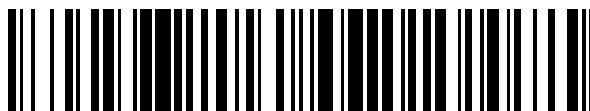


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 674 819**

51 Int. Cl.:

G10L 19/008 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.01.2015 PCT/US2015/013267**

87 Fecha y número de publicación internacional: **06.08.2015 WO15116666**

96 Fecha de presentación y número de la solicitud europea: **28.01.2015 E 15706306 (6)**

97 Fecha y número de publicación de la concesión europea: **04.04.2018 EP 3100263**

54 Título: **Transición de coeficientes ambisónicos ambientales de orden superior**

30 Prioridad:

30.01.2014 US 201461933706 P
30.01.2014 US 201461933714 P
07.03.2014 US 201461949583 P
07.03.2014 US 201461949591 P
28.05.2014 US 201462004067 P
25.07.2014 US 201462029173 P
12.01.2015 US 201514594533

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
04.07.2018

73 Titular/es:

QUALCOMM INCORPORATED (100.0%)
ATTN: International IP Administration, 5775
Morehouse Drive
San Diego, California 92121-1714, US

72 Inventor/es:

PETERS, NILS GÜNTHER y
SEN, DIPANJAN

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 674 819 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

Transición de coeficientes ambisónicos ambientales de orden superior

5 **CAMPO TÉCNICO**

[1] Esta divulgación se refiere a datos de audio y, más específicamente, a la compresión de datos de audio ambisónicos de orden superior.

10 **ANTECEDENTES**

[2] Una señal ambisónica de orden superior (HOA) (a menudo representada por una pluralidad de coeficientes armónicos esféricos (SHC) u otros elementos jerárquicos) es una representación tridimensional de un campo sonoro. Esta representación de HOA o SHC puede representar el campo sonoro de una manera que sea independiente de la geometría del altavoz local utilizado para reproducir una señal de audio multicanal producida a partir de esta señal SHC. Esta señal SHC también puede facilitar la compatibilidad inversa, ya que la señal SHC puede reproducirse en formatos multicanal bien conocidos y sumamente adoptados, tales como un formato de canal de audio 5.1 o un formato de canal de audio 7.1. La representación SHC puede, por lo tanto, permitir una mejor representación de un campo sonoro que también asimila la compatibilidad inversa.

[3] En el artículo "RM1-HOA Working Draft Text" ["Texto borrador de trabajo de RM1-HOA"] de D. Sen et al (107ª conferencia de MPEG; 13-1-2014 al 17-1-2014; San José; ISO / IEC JTC1 / SC29 / WG11 MPEG2014 / M31827, documento núm. m31827, 11 de enero de 2014), se describe una tecnología que constituye el Modelo de Referencia 1 para la codificación de contenido Ambisónico de Orden Superior en una norma prevista de audio tridimensional MPEG-H.

SUMARIO

[4] En general, se describen técnicas para la compresión y descompresión de datos de audio ambisónicos de orden superior. Los datos de audio ambisónicos de orden superior pueden comprender al menos un coeficiente armónico esférico correspondiente a una función de base armónica esférica que tiene un orden mayor que uno.

[5] En un aspecto, un procedimiento de producción de un flujo de bits de datos de audio codificados comprende determinar, en un codificador, cuando un coeficiente ambisónico ambiental de orden superior está en transición durante una trama, siendo el coeficiente ambisónico ambiental de orden superior representativo, al menos en parte, de un componente ambiental de un campo de sonido. El procedimiento comprende además identificar, en el codificador, un elemento de un vector que está asociado al coeficiente ambisónico de orden superior ambiental en transición, siendo el vector representativo, al menos en parte, de un componente espacial del campo de sonido. El procedimiento también comprende generar, en el codificador, y basándose en el vector, un vector reducido para incluir el elemento identificado del vector para la trama, y especificar, en el codificador, el vector reducido y una indicación de la transición del coeficiente ambisónico de orden superior ambiental durante la trama, en el flujo de bits. El procedimiento comprende además mantener la información del estado de transición basándose en el coeficiente ambisónico de orden superior ambiental en transición; y obtener el bit indicativo de la transición en función de la información del estado de transición. La información del estado de transición indica uno entre un estado no de transición, un estado de fundido inicial o un estado de fundido final. El flujo de bits se produce para incluir adicionalmente un bit indicativo de información de estado que incluye la información de estado de transición en la trama, el bit indicativo de la información de estado que permite decodificar el flujo de bits de los datos de audio codificados de la trama, sin referencia a tramas previas del flujo de bits.

[6] En otro aspecto, un dispositivo de codificación de audio está configurado para producir un flujo de bits de datos de audio codificados. El dispositivo de codificación de audio comprende una memoria configurada para almacenar un flujo de bits de datos de audio codificados, y uno o más procesadores configurados para determinar cuándo un coeficiente ambisónico de orden superior ambiental está en transición durante una trama. El coeficiente ambisónico de orden superior ambiental es representativo, al menos en parte, de un componente ambiental de un campo de sonido. Los uno o más procesadores están configurados además para identificar un elemento de un vector que está asociado al coeficiente ambisónico de orden superior ambiental en transición. El vector es representativo, al menos en parte, de un componente espacial del campo de sonido. Los uno o más procesadores también se configuraron para generar, basándose en el vector, un vector reducido para incluir el elemento identificado del vector para la trama, y especificar el vector reducido y una indicación de la transición del coeficiente ambisónico ambiental de orden superior durante la trama, en el flujo de bits. El dispositivo está configurado además para mantener la información del estado de transición en función del coeficiente ambisónico de orden superior ambiental en transición; y el bit indicativo de la transición en función de la información del estado de transición. La información del estado de transición indica uno entre un estado no de transición, un estado de fundido inicial o un estado de fundido final. El flujo de bits se produce

para incluir adicionalmente un bit indicativo de información de estado que incluye la información de estado de transición en la trama, el bit indicativo de la información de estado que permite decodificar el flujo de bits de los datos de audio codificados de la trama, sin referencia a tramas previas del flujo de bits.

5 **[7]** La información de estado incluye preferiblemente información de cuantización.

[8] Alternativamente o adicionalmente, la trama puede ser emitida mediante un protocolo de transmisión por flujo.

10 **[9]** En otro aspecto, un procedimiento de decodificación de un flujo de bits de datos de audio codificados comprende obtener, en un decodificador y a partir de una trama del flujo de bits, un vector reducido representativo, al menos en parte, de un componente espacial de un campo de sonido. El procedimiento también comprende obtener, en el decodificador y a partir de la trama, una indicación de una transición de un coeficiente ambisónico ambiental de orden superior, representativo, al menos en parte, de un componente ambiental de un campo de sonido. El vector reducido incluye un elemento vectorial asociado al coeficiente ambisónico de orden superior ambiental en transición. El procedimiento también comprende mantener información de estado de transición basándose en el bit indicativo de la transición del coeficiente ambisónico de orden superior ambiental; determinar si se realiza una operación de fundido inicial o una operación de fundido final con respecto al coeficiente ambisónico de alto orden ambiental, basándose en la información del estado de transición; y realizar la operación de fundido inicial o la operación de fundido final, con respecto al coeficiente ambisónico ambiental de orden superior, en función de la determinación de si se debe realizar el fundido inicial o el fundido final en el coeficiente ambisónico ambiental de orden superior. El procedimiento comprende además obtener la información de estado de transición a partir de un bit indicativo de información de estado, permitiendo el bit indicativo de la información de estado decodificar el flujo de bits de los datos de audio codificados de la trama, sin referencia a las tramas anteriores del flujo de bits. La información del estado de transición indica uno entre un estado no de transición, un estado de fundido inicial o un estado de fundido final.

30 **[10]** En otro aspecto, un dispositivo de decodificación de audio está configurado para decodificar un flujo de bits de datos de audio codificados. El dispositivo de decodificación de audio comprende una memoria configurada para almacenar una trama de un flujo de bits de datos de audio codificados, y uno o más procesadores configurados para obtener, a partir de la trama, un vector reducido representativo, al menos en parte, de un componente espacial de un campo de sonido. Los uno o más procesadores pueden configurarse adicionalmente para obtener, a partir de la trama, una indicación de una transición de un coeficiente ambisónico ambiental de orden superior, representativo, al menos en parte, de un componente ambiental de un campo de sonido. El vector reducido incluye un elemento vectorial asociado al coeficiente ambisónico de orden superior ambiental en transición. El dispositivo también está configurado para mantener información de estado de transición basándose en el bit indicativo de la transición del coeficiente ambisónico de orden superior ambiental; determinar si se realiza una operación de fundido inicial o una operación de fundido final con respecto al coeficiente ambisónico de alto orden ambiental, basándose en la información del estado de transición; y realizar la operación de fundido inicial o la operación de fundido final, con respecto al coeficiente ambisónico ambiental de orden superior, en función de la determinación de si se realiza el fundido inicial o el fundido final del coeficiente ambisónico ambiental de orden superior. El dispositivo está configurado además para obtener la información de estado de transición desde un bit indicativo de información de estado, permitiendo el bit indicativo de la información de estado decodificar el flujo de bits de los datos de audio codificados de la trama, sin referencia a tramas previas del flujo de bits. La información del estado de transición indica uno entre un estado no de transición, un estado de fundido inicial o un estado de fundido final.

50 **[11]** Preferiblemente, el vector reducido se descuantiza basándose en información de cuantización incluida en el bit indicativo de la información de estado.

[12] Alternativamente o adicionalmente, la trama puede decodificarse para conmutar desde una primera representación de contenido a una segunda representación del contenido, en donde la segunda representación es diferente a la primera representación.

55 **[13]** En otro aspecto, un medio de almacenamiento no transitorio legible por ordenador tiene almacenadas en el mismo instrucciones que cuando se ejecutan provocan que uno o más procesadores de un dispositivo de decodificación de audio o dispositivo de codificación de audio realice uno de los procedimientos descritos anteriormente.

60 **[14]** Los detalles de uno o más aspectos de la divulgación se exponen en los dibujos adjuntos y la descripción a continuación. Otras características, objetivos y ventajas de la divulgación serán evidentes a partir de la descripción y dibujos, y a partir de las reivindicaciones.

65 **BREVE DESCRIPCIÓN DE LOS DIBUJOS**

[15]

La FIG. 1 es un diagrama que ilustra funciones de base armónicas esféricas de varios órdenes y sub-órdenes.

5 La FIG. 2 es un diagrama que ilustra un sistema que puede realizar diversos aspectos de las técnicas descritas en esta divulgación.

