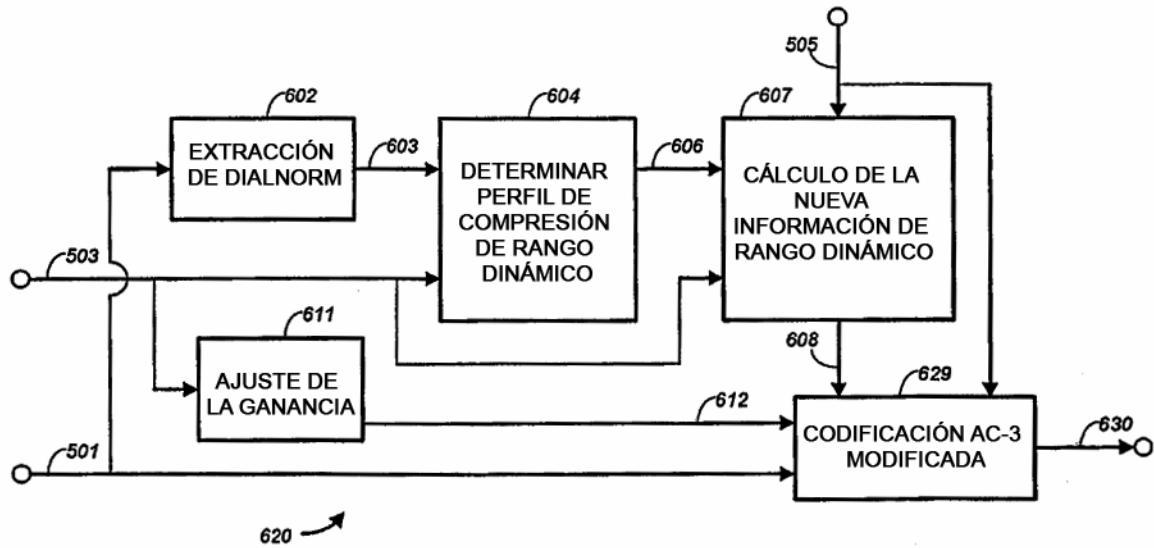
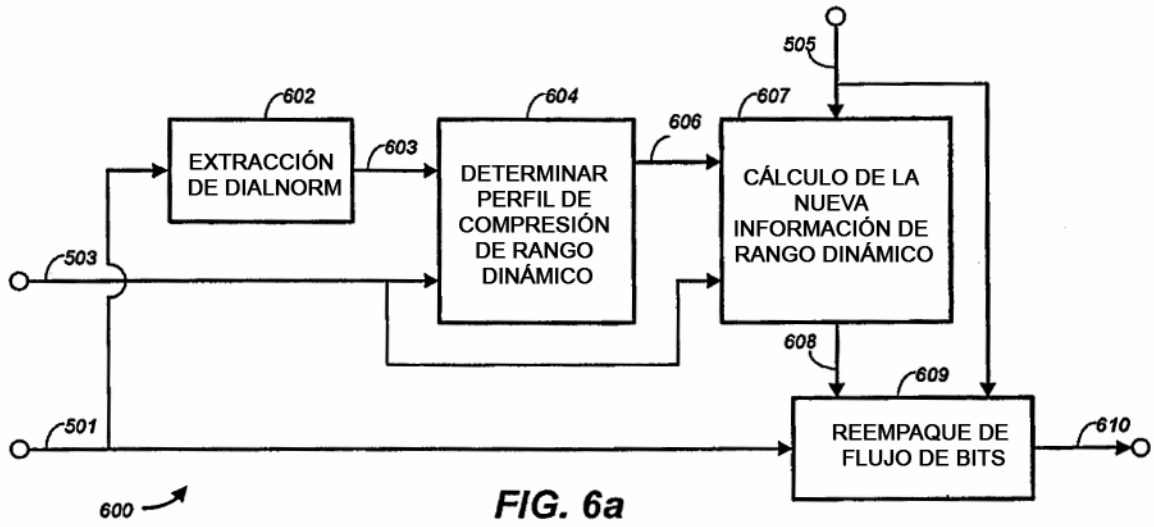