La FIG. 3 es un diagrama de bloques que ilustra, más en detalle, un ejemplo del dispositivo codificador de audio mostrado en el ejemplo

10 de la FIG. 2, que puede realizar diversos aspectos de las técnicas descritas en esta divulgación. La FIG. 4 es un diagrama de bloques que ilustra el dispositivo decodificador de audio de la FIG. 2 en más detalle.

15 La FIG. 5A es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de codificación de audio en la realización de diversos aspectos de las técnicas de síntesis basadas en vectores, descritas en esta divulgación.

20 La FIG. 5B es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de codificación de audio en la realización de diversos aspectos de las técnicas de transición descritas en esta divulgación.

La FIG. 6A es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de decodificación de audio en la realización de varios aspectos de las técnicas descritas en esta divulgación.

25 La FIG. 6B es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de decodificación de audio en la realización de varios aspectos de las técnicas de transición descritas en esta divulgación.

Las FIGs. 7A a 7J son diagramas que ilustran una parte del flujo de bits o de la información del canal lateral que puede especificar los componentes espaciales comprimidos con más detalle.

30 La FIG. 8 es un diagrama que ilustra canales de audio a los que un dispositivo de decodificación de audio puede aplicar las técnicas descritas en esta divulgación.

35 La FIG. 9 es un diagrama que ilustra el fundido final de un coeficiente HOA ambiental adicional, el fundido inicial de una correspondiente contribución reconstruida de los distintos componentes y una suma de los coeficientes HOA y la contribución reconstruida.

DESCRIPCIÓN DETALLADA

40 [16] La evolución del sonido envolvente ha dejado disponibles muchos formatos de salida para el entretenimiento de hoy en día. Los ejemplos de tales formatos de sonido envolvente de consumo son en su mayoría basados en "canales" en cuanto a que, implícitamente, especifican las señales a los altavoces en ciertas coordenadas geométricas. Los formatos de sonido envolvente de consumo incluyen el popular formato 5.1 (que incluye los siguientes seis canales: delantero izquierdo (FL), delantero derecho (FR), central o delantero central, posterior izquierdo o envolvente izquierdo, posterior derecho o envolvente derecho, y los efectos de baja frecuencia (LFE)), el creciente formato 7.1, varios formatos que incluyen altavoces de altura tales como el formato 7.1.4 y el formato 22.2 (por ejemplo, para usar con el estándar de Televisión de Ultra Alta Definición). Los formatos no de consumo pueden abarcar cualquier número de altavoces (en geometrías simétricas y no simétricas) a menudo denominados "formaciones envolventes". Un ejemplo de una configuración de este tipo incluye 32 altavoces colocados en coordenadas en las esquinas de un icosaedro truncado.

55 [17] La entrada a un futuro codificador de MPEG es, optativamente, uno de tres formatos posibles: (i) audio tradicional basado en canales (como se ha expuesto anteriormente), que está destinado a ser reproducido mediante altavoces en posiciones preestablecidas; (ii) audio basado en objetos, que implica datos discretos de modulación por código de pulso (PCM) para objetos de audio individuales con metadatos asociados que contienen sus coordenadas de ubicación (entre otra información); y (iii) audio basado en la escena, que implica la representación del campo sonoro utilizando coeficientes de funciones de base armónicas esféricas (también llamados "coeficientes armónicos esféricos" o SHC, "Ambisónicos de Orden Superior" o HOA y "coeficientes HOA"). El futuro codificador de MPEG se puede describir con mayor detalle en un documento titulado 60 "Convocatoria de propuestas para audio tridimensional", de la Organización Internacional para la Estandarización/Comisión Electrotécnica Internacional (ISO)/(IEC) JTC1/SC29/WG11/N13411, publicado en enero de 2013 en Ginebra, Suiza, y disponible en <http://mpeg.chiariglione.org/sites/default/files/files/standards/parts/docs/w13411.zip>.

65 [18] Hay varios formatos basados en canales de "sonido envolvente" en el mercado. Varían, por ejemplo, desde el sistema de cine en casa 5.1 (que es el que ha tenido el mayor éxito en cuanto a irrumpir en las salas

de estar, más allá del estéreo) hasta el sistema 22.2 desarrollado por NHK (Nippon Hoso Kyokai o Corporación Difusora de Japón). Los creadores de contenido (por ejemplo, los estudios de Hollywood) desearían producir la banda sonora de una película una vez y no gastar esfuerzos en remezclarla para cada configuración de altavoces. Recientemente, las Organizaciones de Desarrollo de Normas han estado considerando maneras de proporcionar una codificación en un flujo de bits estandarizado y una decodificación subsiguiente que sea adaptable e indiferente con respecto a la geometría (y el número) de los altavoces y a las condiciones acústicas en la ubicación de la reproducción (que implica un reproductor).

[19] Para proporcionar dicha flexibilidad a los creadores de contenido, se puede usar un conjunto jerárquico de elementos para representar un campo sonoro. El conjunto jerárquico de elementos puede referirse a un conjunto de elementos en el que los elementos están ordenados de tal manera que un conjunto básico de elementos de orden inferior proporciona una representación completa del campo sonoro modelado. A medida que el conjunto se extiende para incluir elementos de orden superior, la representación se torna más detallada, aumentando la resolución.

[20] Un ejemplo de un conjunto jerárquico de elementos es un conjunto de coeficientes armónicos esféricos (SHC). La siguiente expresión demuestra una descripción o representación de un campo sonoro usando SHC:

$$p_i(t, r_r, \theta_r, \varphi_r) = \sum_{\omega=0}^{\infty} \left[4\pi \sum_{n=0}^{\infty} j_n(kr_r) \sum_{m=-n}^n A_n^m(k) Y_n^m(\theta_r, \varphi_r) \right] e^{j\omega t},$$

[21] La expresión muestra que la presión p_i en cualquier punto $\{r_r, \theta_r, \varphi_r\}$ del campo de sonido, en el momento t , puede ser representada únicamente por los SHC. $A_n^m(k)$. Aquí, $k = \frac{\omega}{c}$, c es la velocidad del sonido (~343 m/s), $\{r_r, \theta_r, \varphi_r\}$ es un punto de referencia (o punto de observación), $j_n(\cdot)$ es la función esférica de Bessel de orden n y $Y_n^m(\theta_r, \varphi_r)$ son las funciones básicas armónicas esféricas de orden n y sub-orden m . Se puede reconocer que el término entre corchetes es una representación en el dominio de la frecuencia de la señal (es decir, $S(\omega, r_r, \theta_r, \varphi_r)$) que puede ser aproximada por varias transformaciones de tiempo-frecuencia, tales como la transformación discreta de Fourier (DFT), la transformación de coseno discreta (DCT) o una transformación de ondículas. Otros ejemplos de conjuntos jerárquicos incluyen conjuntos de coeficientes de transformación de ondículas y otros conjuntos de coeficientes de funciones de base de multi-resolución.

[22] La FIG. 1 es un diagrama que ilustra las funciones de base armónicas esféricas desde el orden cero ($n = 0$) hasta el cuarto orden ($n = 4$). Como puede verse, para cada orden, hay una expansión de los sub-órdenes m que se muestran pero no se señalan explícitamente en el ejemplo de la FIG. 1 para facilitar la ilustración.

[23] Los SHC $A_n^m(k)$ pueden ser físicamente adquiridos (por ejemplo, registrados) con varias configuraciones de formaciones de micrófonos o, alternativamente, se pueden obtener de descripciones basadas en canales o basadas en objetos del campo sonoro. Los SHC representan el audio basado en la escena, donde los SHC pueden ser introducidos en un codificador de audio para obtener SHC codificados que pueden fomentar una transmisión o un almacenamiento más eficaz. Por ejemplo, puede utilizarse una representación de cuarto orden que implique coeficientes $(1 + 4)^2$ (25 y, por lo tanto, de cuarto orden).

[24] Como se ha señalado anteriormente, los SHC pueden obtenerse de una grabación con micrófono utilizando una formación de micrófonos. En el artículo de Poletti, M., "Three-Dimensional Surround Sound Systems Based on Spherical Harmonics" ["Sistemas tridimensionales de sonido envolvente basados en armónicos esféricos"], J. Audio Eng. Soc., Vol. 53, N° 11, noviembre de 2005, págs. 1004 a 1025, se describen varios ejemplos de cómo se pueden obtener los SHC a partir de formaciones de micrófonos.

[25] Para ilustrar cómo los SHC se pueden obtener a partir de una descripción basada en objetos, considérese la siguiente ecuación. Los coeficientes $A_n^m(k)$ para el campo sonoro correspondiente a un objeto de audio individual pueden expresarse como:

$$A_n^m(k) = g(\omega)(-4\pi ik)h_n^{(2)}(kr_s)Y_n^{m*}(\theta_s, \varphi_s),$$

donde i es $\sqrt{-1}$, $h_n^{(2)}(\cdot)$ la función esférica de Hankel (del segundo tipo) de orden n y $\{r_s, \theta_s, \varphi_s\}$ es la ubicación del objeto. Conocer la energía fuente del objeto $g(\omega)$ como función de la frecuencia (por ejemplo, utilizando técnicas de análisis de tiempo-frecuencia, tales como realizar una transformación rápida de Fourier en el flujo de PCM) nos permite convertir cada objeto de PCM y la ubicación correspondiente en el SHC $A_n^m(k)$. Además, se puede mostrar (dado que lo anterior es una descomposición lineal y ortogonal) que los coeficientes $A_n^m(k)$ para cada objeto son

aditivos. De esta manera, una multitud de objetos de PCM puede ser representada por los coeficientes $A_n^m(k)$ (por ejemplo, como una suma de los vectores de coeficientes para los objetos individuales). Esencialmente, los coeficientes contienen información sobre el campo sonoro (la presión como una función de las coordenadas tridimensionales), y lo anterior representa la transformación de los objetos individuales a una representación del campo sonoro global en las proximidades del punto de observación $\{r_r, \theta_r, \varphi_r\}$. Las cifras restantes se describen a continuación en el contexto de la codificación de audio basada en objetos y basada en SHC.

[26] La FIG. 2 es un diagrama que ilustra un sistema 10 que puede realizar diversos aspectos de las técnicas descritas en esta divulgación. Como se muestra en el ejemplo de la FIG. 2, el sistema 10 incluye un dispositivo creador de contenido 12 y un dispositivo consumidor de contenido 14. Aunque se describen en el contexto del dispositivo creador de contenido 12 y del dispositivo consumidor de contenido 14, las técnicas pueden implementarse en cualquier contexto en el que los SHC (que también pueden denominarse coeficientes HOA), o cualquier otra representación jerárquica de un campo sonoro, se codifican para formar un flujo de bits representativo de los datos de audio. Además, el dispositivo creador de contenido 12 puede representar cualquier forma de dispositivo informático capaz de implementar las técnicas descritas en esta divulgación, incluido un equipo manual (o teléfono celular), un ordenador de tableta, un teléfono inteligente o un ordenador de sobremesa, para proporcionar unos pocos ejemplos. Del mismo modo, el dispositivo consumidor de contenido 14 puede representar cualquier forma de dispositivo informático capaz de implementar las técnicas descritas en esta divulgación, incluido un equipo manual (o teléfono celular), un ordenador de tableta, un teléfono inteligente, un decodificador o un ordenador de sobremesa, para proporcionar unos pocos ejemplos.

[27] El dispositivo creador de contenido 12 puede ser operado por un estudio de cine u otra entidad que pueda generar contenido de audio multicanal para el consumo por los operadores de los consumidores de contenido, tales como el dispositivo consumidor de contenido 14. En algunos ejemplos, el dispositivo creador de contenido 12 puede ser operado por un usuario individual que desea comprimir los coeficientes HOA 11. A menudo, este creador de contenido genera contenido de audio junto con el contenido de vídeo. El dispositivo consumidor de contenido 14 puede ser operado por un individuo. El dispositivo consumidor de contenido 14 puede incluir un sistema de reproducción de audio 16, que puede referirse a cualquier forma de sistema de reproducción de audio capaz de representar los SHC para su reproducción como contenido de audio multicanal.

[28] El dispositivo creador de contenido 12 incluye un sistema de edición de audio 18. El dispositivo creador de contenido 12 obtiene grabaciones en vivo 7 en diversos formatos (incluso directamente como coeficientes HOA) y objetos de audio 9, que el dispositivo creador de contenido 12 puede editar utilizando el sistema de edición de audio 18. El creador de contenido puede, durante el proceso de edición, representar los coeficientes HOA 11 de los objetos de audio 9, escuchando las señales de altavoces representadas en un intento de identificar varios aspectos del campo sonoro que requieran una edición adicional. El dispositivo creador de contenido 12 puede entonces editar los coeficientes HOA 11 (potencialmente, de manera indirecta mediante la manipulación de diferentes objetos entre los objetos de audio 9 a partir de los cuales pueden obtenerse los coeficientes HOA originales, de la manera descrita anteriormente). El dispositivo creador de contenido 12 puede emplear el sistema de edición de audio 18 para generar los coeficientes HOA 11. El sistema de edición de audio 18 representa cualquier sistema capaz de editar datos de audio y emitir los datos de audio como uno o más coeficientes armónicos esféricos originales.

[29] Cuando se completa el proceso de edición, el dispositivo creador de contenido 12 puede generar un flujo de bits 21 basándose en los coeficientes HOA 11. Es decir, el dispositivo creador de contenido 12 incluye un dispositivo de codificación de audio 20 que representa un dispositivo configurado para codificar o comprimir de otro modo los coeficientes HOA 11 de acuerdo a diversos aspectos de las técnicas descritas en esta divulgación para generar el flujo de bits 21. El dispositivo de codificación de audio 20 puede generar el flujo de bits 21 para su transmisión, por ejemplo, a través de un canal de transmisión, que puede ser un canal cableado o inalámbrico, un dispositivo de almacenamiento de datos o algo similar. El flujo de bits 21 puede representar una versión codificada de los coeficientes HOA 11 y puede incluir un flujo de bits primario y otro flujo de bits lateral, que se puede denominar información de canal lateral.

[30] Aunque se describe con más detalle a continuación, el dispositivo de codificación de audio 20 puede configurarse para codificar los coeficientes HOA 11 basándose en una síntesis basada en vectores o una síntesis basada en la dirección. Para determinar si se realiza la metodología de descomposición basada en vectores o una metodología de descomposición basada en la dirección, el dispositivo de codificación de audio 20 puede determinar, basándose, al menos en parte, en los coeficientes HOA 11, ya sea que los coeficientes HOA 11 se generaran mediante una grabación natural de un campo sonoro (por ejemplo, grabación en vivo 7) o se produjeran artificialmente (es decir, sintéticamente) a partir de, por ejemplo, objetos de audio 9, tales como un objeto PCM. Cuando los coeficientes HOA 11 se generaron a partir de los objetos de audio 9, el dispositivo de codificación de audio 20 puede codificar los coeficientes HOA 11 usando la metodología de descomposición basada en la dirección. Cuando los coeficientes HOA 11 se capturaron en vivo usando, por

ejemplo, un eigenmike, el dispositivo de codificación de audio 20 puede codificar los coeficientes HOA 11 basándose en la metodología de descomposición basada en vectores. La distinción anterior representa un ejemplo de dónde puede desplegarse la metodología de descomposición basada en vectores o direccional. Puede haber otros casos en que una o ambas pueden ser útiles para las grabaciones naturales, el contenido generado artificialmente o una mezcla de los dos (contenido híbrido). Además, también es posible utilizar ambas metodologías simultáneamente para codificar una única trama temporal de los coeficientes HOA.

[31] Suponiendo, con fines de ilustración, que el dispositivo de codificación de audio 20 determina que los coeficientes HOA 11 se capturaran en vivo o que representan de otro modo grabaciones en vivo, tales como la grabación en vivo 7, el dispositivo de codificación de audio 20 puede configurarse para codificar los coeficientes HOA 11 utilizando una metodología de descomposición basada en vectores que implica la aplicación de una transformación lineal invertible (LIT). Un ejemplo de la transformación invertible lineal se denomina "descomposición en valores singulares" (o "SVD"). En este ejemplo, el dispositivo de codificación de audio 20 puede aplicar una SVD a los coeficientes HOA 11 para determinar una versión descompuesta de los coeficientes HOA 11. El dispositivo de codificación de audio 20 puede entonces analizar la versión descompuesta de los coeficientes HOA 11 para identificar varios parámetros, lo que puede facilitar el reordenamiento de la versión descompuesta de los coeficientes HOA 11. El dispositivo de codificación de audio 20 puede entonces reordenar la versión descompuesta de los coeficientes HOA 11 basándose en los parámetros identificados, donde dicho reordenamiento, tal como se describe con más detalle a continuación, puede mejorar la eficacia de la codificación, dado que la transformación puede reordenar los coeficientes HOA entre las tramas de los coeficientes HOA (donde una trama puede incluir M muestras de los coeficientes HOA 11 y M está, en algunos ejemplos, fijado en 1024). Después de reordenar la versión descompuesta de los coeficientes HOA 11, el dispositivo de codificación de audio 20 puede seleccionar la versión descompuesta de los coeficientes HOA 11 representativos de los componentes de primer plano (o, en otras palabras, distintos, predominantes o salientes) del campo sonoro. El dispositivo de codificación de audio 20 puede especificar la versión descompuesta de los coeficientes HOA 11 representativos de los componentes de primer plano, como un objeto de audio y la información de dirección asociada.

[32] El dispositivo de codificación de audio 20 también puede realizar un análisis de campo sonoro con respecto a los coeficientes HOA 11 con el fin, al menos en parte, de identificar los coeficientes HOA 11 representativos de uno o más componentes de trasfondo (o, en otras palabras, ambientales) del campo sonoro. El dispositivo de codificación de audio 20 puede realizar compensación de energía con respecto a los componentes de trasfondo, dado que, en algunos ejemplos, los componentes de trasfondo solo pueden incluir un subconjunto de cualquier muestra dada de los coeficientes HOA 11 (por ejemplo, tales como los coeficientes HOA 11 correspondientes a las funciones de base esférica de orden cero y de primer orden, y no los coeficientes HOA 11 correspondientes a las funciones de base esférica de segundo orden o de orden superior). Cuando se realiza la reducción de orden, en otras palabras, el dispositivo de codificación de audio 20 puede aumentar (por ejemplo, sumar/restar energía a/desde) los coeficientes HOA de trasfondo restantes de los coeficientes HOA 11 para compensar el cambio en la energía total que resulta de realizar la reducción de orden.

[33] El dispositivo de codificación de audio 20 puede realizar, a continuación, una forma de codificación psicoacústica (tal como la envolvente de MPEG, MPEG-AAC, MPEG-USAC u otras formas conocidas de codificación psicoacústica) con respecto a cada uno de los coeficientes HOA 11 representativos de los componentes de trasfondo y a cada uno de los objetos de audio de primer plano. El dispositivo de codificación de audio 20 puede realizar una forma de interpolación con respecto a la información direccional de primer plano y luego realizar una reducción de orden con respecto a la información direccional de primer plano interpolada, para generar información direccional de primer plano reducida en orden. El dispositivo de codificación de audio 20 puede realizar adicionalmente, en algunos ejemplos, una cuantización con respecto a la información direccional de primer plano reducida en orden, emitiendo información direccional de primer plano codificada. En algunos casos, la cuantización puede comprender una cuantización escalar o por entropía. El dispositivo de codificación de audio 20 puede formar entonces el flujo de bits 21 para incluir los componentes de trasfondo codificados, los objetos de audio de primer plano codificados y la información direccional cuantizada. El dispositivo de codificación de audio 20 puede entonces transmitir o emitir de otro modo el flujo de bits 21 al dispositivo consumidor de contenido 14.

[34] Si bien se muestra en la FIG. 2 como transmitido directamente al dispositivo consumidor de contenido 14, el creador de contenido 12 puede emitir el flujo de bits 21 a un dispositivo intermedio situado entre el dispositivo creador de contenido 12 y el dispositivo consumidor de contenido 14. Este dispositivo intermedio puede almacenar el flujo de bits 21 para su posterior suministro al dispositivo consumidor de contenido 14, que puede solicitar el flujo de bits. El dispositivo intermedio puede comprender un servidor de ficheros, un servidor de la Red, un ordenador de sobremesa, un ordenador portátil, un ordenador de tableta, un teléfono móvil, un teléfono inteligente o cualquier otro dispositivo capaz de almacenar el flujo de bits 21 para su posterior recuperación mediante un decodificador de audio. Este dispositivo intermedio puede residir en una red de suministro de contenido, capaz de transmitir en flujo continuo el flujo de bits 21 (y posiblemente junto con la transmisión de un correspondiente flujo de bits de datos de vídeo) a los abonados, tales como el dispositivo

consumidor de contenido 14 que solicita el flujo de bits 21.