FIG. 4



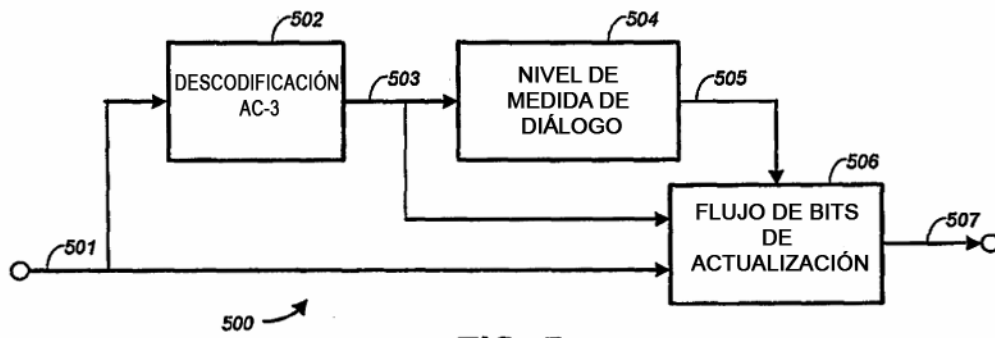


FIG. 5

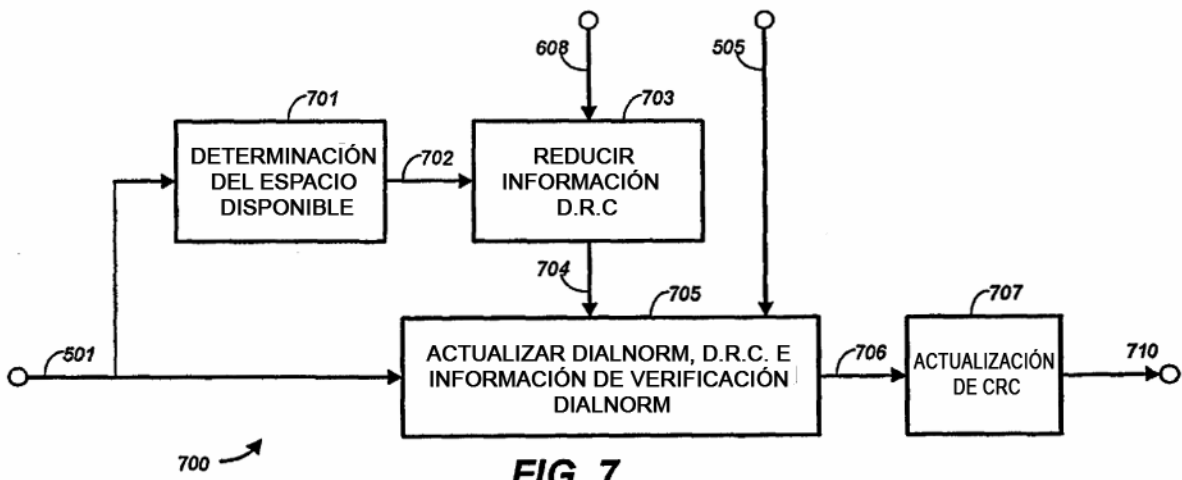


FIG. 7

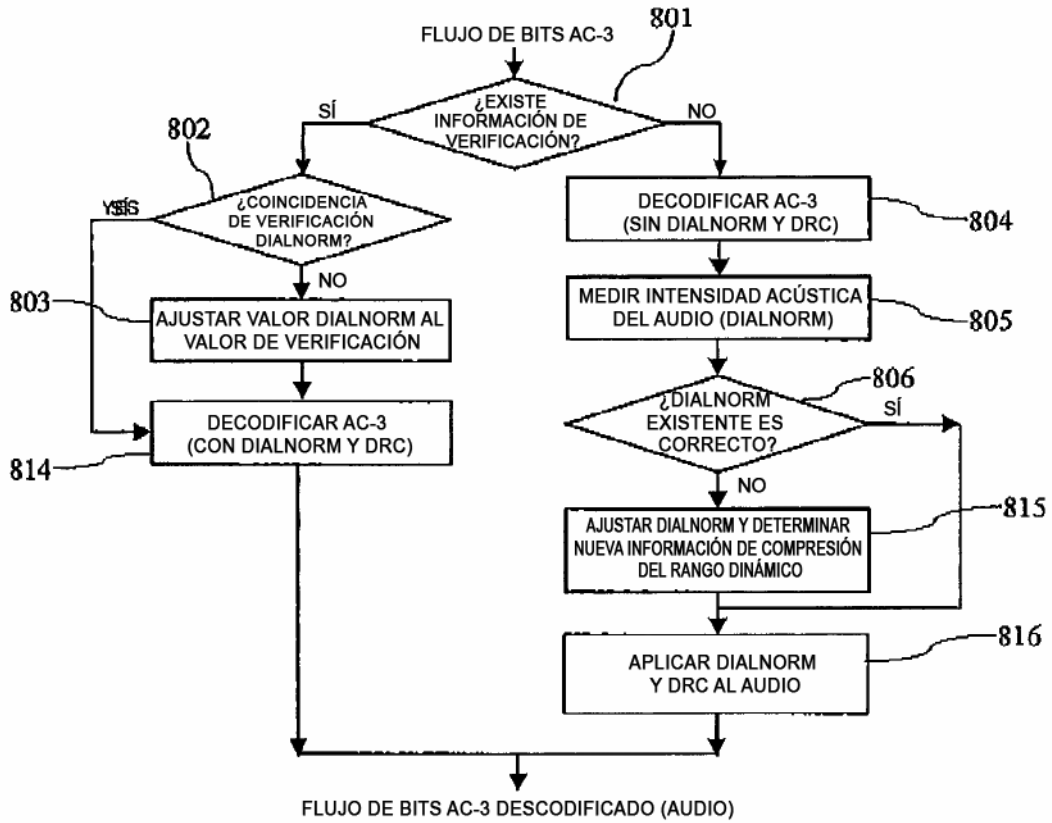


FIG. 8

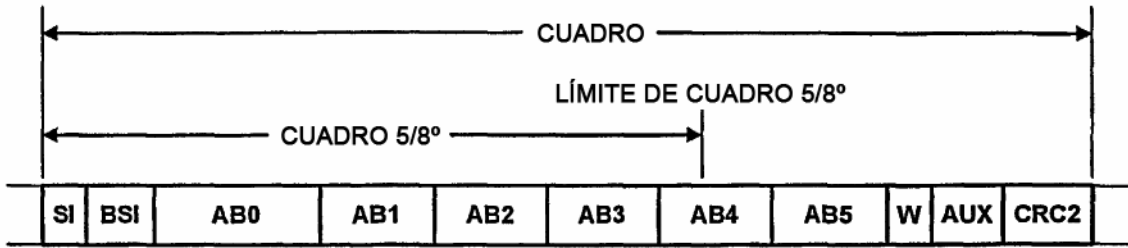


FIG. 9a