[35] Alternativamente, el dispositivo creador de contenido 12 puede almacenar el flujo de bits 21 en un medio de almacenamiento, tal como un disco compacto, un disco de vídeo digital, un disco de vídeo de alta definición u otros medios de almacenamiento, la mayoría de los cuales pueden ser leídos por un ordenador y, por lo tanto, pueden denominarse medios de almacenamiento legibles por ordenador o medios de almacenamiento no transitorio legibles por ordenador. En este contexto, el canal de transmisión puede referirse a aquellos canales por los cuales se transmiten contenidos almacenados en los medios (y pueden incluir tiendas minoristas y otros mecanismos de suministro basados en almacenes). En cualquier caso, las técnicas de esta divulgación no deberían, por lo tanto, limitarse, a este respecto, al ejemplo de la FIG. 2.

[36] Como se muestra adicionalmente en el ejemplo de la FIG. 2, el dispositivo consumidor de contenido 14 incluye el sistema de reproducción de audio 16. El sistema de reproducción de audio 16 puede representar cualquier sistema de reproducción de audio capaz de reproducir datos de audio multicanal. El sistema de reproducción de audio 16 puede incluir un número de diferentes reproductores 22. Los reproductores 22 pueden proporcionar, cada uno, una forma diferente de reproducción, en donde las diferentes formas de reproducción pueden incluir una o más de las diversas formas de realizar la exploración de amplitud de base vectorial (VBAP) y/o una o más de las diversas formas de realizar la síntesis del campo sonoro. Como se usa en el presente documento, "A y/o B" significa "A o B" o ambos "A y B".

[37] El sistema de reproducción de audio 16 puede incluir además un dispositivo de decodificación de audio 24. El dispositivo de decodificación de audio 24 puede representar un dispositivo configurado para decodificar los coeficientes HOA 11' desde el flujo de bits 21, donde los coeficientes HOA 11' pueden ser similares a los coeficientes HOA 11, pero difieren debido a las operaciones con pérdidas (por ejemplo, cuantización) y/o a la transmisión mediante el canal de transmisión. Es decir, el dispositivo de decodificación de audio 24 puede descuantizar la información direccional de primer plano especificada en el flujo de bits 21, al mismo tiempo que realiza la decodificación psicoacústica con respecto a los objetos de audio de primer plano especificados en el flujo de bits 21 y los coeficientes HOA codificados, representativos de componentes de trasfondo. El dispositivo de decodificación de audio 24 puede realizar además la interpolación con respecto a la información direccional de primer plano decodificada y luego determinar los coeficientes HOA representativos de los componentes de primer plano, basándose en los objetos de audio de primer plano decodificados y en la información direccional de primer plano interpolada. El dispositivo de decodificación de audio 24 puede entonces determinar los coeficientes HOA 11' basándose en los coeficientes HOA determinados, representativos de los componentes de primer plano, y los coeficientes HOA decodificados, representativos de los componentes de trasfondo.

[38] El sistema de reproducción de audio 16 puede, después de decodificar el flujo de bits 21, obtener los coeficientes HOA 11' y reproducir los coeficientes HOA 11' en las señales de salida de altavoces 25. Las señales de los altavoces 25 pueden accionar uno o más altavoces (que no se muestran en el ejemplo de la FIG. 2 para facilitar la ilustración).

[39] Para seleccionar el reproductor adecuado o, en algunos casos, generar un reproductor adecuado, el sistema de reproducción de audio 16 puede obtener información de los altavoces 13, indicativa de un número de altavoces y/o una geometría espacial de los altavoces. En algunos casos, el sistema de reproducción de audio 16 puede obtener la información de los altavoces 13 usando un micrófono de referencia y accionando los altavoces de tal manera como para determinar dinámicamente la información de altavoces 13. En otros casos, o conjuntamente con la determinación dinámica de la información de altavoces 13, el sistema de reproducción de audio 16 puede inducir a un usuario a interactuar con el sistema de reproducción de audio 16 e introducir la información de altavoces 13.

[40] El sistema de reproducción de audio 16 puede entonces seleccionar uno de los reproductores de audio 22 basándose en la información de altavoces 13. En algunos casos, el sistema de reproducción de audio 16 puede, cuando ninguno de los reproductores de audio 22 está dentro de alguna medición de similitud de umbral (en cuanto a geometría de altavoces) a lo especificado en la información de altavoces 13, generar el reproductor entre los reproductores de audio 22 basándose en la información de altavoces 13. El sistema de reproducción de audio 16 puede, en algunos casos, generar uno de los procesadores de audio 22 basándose en la información de altavoces 13 sin intentar primero seleccionar un reproductor existente entre los reproductores de audio 22.

[41] La FIG. 3 es un diagrama de bloques que ilustra, en más detalle, un ejemplo del dispositivo de codificación de audio 20 mostrado en el ejemplo de la FIG. 2, que puede realizar diversos aspectos de las técnicas descritas en esta divulgación. El dispositivo de codificación de audio 20 incluye una unidad de análisis de contenido 26, una unidad de descomposición basada en vectores 27 y una unidad de descomposición basada en la dirección 28. Aunque se describe brevemente a continuación, en la Publicación de Solicitud de Patente Internacional N° WO 2014/194099, titulada "INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD" ["INTERPOLACIÓN PARA REPRESENTACIONES

DESCOMPUESTAS DE UN CAMPO DE SONIDO"], presentada el 29 de mayo de 2014, está disponible más información sobre el dispositivo de codificación de audio 20 y los diversos aspectos de compresión o codificación de otro modo de los coeficientes HOA.

5 **[42]** La unidad de análisis de contenido 26 representa una unidad configurada para analizar el contenido de los coeficientes HOA 11, para identificar si los coeficientes HOA 11 representan contenido generado a partir de una grabación en vivo o un objeto de audio. La unidad de análisis de contenido 26 puede determinar si los coeficientes HOA 11 se generaron a partir de una grabación de un campo sonoro real o de un objeto de audio artificial. En algunos casos, cuando los coeficientes HOA en tramas 11 se generaron a partir de una grabación, la unidad de análisis de contenido 26 pasa los coeficientes HOA 11 a la unidad de descomposición basada en vectores 27. En algunos casos, cuando los coeficientes HOA en tramas 11 se generaron a partir de un objeto de audio sintético, la unidad de análisis de contenido 26 pasa los coeficientes HOA 11 a la unidad de síntesis basada en la dirección 28. La unidad de síntesis basada en la dirección 28 puede representar una unidad configurada para realizar una síntesis basada en la dirección de los coeficientes HOA 11 para generar un flujo de bits basado en la dirección 21.

10 **[43]** Como se muestra en el ejemplo de la FIG. 3, la unidad de descomposición basada en vectores 27 puede incluir una unidad de transformación invertible lineal (LIT) 30, una unidad de cálculo de parámetros 32, una unidad de reordenamiento 34, una unidad de selección del primer plano 36, una unidad de compensación de energía 38, una unidad codificadora de audio psicoacústico 40, una unidad de generación de flujos de bits 42, una unidad de análisis del campo sonoro 44, una unidad de reducción de coeficientes 46, una unidad de selección de trasfondo (BG) 48, una unidad de interpolación espacio-temporal 50 y una unidad de cuantización 52.

25 **[44]** La unidad de transformación invertible lineal (LIT) 30 recibe los coeficientes HOA 11 en forma de canales HOA, siendo cada canal representativo de un bloque o trama de un coeficiente asociado a un orden dado, sub-orden de las funciones de base esférica (que se puede indicar como HOA $[k]$, donde k puede indicar la trama o bloque actual de muestras). La matriz de coeficientes HOA 11 puede tener las dimensiones D : $M \times (N+1)^2$.

30 **[45]** Es decir, la unidad de LIT 30 puede representar una unidad configurada para realizar una forma de análisis denominada descomposición en valores singulares. Aunque se describen con respecto a la SVD, las técnicas descritas en esta divulgación pueden realizarse con respecto a cualquier transformación o descomposición similar que proporcione conjuntos de salida de energía compactada y no correlacionada linealmente. También, la referencia a "conjuntos" en esta divulgación está generalmente concebida para referirse a conjuntos no nulos, a menos que se indique específicamente lo contrario y no se pretenda referirse a la definición matemática clásica de conjuntos que incluye el denominado "conjunto vacío".

35 **[46]** Una transformación alternativa puede comprender un análisis de componentes principales, que se denomina a menudo "PCA". Un PCA se refiere a un procedimiento matemático que emplea una transformación ortogonal para convertir un conjunto de observaciones de variables posiblemente correlacionadas en un conjunto de variables no correlacionadas linealmente, denominadas componentes principales. Las variables no correlacionadas linealmente representan variables que no tienen una relación (o dependencia) estadística lineal entre sí. Los componentes principales pueden ser descritos como que tienen un pequeño grado de correlación estadística entre sí. En cualquier caso, el número de componentes denominados principales es menor o igual que el número de variables originales. En algunos ejemplos, la transformación se define de tal manera que el primer componente principal tenga la varianza más grande posible (o, en otras palabras, represente tanto de la variabilidad en los datos como sea posible) y cada componente sucesivo a su vez tenga la mayor varianza posible, bajo la restricción de que el componente sucesivo sea ortogonal a (lo que puede ser re-expresado como no correlacionado con) los componentes anteriores. Un PCA puede realizar una forma de reducción de orden que, en términos de los coeficientes HOA 11, puede dar como resultado la compresión de los coeficientes HOA 11. En función del contexto, el PCA puede ser mencionado con varios nombres diferentes, tales como la transformación discreta de Karhunen-Loeve, la transformación de Hotelling, la descomposición ortogonal adecuada (POD) y la descomposición en autovalores (EVD), para nombrar unos pocos ejemplos. Las propiedades de dichas operaciones que conducen a la meta subyacente de compresión de datos de audio son la "compactación energética" y la "decorrelación" de los datos de audio multicanal.

40 **[47]** En cualquier caso, suponiendo que la unidad de LIT 30 realiza una descomposición en valores singulares (que, de nuevo, puede denominarse "SVD") con fines de ejemplo, la unidad de LIT 30 puede transformar los coeficientes HOA 11 en dos o más conjuntos de coeficientes HOA transformados. Estos "conjuntos" de coeficientes HOA transformados pueden incluir vectores de coeficientes HOA transformados. En el ejemplo de la FIG. 3, la unidad de LIT 30 puede realizar la SVD con respecto a los coeficientes HOA 11 para generar una matriz llamada V , una matriz S y una matriz U . La SVD, en álgebra lineal, puede representar una factorización de una matriz X real o compleja de y por z (donde X puede representar datos de audio multicanal, tales como los coeficientes HOA 11) de la siguiente forma:

$$X = USV^*$$

U puede representar una matriz unitaria real o compleja de y por y , donde las columnas y de U se conocen como los vectores singulares izquierdos de los datos de audio multicanal. S puede representar una matriz diagonal rectangular de y por z con números reales no negativos en la diagonal, donde los valores diagonales de S se conocen comúnmente como los valores singulares de los datos de audio multicanal. V^* (que puede indicar una traspuesta conjugada de V) puede representar una matriz unitaria real o compleja de z por z , donde las columnas z de V^* son conocidas como vectores singulares derechos de los datos de audio multicanal.

[48] Aunque se describen en esta divulgación como aplicadas a datos de audio multicanal que comprenden coeficientes HOA 11, las técnicas pueden aplicarse a cualquier forma de datos de audio multicanal. De esta manera, el dispositivo de codificación de audio 20 puede realizar una descomposición en valores singulares con respecto a datos de audio multicanal representativos de al menos una parte del campo sonoro para generar una matriz U representativa de los vectores singulares izquierdos de los datos de audio multicanal, una matriz S representativa de los valores singulares de los datos de audio multicanal y una matriz V representativa de los vectores singulares derechos de los datos de audio multicanal, y que representa los datos de audio multicanal como una función de al menos una parte de una o más entre la matriz U, la matriz S y la matriz V.

[49] En algunos ejemplos, la matriz V^* en la expresión matemática de SVD mencionada anteriormente se indica como la traspuesta conjugada de la matriz V para reflejar que la SVD se puede aplicar a matrices que comprenden números complejos. Cuando se aplica a matrices que comprenden solo números reales, la conjugada compleja de la matriz V (o, en otras palabras, la matriz V^*) puede considerarse como la traspuesta de la matriz V. A continuación se supone, con fines de facilitar la ilustración, que los coeficientes HOA 11 comprenden números reales, con el resultado de que la matriz V se emite mediante la SVD, en lugar de la matriz V^* . Además, aunque se indica como la matriz V en esta divulgación, la referencia a la matriz V debería entenderse como referencia a la traspuesta de la matriz V cuando corresponda. Aunque se supone que es la matriz V, las técnicas pueden aplicarse de manera similar a los coeficientes HOA 11 que tienen coeficientes complejos, donde la salida de la SVD es la matriz V^* . Por consiguiente, las técnicas no deberían estar limitadas a este respecto para proporcionar solamente la aplicación de la SVD para generar una matriz V, sino que pueden incluir la aplicación de la SVD a coeficientes HOA 11 que tengan componentes complejos para generar una matriz V^* .

[50] En cualquier caso, la unidad de LIT 30 puede realizar una forma en bloque de la SVD con respecto a cada bloque (que puede referirse a una trama) de datos de audio ambisónicos de orden superior (HOA) (donde los datos de audio ambisónicos incluyen bloques o muestras de los coeficientes HOA 11 o cualquier otra forma de datos de audio multicanal). Como se ha indicado anteriormente, se puede usar una variable M para indicar la longitud de una trama de audio en muestras. Por ejemplo, cuando una trama de audio incluye 1024 muestras de audio, M es igual a 1024. Aunque se describen con respecto al valor típico para M, las técnicas de la divulgación no deberían limitarse al valor típico para M. La unidad de LIT 30 puede realizar, por lo tanto, una SVD en bloques con respecto a un bloque, teniendo los coeficientes HOA 11 M por $(N+1)^2$ coeficientes HOA, donde N, de nuevo, indica el orden de los datos de audio HOA. La unidad de LIT 30 puede generar, mediante la realización de la SVD, una matriz V, una matriz S y una matriz U, donde cada una de las matrices puede representar las respectivas matrices V, S y U descritas anteriormente. De esta manera, la unidad de transformación invertible lineal 30 puede realizar una SVD con respecto a los coeficientes HOA 11 para emitir vectores $US[k]$ 33 (que pueden representar una versión combinada de los vectores S y los vectores U) que tienen dimensiones D: $M \times (N+1)^2$, y vectores $V[k]$ 35 con dimensiones D: $(N+1)^2 \times (N+1)^2$. Los elementos vectoriales individuales en la matriz $US[k]$ también pueden denominarse $X_{PS}(k)$, mientras que los vectores individuales de la matriz $V[k]$ también pueden denominarse $v(k)$.

[51] Un análisis de las matrices U, S y V puede revelar que las matrices llevan o representan características espaciales y temporales del campo sonoro subyacente representado anteriormente por X. Cada uno de los N vectores en U (de muestras de longitud M) puede representar señales de audio normalizadas separadas como una función del tiempo (para el período de tiempo representado por M muestras), que son ortogonales entre sí y que han sido desacopladas de cualquier característica espacial (que también puede denominarse información direccional). Las características espaciales, que representan la forma espacial y la anchura de la posición (r, theta, phi), pueden representarse en cambio por vectores $i^{\text{ésimos}}$ individuales, $v^{(i)}(k)$, en la matriz V (cada uno de longitud $(N+1)^2$). Los elementos individuales de cada uno de los vectores $v^{(i)}(k)$ pueden representar un coeficiente HOA que describe la forma y dirección del campo de sonido para un objeto de audio asociado. Tanto los vectores de la matriz U como los de la matriz V se normalizan de tal manera que sus energías de raíces de cuadrados medios son iguales a la unidad. La energía de las señales de audio en U está así representada por los elementos diagonales en S. Multiplicando U y S para formar $US[k]$ (con los elementos vectoriales individuales $X_{PS}(k)$), representan por tanto la señal de audio con energías verdaderas. La capacidad de la descomposición SVD para desacoplar las señales de tiempo de audio (en U), sus energías (en S) y sus características espaciales (en V) puede dar soporte a varios aspectos de las técnicas descritas en esta divulgación. Además, el modelo de síntesis de los coeficientes HOA[k] subyacentes, X, por una multiplicación de vectores de $US[k]$ y $V[k]$ da lugar al término "descomposición basada en vectores" que se

utiliza en todo este documento.

[52] Aunque se describe como realizada directamente con respecto a los coeficientes HOA 11, la unidad de LIT 30 puede aplicar la transformación invertible lineal a los derivados de los coeficientes HOA 11. Por ejemplo, la unidad de LIT 30 puede aplicar una SVD con respecto a una matriz de densidad espectral de potencia obtenida a partir de los coeficientes HOA 11. La matriz de densidad espectral de potencia se puede indicar como PSD y se obtiene mediante la multiplicación matricial de la traspuesta de la trama HoA y la tramaHoA, como se esboza en el seudo-código que sigue a continuación. La notación tramaHoA se refiere a una trama de los coeficientes HOA 11.

[53] La unidad de LIT 30 puede, después de aplicar la SVD (svd) a la PSD, puede obtener una matriz $S[k]^2$ ($S_al_cuadrado$) y una matriz $V[k]$. La matriz $S[k]^2$ puede indicar una matriz $S[k]$ al cuadrado, con lo que la unidad de LIT 30 puede aplicar una operación de raíz cuadrada a la matriz $S[k]^2$ para obtener la matriz $S[k]$. La unidad de LIT 30 puede, en algunos casos, realizar la cuantización con respecto a la matriz $V[k]$ para obtener una matriz $V[k]$ cuantizada (que puede estar indicada como matriz $V[k]'$). La unidad de LIT 30 puede obtener la matriz $U[k]$ multiplicando primero la matriz $S[k]$ por la matriz cuantizada $V[k]'$ para obtener una matriz $SV[k]'$. La unidad de LIT 30 puede obtener a continuación la seudo-inversa (pinv) de la matriz $SV[k]'$ y luego multiplicar los coeficientes HOA 11 por la seudo-inversa de la matriz $SV[k]'$ para obtener la matriz $U[k]$. Lo anterior puede estar representado por el siguiente seudo-código:

$$PSD = \text{hoaFrame}' * \text{hoaFrame};$$

$$[V, S_squared] = \text{svd}(PSD, 'econ');$$

$$S = \text{sqrt}(S_squared);$$

$$U = \text{hoaFrame} * \text{pinv}(S * V');$$

[54] Mediante la realización de la SVD con respecto a la densidad espectral de potencia (PSD) de los coeficientes HOA, en lugar de los coeficientes en sí, la unidad de LIT 30 puede reducir potencialmente la complejidad de cálculo para realizar la SVD en términos de uno o más ciclos de procesador y espacio de almacenamiento, mientras se consigue la misma eficacia en la codificación de la fuente de audio como si la SVD se aplicara directamente a los coeficientes HOA. Es decir, la SVD de tipo PSD anteriormente descrita puede ser potencialmente menos exigente en términos de cálculo porque la SVD se realiza sobre una matriz $F * F$ (siendo F el número de coeficientes HOA), en comparación con una matriz $M * F$, siendo M la longitud de trama, es decir, 1024 o más muestras. La complejidad de una SVD puede ahora, mediante la aplicación a la PSD en lugar de los coeficientes HOA 11, estar alrededor de $O(L^3)$ en comparación con $O(M * L^2)$ cuando se aplica a los coeficientes HOA 11 (donde $O(*)$ indica la notación con O mayúscula de la complejidad de cálculo común a las ciencias informáticas).

[55] La unidad de cálculo de parámetros 32 representa una unidad configurada para calcular varios parámetros, tales como un parámetro de correlación (R), parámetros de propiedades direccionales (θ , φ , r) y una propiedad de energía (e). Cada uno de los parámetros para la trama actual puede indicarse como $R[k]$, $\theta[k]$, $\varphi[k]$, $r[k]$ y $e[k]$. La unidad de cálculo de parámetros 32 puede realizar un análisis de energía y/o correlación (o lo que se denomina correlación cruzada) con respecto a los vectores $US[k]$ 33 para identificar los parámetros. La unidad de cálculo de parámetros 32 también puede determinar los parámetros para la trama anterior, donde los parámetros de tramas anteriores pueden indicarse como $R[k-1]$, $\theta[k-1]$, $\varphi[k-1]$, $r[k-1]$ y $e[k-1]$, basándose en la trama anterior del vector $US[k-1]$ y los vectores $V[k-1]$. La unidad de cálculo de parámetros 32 puede emitir los parámetros actuales 37 y los parámetros anteriores 39 para reordenar la unidad 34.

[56] La descomposición SVD no garantiza que la señal/objeto de audio, representados por el p -ésimo vector en los vectores $US[k-1]$ 33, que se puede indicar como el vector $US[k-1][p]$ (o, alternativamente, como $X_{PS}^{(p)}(k-1)$), serán los mismos señal/objeto de audio (adelantados en el tiempo) representados por el vector p -ésimo en los vectores $US[k]$ 33, que también se pueden indicar como vectores $US[k][p]$ 33 (o, alternativamente, como $X_{PS}^{(p)}(k)$). Los parámetros calculados por la unidad de cálculo de parámetros 32 pueden ser utilizados por la unidad de reordenamiento 34 para reordenar los objetos de audio para representar su evaluación natural o su continuidad en el tiempo.