LÍMITE AB1-AB2

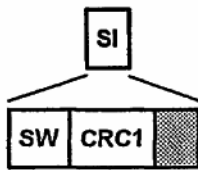


FIG. 9b

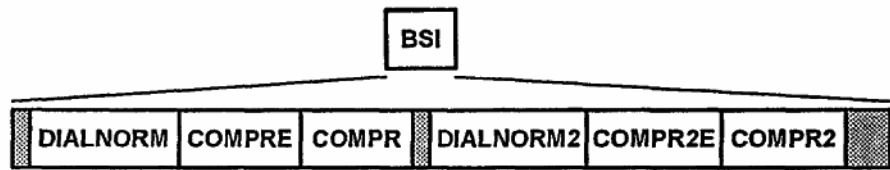


FIG. 9c

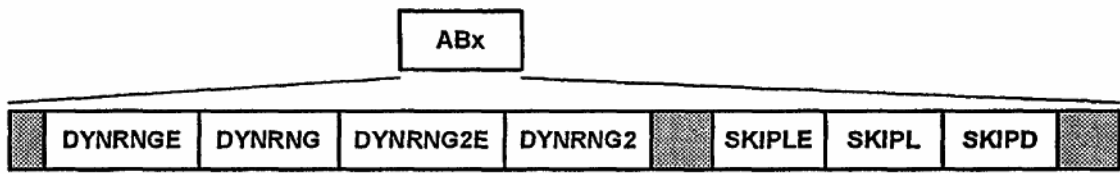


FIG. 9d

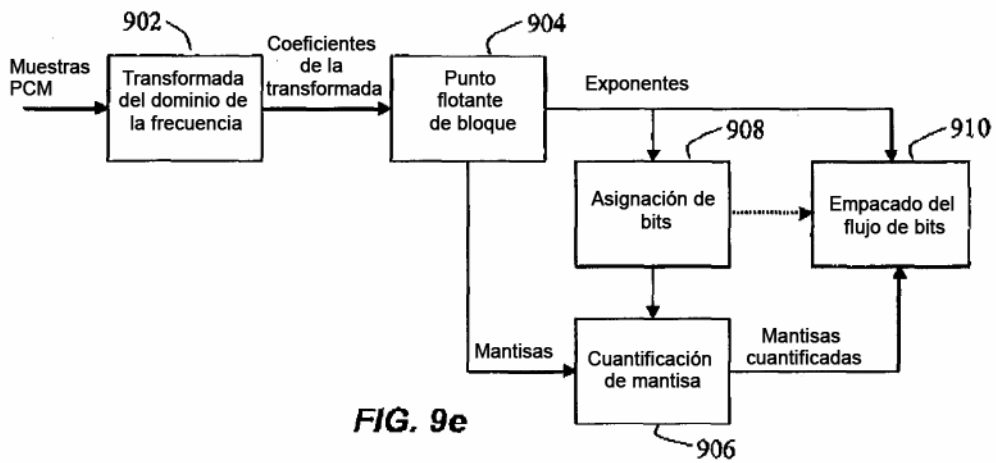


FIG. 9e