[57] Es decir, la unidad de reordenamiento 34 puede comparar cada uno de los parámetros 37 de los primeros vectores $US[k]$ 33, por orden, ante cada uno de los parámetros 39 para los segundos vectores $US[k-1]$ 33. La unidad de reordenamiento 34 puede reordenar (utilizando como ejemplo un algoritmo húngaro) los diversos vectores dentro de la matriz $US[k]$ 33 y la matriz $V[k]$ 35 basándose en los parámetros actuales 37 y los parámetros anteriores 39 para emitir una matriz $US[k]$ reordenada 33' (que se puede indicar matemáticamente como $US[k]$) y una matriz $V[k]$ reordenada 35' (que se puede indicar matemáticamente como $V[k]$) a una unidad de selección de sonido de primer plano 36 (o sonido predominante - PS) ("unidad de selección del primer plano 36") y una unidad de compensación de energía 38.

- 5 [58] La unidad de análisis del campo sonoro 44 puede representar una unidad configurada para realizar un análisis de campo sonoro con respecto a los coeficientes HOA 11 para alcanzar potencialmente una velocidad de bits deseada 41. La unidad de análisis del campo sonoro 44 puede, basándose en este análisis y/o en una velocidad de bits deseada recibida 41, determinar el número total de instanciaciones del codificador psicoacústico (que puede ser una función del número total de canales ambientales o de trasfondo (BG_{TOT}) y el número de canales de primer plano o, en otras palabras, canales predominantes. El número total de instanciaciones del codificador psicoacústico se puede indicar como `númCanalesTransporteHOA`.
- 10 [59] La unidad de análisis del campo sonoro 44 también puede determinar, de nuevo para alcanzar potencialmente la velocidad de bits deseada 41, el número total de canales de primer plano (nFG) 45, el orden mínimo del campo sonoro de trasfondo (o, en otras palabras, ambiental) (N_{BG} o, alternativamente, `MínOrdenHoaAmb`), el número correspondiente de canales reales, representativo del orden mínimo del campo sonoro de trasfondo ($nBGa = (\text{MínOrdenHoa} + 1)^2$), y los índices (i) de los canales adicionales de HOA de BG a enviar (que pueden ser colectivamente indicados como información de canal de trasfondo 43 en el ejemplo de la FIG. 3). La información de canal de trasfondo 42 también puede denominarse información de canal ambiental 43. Cada uno de los canales que queda de `númCanalesTransporteHOA - nBGa` puede ser un "canal de trasfondo/ambiental adicional", un "canal predominante basado en un vector activo", una "señal predominante basada en dirección activa" o estar "completamente inactivo". En un aspecto, los tipos de canales se pueden indicar como un elemento sintáctico ("TipoCanal") con dos bits (por ejemplo, 00: señal basada en dirección; 01: señal predominante basada en vector; 10: señal ambiental adicional; 11: señal inactiva). El número total de señales de trasfondo o de ambiente, $nBGa$, puede estar dado por $(\text{MínOrdenHoaAmb} + 1)^2 +$ el número de veces que el índice 10 (en el ejemplo anterior) aparece como un tipo de canal en el flujo de bits para esa trama.
- 25 [60] En cualquier caso, la unidad de análisis del campo sonoro 44 puede seleccionar el número de canales de trasfondo (o, en otras palabras, ambientales) y el número de canales de primer plano (o, en otras palabras, predominantes) basándose en la velocidad de bits deseada 41, seleccionando más canales de trasfondo y/o de primer plano cuando la velocidad de bits deseada 41 es relativamente mayor (por ejemplo, cuando la velocidad de bits deseada 41 es igual o mayor que 512 Kbps). En un aspecto, el `númCanalesTransporteHOA` se puede fijar en 8 mientras que el `MínOrdenHoaAmb` puede ser fijado en 1 en la sección de cabecera del flujo de bits. En este escenario, en cada trama se pueden dedicar cuatro canales para representar la parte de trasfondo o ambiental del campo sonoro, mientras que los otros 4 canales pueden variar, de trama a trama, según el tipo de canal, por ejemplo, usado como un canal de trasfondo/ambiental adicional o un canal de primer plano/predominante. Las señales en primer plano/predominantes pueden ser una entre señales basadas en vectores o en la dirección, como se ha descrito anteriormente.
- 30 [61] En algunos casos, el número total de señales predominantes basadas en vectores para una trama puede estar dado por el número de veces que el índice TipoCanal es 01 en el flujo de bits de dicha trama. En el aspecto anterior, para cada canal de trasfondo/ambiental adicional (por ejemplo, correspondiente a un TipoCanal de 10), información correspondiente de cuál de los posibles coeficientes HOA (más allá de los cuatro primeros) se puede representar en ese canal. La información, para contenido HOA de cuarto orden, puede ser un índice para indicar los coeficientes HOA 5 a 25. Los cuatro primeros coeficientes HOA ambientales 1 a 4 pueden enviarse todo el tiempo cuando `mínOrdenHOAAmb` se fija en 1, por lo tanto, el dispositivo de codificación de audio solo puede necesitar indicar uno de los coeficientes HOA ambientales adicionales que tengan un índice de 5 a 25. Esta información, por tanto, podría ser enviada utilizando un elemento sintáctico de 5 bits (para el contenido de 4 orden), que puede denominarse "ÍndCoefAmbCodificado".
- 45 [62] Para ilustrarlo, supongamos que el `mínOrdenHOAAmb` se fija en 1 y que se envía un coeficiente HOA ambiental adicional con un índice de seis, mediante el flujo de bits 21, como ejemplo. En este ejemplo, el `mínOrdenHOAAmb` de 1 indica que los coeficientes HOA ambientales tienen un índice de 1, 2, 3 y 4. El dispositivo de codificación de audio 20 puede seleccionar los coeficientes HOA ambientales porque los coeficientes HOA ambientales tienen un índice menor o igual a $(\text{mínOrdenHOAAmb} + 1)^2$ o 4 en este ejemplo. El dispositivo de codificación de audio 20 puede especificar los coeficientes HOA ambientales asociados a los índices de 1, 2, 3 y 4 en el flujo de bits 21. El dispositivo de codificación de audio 20 también puede especificar el coeficiente HOA ambiental adicional con un índice de 6 en el flujo de bits, como un canalHOAAmbientalAdicional con un TipoCanal de 10. El dispositivo de codificación de audio 20 puede especificar el índice usando el elemento sintáctico `ÍndCoefAmbCodificado`. Como una cuestión práctica, el elemento `ÍndCoefAmbCodificado` puede especificar todos los índices de 1 a 25. Sin embargo, debido a que el `mínOrdenHOAAmb` está fijado en uno, el dispositivo de codificación de audio 20 no puede especificar ninguno de los primeros cuatro índices (ya que se sabe que los primeros cuatro índices están especificados en el flujo de bits 21 mediante el elemento sintáctico `mínOrdenHOAAmb`). En cualquier caso, debido a que el dispositivo de codificación de audio 20 especifica los cinco coeficientes HOA ambientales mediante el `mínOrdenHOAAmb` (para los primeros cuatro) y el `ÍndCoefAmbCodificado` (para el coeficiente HOA ambiental adicional), el dispositivo de codificación de audio 20 no puede especificar los correspondientes elementos del vector V, asociados a los coeficientes ambientales HOA que tienen un índice de 1, 2, 3, 4 y 6. Como resultado, el

dispositivo de codificación de audio 20 puede especificar el vector V con elementos [5, 7:25],

[63] En un segundo aspecto, todas las señales de primer plano/predominantes son señales basadas en vectores. En este segundo aspecto, el número total de señales en primer plano / predominantes puede estar dado por $nFG = \text{númCanalesTransporteHOA} - [(\text{MínOrdenHoaAmb} + 1)^2 + \text{cada uno de los canalAdicionalHOAAmbiental}]$.

[64] La unidad de análisis del campo sonoro 44 emite la información de canal de trasfondo 43 y los coeficientes HOA 11 a la unidad de selección de trasfondo (BG) 36, la información de canal de trasfondo 43 a la unidad de reducción de coeficientes 46 y a la unidad de generación del flujo de bits 42, y el nFG 45 a una unidad de selección de primer plano 36.

[65] La unidad de selección de trasfondo 48 puede representar una unidad configurada para determinar los coeficientes HOA de trasfondo o ambientales 47 basándose en la información de canal de trasfondo (por ejemplo, el campo sonoro de trasfondo (N_{BG}) y el número ($nBGa$) y los índices (i) de los canales HOA de BG adicionales a enviar). Por ejemplo, cuando N_{BG} es igual a uno, la unidad de selección de trasfondo 48 puede seleccionar los coeficientes HOA 11 para cada muestra de la trama de audio que tenga un orden igual o inferior a uno. La unidad de selección de trasfondo 48 puede, en este ejemplo, seleccionar entonces los coeficientes HOA 11 que tengan un índice identificado por uno de los índices (i) como coeficientes adicionales HOA de BG, en donde el $nBGa$ se proporciona a la unidad de generación del flujo de bits 42 para que se especifique en el flujo de bits 21 para permitir que el dispositivo de decodificación de audio, tal como el dispositivo de decodificación de audio 24, que se muestra en el ejemplo de las FIGs. 2 y 4, pueda analizar sintácticamente los coeficientes HOA 47 a partir del flujo de bits 21. La unidad de selección de trasfondo 48 puede emitir entonces los coeficientes HOA ambientales 47 a la unidad de compensación de energía 38. Los coeficientes HOA ambientales 47 pueden tener dimensiones D: $M \times [(N_{BG}+1)^2 + nBGa]$. Los coeficientes HOA ambientales 47 también pueden denominarse "coeficientes HOA ambientales 47", donde cada uno de los coeficientes HOA ambientales 47 corresponde a un canal HOA ambiental independiente 47, a codificar por la unidad codificadora de audio psicoacústica 40.

[66] La unidad de selección del primer plano 36 puede representar una unidad configurada para seleccionar la matriz US[k] reordenada 33' y la matriz V[k] reordenada 35', que representan componentes de primer plano, o distintos, del campo sonoro basándose en nFG 45 (que pueden representar uno o más índices que identifican los vectores de primer plano). La unidad de selección del primer plano 36 puede emitir nFG

señales 49 (que se pueden indicar como una $US[k]_{1, \dots, nFG}$ reordenada 49, $FG_1, \dots, nFG[k]$ 49, o $X_{PS}^{(1..nFG)}(k)$ 49) a la unidad codificadora de audio psicoacústica 40, en donde las nFG señales 49 pueden tener dimensiones D: $M \times nFG$ y cada una representa objetos de mono-audio. La unidad de selección del primer plano 36 también puede emitir la matriz V[k] reordenada 35' (o $v^{(1..nFG)}(k)$ 35'), correspondiente a los componentes de primer plano del campo sonoro, a la unidad de interpolación espacio-temporal 50, donde un subconjunto de la matriz V[k] reordenada 35', correspondiente a los componentes de primer plano, se puede indicar como la matriz V[k] de primer plano 51_k (que se puede indicar matemáticamente como $V_{1, \dots, nFG}[k]$) con dimensiones D: $(N+1)^2 \times nFG$.

[67] La unidad de compensación de energía 38 puede representar una unidad configurada para realizar compensación de energía con respecto a los coeficientes HOA ambientales 47, para compensar la pérdida de energía debida a la eliminación de varios de los canales HOA por la unidad de selección de trasfondo 48. La unidad de compensación de energía 38 puede realizar un análisis energético con respecto a una o más entre la matriz reordenada US[k] 33', la matriz V[k] reordenada 35', las nFG señales 49, los vectores V[k] 51_k de primer plano y los coeficientes HOA ambientales 47, y luego realizar una compensación de energía basándose en el análisis de energía para generar los coeficientes HOA ambientales compensados en energía 47'. La unidad de compensación de energía 38 puede emitir los coeficientes HOA ambientales compensados en energía 47' a la unidad codificadora de audio psicoacústica 40.

[68] La unidad de interpolación espacio-temporal 50 puede representar una unidad configurada para recibir los vectores V[k] de primer plano 51_k para la $k^{\text{ésima}}$ trama y los vectores de primer plano V[k-1] 51_{k-1} para la trama anterior (de ahí la notación k-1) y realizar una interpolación espacio-temporal para generar los vectores V[k] de primer plano interpolados. La unidad de interpolación espacio-temporal 50 puede recombinar las nFG señales 49 con los vectores V[k] de primer plano 51_k para recuperar los coeficientes HOA de primer plano reordenados. La unidad de interpolación espacio-temporal 50 puede entonces dividir los coeficientes HOA de primer plano reordenados por los vectores V[k] interpolados para generar nFG señales interpoladas 49'. La unidad de interpolación espacio-temporal 50 también puede emitir los vectores V[k] de primer plano 51_k que se usaron para generar los vectores V[k] de primer plano interpolados, de modo que un dispositivo de decodificación de audio, tal como el dispositivo de decodificación de audio 24, pueda generar los vectores V[k] de primer plano interpolados y recuperar así los vectores V[k] de primer plano 51_k . Los vectores V[k] de primer plano 51_k utilizados para generar los vectores V[k] de primer plano interpolados se indican como los vectores V[k] de primer plano restantes 53. Con el fin de asegurar que los mismos V[k] y V[k-1] se usen en el codificador

y el decodificador (para crear los vectores interpolados $V[k]$) se pueden usar versiones cuantizadas/descuantizadas de los vectores en el codificador y el decodificador.

[69] En funcionamiento, la unidad de interpolación espacio-temporal 50 puede interpolar una o más subtramas de una primera trama de audio a partir de una primera descomposición, por ejemplo, vectores $V[k]$ de primer plano 51_k , de una parte de una primera pluralidad de los coeficientes HOA 11 incluidos en la primera trama, y una segunda descomposición, por ejemplo, vectores $V[k]$ de primer plano 51_{k-1} , de una parte de una segunda pluralidad de los coeficientes HOA 11 incluidos en una segunda trama, para generar coeficientes armónicos esféricos interpolados descompuestos para las una o más sub-tramas.

[70] En algunos ejemplos, la primera descomposición comprende los primeros vectores $V[k]$ de primer plano 51_k , representativos de los vectores singulares derechos de la parte de los coeficientes HOA 11. De manera similar, en algunos ejemplos, la segunda descomposición comprende los segundos vectores $V[k]$ de primer plano 51_k , representativos de los vectores singulares derechos de la parte de los coeficientes HOA 11.

[71] En otras palabras, el audio tridimensional basado en armónicos esféricos puede ser una representación paramétrica del campo de presión tridimensional en términos de funciones de base ortogonales sobre una esfera. Cuanto mayor sea el orden N de la representación, la resolución espacial será potencialmente mayor y, a menudo, mayor será el número de coeficientes armónicos esféricos (SH) (para un total de $(N+1)^2$ coeficientes). Para muchas aplicaciones, puede requerirse una compresión de ancho de banda de los coeficientes para poder transmitir y almacenar los coeficientes de manera eficaz. Las técnicas dirigidas en esta divulgación pueden proporcionar un proceso de reducción de la dimensionalidad basado en tramas, utilizando la descomposición en valores singulares (SVD). El análisis de la SVD puede descomponer cada trama de coeficientes en tres matrices U , S y V . En algunos ejemplos, las técnicas pueden manipular algunos de los vectores en la matriz $US[k]$ como componentes de primer plano del campo sonoro subyacente. Sin embargo, cuando se manipulan de esta manera, los vectores (en la matriz $US[k]$) son discontinuos de trama en trama, incluso aunque representan el mismo componente de audio distinto. Las discontinuidades pueden conducir a distorsiones significativas cuando los componentes son suministrados mediante codificadores de audio transformado.

[72] En algunos aspectos, la interpolación espacio-temporal puede apoyarse en la observación de que la matriz V puede interpretarse como ejes espaciales ortogonales en el dominio de las armónicas esféricas. La matriz $U[k]$ puede representar una proyección de los datos de armónicos esféricos (HOA) en términos de las funciones de base, donde la discontinuidad puede atribuirse a los ejes espaciales ortogonales ($V[k]$) que cambian cada trama y, por lo tanto, son discontinuos en sí mismos. Esto es diferente a algunas otras descomposiciones, tales como la Transformación de Fourier, donde las funciones de base son, en algunos ejemplos, constantes de una trama a otra. En estos términos, la SVD puede considerarse como un algoritmo de búsqueda de coincidencia. La unidad de interpolación espacio-temporal 50 puede realizar la interpolación para mantener, potencialmente, la continuidad entre las funciones de base ($V[k]$) de una trama a otra, mediante la interpolación entre ellas.

[73] Como se ha indicado anteriormente, la interpolación puede realizarse con respecto a las muestras. El caso se generaliza en la descripción anterior cuando las sub-tramas comprenden un solo conjunto de muestras. Tanto en el caso de la interpolación sobre muestras como sobre sub-tramas, la operación de interpolación puede adoptar la forma de la siguiente ecuación:

$$\overline{v(l)} = w(l)v(k) + (1 - w(l))v(k - 1).$$

En esta ecuación anterior, la interpolación puede realizarse con respecto al único vector $V v(k)$ a partir del único vector $Vv(k-1)$, que en un aspecto podría representar vectores V a partir de las tramas adyacentes k y $k-1$. En la ecuación anterior, l representa la resolución sobre la que se está llevando a cabo la interpolación, donde l puede indicar una muestra entera y $l = 1, \dots, T$ (donde T es la longitud de las muestras sobre las que se

está realizando la interpolación y sobre las cuales los vectores interpolados de salida $\overline{v(l)}$ son necesarios, y también indica que la salida del proceso produce l de los vectores). Alternativamente, l podría indicar sub-tramas consistentes en múltiples muestras. Cuando, por ejemplo, una trama está dividida en cuatro sub-tramas, l puede comprender valores de 1, 2, 3 y 4 para cada una de las sub-tramas. El valor de l se puede señalar como un campo denominado "TiempoInterpolaciónEspacialCodificada" mediante un flujo de bits, de modo que la operación de interpolación pueda ser replicada en el decodificador. El $w(l)$ puede comprender valores de las ponderaciones de interpolación. Cuando la interpolación es lineal, $w(l)$ puede variar linealmente y monótonamente entre 0 y 1, en función de l . En otros casos, $w(l)$ puede variar entre 0 y 1 de forma no lineal pero monótona (tal como un cuarto de ciclo de un coseno elevado) en función de l . La función, $w(l)$, puede indicarse entre unas pocas posibilidades diferentes de funciones y señalizarse en el flujo de bits como un campo denominado "ProcedimientoInterpolaciónEspacial", de modo que la operación de interpolación idéntica pueda ser replicada por el decodificador. Cuando $w(l)$ tiene un valor cercano a 0, la salida, $\overline{v(l)}$, puede estar

sumamente ponderada o influenciada por $v(k-1)$. Mientras que, cuando $w(l)$ tiene un valor cercano a 1, asegura que la salida, $\overline{v(l)}$, está sumamente ponderada o influenciada por $v(k-1)$.

5 **[74]** La unidad de reducción de coeficientes 46 puede representar una unidad configurada para realizar una reducción de coeficientes con respecto a los restantes vectores $V[k]$ de primer plano 53, basándose en la información de canal de trasfondo 43 para emitir los vectores de primer plano reducidos $V[k]$ 55 a la unidad de cuantización 52. Los vectores $V[k]$ de primer plano reducidos 55 pueden tener dimensiones D: $[(N+1)^2 - (N_{BG}+1)^2 - BG_{TOT}] \times nFG$.

10 **[75]** La unidad de reducción de coeficientes 46 puede, a este respecto, representar una unidad configurada para reducir el número de coeficientes de los restantes vectores $V[k]$ de primer plano 53. En otras palabras, la unidad de reducción de coeficientes 46 puede representar una unidad configurada para eliminar los coeficientes en los vectores $V[k]$ de primer plano (que forman los restantes vectores $V[k]$ de primer plano 53) que tienen poca o ninguna información direccional. Como se ha descrito anteriormente, en algunos ejemplos, 15 los coeficientes de los vectores $V[k]$ distintos o, en otras palabras, de primer plano que corresponden a funciones de base de primer orden y de orden cero (que pueden indicarse como N_{BG}) proporcionan poca información direccional y por lo tanto pueden ser retirados de los V vectores de primer plano (mediante un proceso que se puede denominar "reducción de coeficientes"). En este ejemplo, se puede proporcionar mayor flexibilidad para identificar no solo los coeficientes que corresponden a N_{BG} , sino para identificar canales HOA 20 adicionales (que pueden indicarse con la variable TotalDeCanHOAAmbAdic) a partir del conjunto de $[(N_{BG}+1)^2+1, (N+1)^2]$. La unidad de análisis del campo sonoro 44 puede analizar los coeficientes HOA 11 para determinar BG_{TOT} , que puede identificar no solo el $(N_{BG}+1)^2$, sino el TotalDeCanHOAAmbAdic, que se puede denominar colectivamente la información de canal de trasfondo 43. La unidad de reducción de coeficientes 46 puede entonces eliminar los coeficientes correspondientes a $(N_{BG}+1)^2$ y el TotalDeCanHOAAmbAdic de los 25 restantes vectores $V[k]$ de primer plano 53, para generar una matriz $V[k]$ de menores dimensiones 55, de tamaño $((N+1)^2 - (BG_{TOT}) \times nFG$, que también se puede mencionar como los vectores $V[k]$ de primer plano reducidos 55.

30 **[76]** La unidad de cuantización 52 puede representar una unidad configurada para realizar cualquier forma de cuantización para comprimir los vectores $V[k]$ de primer plano reducidos 55 y generar vectores de primer plano $V[k]$ codificados 57, emitiendo los vectores de primer plano codificados $V[k]$ 57 a la unidad de generación del flujo de bits 42. En funcionamiento, la unidad de cuantización 52 puede representar una unidad configurada para comprimir un componente espacial del campo sonoro, es decir, uno o más de los vectores $V[k]$ de primer plano reducidos 55 en este ejemplo. A efectos de ejemplo, se supone que los vectores $V[k]$ de primer plano 35 reducidos 55 incluyen dos vectores de fila que tienen, como resultado de la reducción de coeficientes, menos de 25 elementos cada uno (lo que implica una representación HOA de cuarto orden del campo sonoro). Aunque se describe con respecto a dos vectores de fila, se puede incluir cualquier número de vectores en los vectores $V[k]$ de primer plano reducidos 55, hasta $(n+1)^2$, donde n indica el orden de la representación HOA del campo sonoro. Además, aunque se describe a continuación como realizando una cuantización escalar y/o 40 entrópica, la unidad de cuantización 52 puede realizar cualquier forma de cuantización que dé como resultado la compresión de los vectores $V[k]$ de primer plano reducidos 55.

[77] La unidad de cuantización 52 puede recibir los vectores $V[k]$ de primer plano reducidos 55 y realizar un esquema de compresión para generar los vectores $V[k]$ de primer plano codificados 57. El esquema de 45 compresión puede implicar cualquier esquema de compresión concebible para comprimir elementos de un vector, o datos en general, y no debería limitarse al ejemplo descrito más adelante con más detalle. La unidad de cuantización 52 puede realizar, como ejemplo, un esquema de compresión que incluye uno o más entre una transformación de representaciones de coma flotante de cada elemento de los vectores reducidos $V[k]$ de primer plano 55 en representaciones enteras de cada elemento de los vectores reducidos $V[k]$ de primer plano 50 55, una cuantización uniforme de las representaciones enteras de los vectores reducidos $V[k]$ de primer plano 55 y la categorización y codificación de las representaciones enteras cuantizadas de los vectores restantes $V[k]$ de primer plano 55.

55 **[78]** En algunos ejemplos, varios de los uno o más procesos del esquema de compresión pueden ser controlados dinámicamente por parámetros para conseguir, o casi conseguir, como ejemplo, una velocidad de bits deseada 41 para el flujo de bits resultante 21. Dado que cada uno de los vectores reducidos $V[k]$ de primer plano 55 son ortonormales entre sí, cada uno de los vectores reducidos $V[k]$ de primer plano 55 puede codificarse independientemente. En algunos ejemplos, tal como se describe con más detalle a continuación, cada elemento de cada vector reducido $V[k]$ de primer plano 55 puede codificarse utilizando la misma 60 modalidad de codificación (definida por varias sub-modalidades).

[79] Como se describe en la publicación nº WO 2014 / 194099, la unidad de cuantización 52 puede realizar la cuantización escalar y/o la codificación de Huffman para comprimir los vectores reducidos $V[k]$ de primer plano 55, emitiendo los vectores codificados $V[k]$ de primer plano 57, que también se pueden mencionar como 65 información de canal lateral 57. La información de canal lateral 57 puede incluir elementos sintácticos usados

para codificar los restantes vectores $V[k]$ de primer plano 55.

[80] Como se ha indicado en la publicación nº WO 2014/194099, la unidad de cuantización 52 puede generar elementos sintácticos para la información de canal lateral 57. Por ejemplo, la unidad de cuantización 52 puede especificar un elemento sintáctico en una cabecera de una unidad de acceso (que puede incluir una o más tramas) que indica cuál de la pluralidad de modalidades de configuración ha sido seleccionada. Aunque se describe como especificada por cada unidad de acceso, la unidad de cuantización 52 puede especificar el elemento sintáctico por cada trama o sobre cualquier otra base periódica o no periódica (como una vez para el flujo de bits completo). En cualquier caso, el elemento sintáctico puede comprender dos bits que indican cuál de las tres modalidades de configuración fueron seleccionadas para especificar el conjunto no nulo de coeficientes de los vectores reducidos $V[k]$ de primer plano 55 para representar los aspectos direccionales del componente distinto. El elemento sintáctico se puede indicar como "longitudVectVcodificado". De esta manera, la unidad de cuantización 52 puede señalar o especificar de otro modo en el flujo de bits cuáles de las tres modalidades de configuración se usaron para especificar los vectores codificados $V[k]$ de primer plano 57 en el flujo de bits.

[81] Por ejemplo, se pueden presentar tres modalidades de configuración en la tabla de sintaxis para DatosVecV (mencionado más adelante en este documento). En ese ejemplo, las modalidades de configuración son las siguientes: (Modalidad 0), se transmite una longitud de vector-V completa en el campo DatosVecV; (Modalidad 1), los elementos del vector V asociados al número mínimo de coeficientes para los coeficientes HOA Ambientales y todos los elementos del vector V que incluían canales HOA adicionales que no se transmiten; y (Modalidad 2), los elementos del vector V asociados al número mínimo de coeficientes para los coeficientes HOA Ambientales no se transmiten. La tabla de sintaxis de DatosVecV ilustra las modalidades en relación con un conmutador y una sentencia de casos. Aunque se describen con respecto a tres modalidades de configuración, las técnicas no deberían limitarse a tres modalidades de configuración, y pueden incluir cualquier número de modalidades de configuración, incluido una única modalidad de configuración o una pluralidad de modalidades. La publicación nº WO 2014/194099 proporciona un ejemplo diferente con cuatro modalidades. La unidad de cuantización escalar/entrópica 53 también puede especificar el indicador 63 como otro elemento sintáctico en la información del canal lateral 57.

[82] Además, aunque se describe con respecto a una forma de cuantización escalar, la unidad de cuantización 52 puede realizar la cuantización del vector o cualquier otra forma de cuantización. En algunos casos, la unidad de cuantización 52 puede alternar entre la cuantización vectorial y la cuantización escalar. Durante la cuantización escalar descrita anteriormente, la unidad de cuantización 52 puede calcular la diferencia entre dos vectores V sucesivos (sucesivos como de trama a trama) y codificar la diferencia (o, en otras palabras, el residuo). La cuantización vectorial no implica dicha codificación de diferencia (que, en cierto sentido, puede ser una forma predictiva de codificación en cuanto a que la cuantización escalar predice el vector V actual basándose en un vector V previo y una diferencia señalizada).

[83] La unidad codificadora de audio psicoacústica 40 incluida dentro del dispositivo de codificación de audio 20 puede representar múltiples instancias de un codificador de audio psicoacústico, cada una de las cuales se utiliza para codificar un objeto de audio diferente o canal HOA de cada uno de los coeficientes HOA ambientales compensados en energía 47' y las nFG señales interpoladas 49' para generar los coeficientes HOA ambientales codificados 59 y las nFG señales codificadas 61. La unidad codificadora de audio psicoacústico 40 puede emitir los coeficientes HOA ambientales codificados 59 y las nFG señales codificadas 61 a la unidad de generación de flujos de bits 42.

[84] La unidad de generación del flujo de bits 42 incluida dentro del dispositivo de codificación de audio 20 representa una unidad que formatea datos para adaptarse a un formato conocido (que puede referirse a un formato conocido por un dispositivo de decodificación), generando de este modo el flujo de bits basado en vectores 21. El flujo de bits 21 puede, en otras palabras, representar datos de audio codificados, que se han codificado de la manera descrita anteriormente. La unidad de generación de flujos de bits 42 puede representar un multiplexor en algunos ejemplos, que pueden recibir los vectores codificados $V[k]$ de primer plano 57, los coeficientes HOA codificados 59, las señales nFG codificadas 61 y la información de canal de trasfondo 43. La unidad de generación del flujo de bits 42 puede generar entonces un flujo de bits 21 basado en los vectores codificados $V[k]$ de primer plano 57, los coeficientes HOA ambientales codificados 59, las nFG señales codificadas 61 y la información de canal de trasfondo 43. El flujo de bits 21 puede incluir un flujo de bits primario o principal y uno o más flujos de bits de canal lateral.

[85] Aunque no se muestra en el ejemplo de la Fig. 3, el dispositivo de codificación de audio 20 también puede incluir una unidad de salida de flujo de bits que conmuta la salida del flujo de bits desde el dispositivo de codificación de audio 20 (por ejemplo, entre el flujo de bits basado en la dirección 21 y el flujo de bits basado en vectores 21), basándose en si una trama actual va a ser codificada usando la síntesis basada en la dirección o la síntesis basada en vectores. La unidad de salida de flujo de bits puede realizar la conmutación basándose en el elemento sintáctico emitido por la unidad de análisis de contenido 26, indicando si se realizó una síntesis basada en la dirección (como resultado de detectar que los coeficientes HOA 11 se generaron a

partir de un objeto de audio sintético) o si se realizó una síntesis basada en vectores (como resultado de detectar que se registraron los coeficientes HOA). La unidad de salida del flujo de bits puede especificar la sintaxis de cabecera correcta para indicar la conmutación o codificación actual utilizada para la trama actual, junto con la respectiva de los flujos de bits 21.

[86] Además, como se ha indicado anteriormente, la unidad de análisis de campo de sonido 44 puede identificar BG_{TOT} coeficientes HOA ambientales 47, que pueden cambiar de trama a trama (aunque a veces BG_{TOT} puede permanecer constante o igual entre dos o más tramas adyacentes (en el tiempo)). El cambio en BG_{TOT} puede dar como resultado cambios en los coeficientes expresados en los vectores $V[k]$ en primer plano reducidos 55. El cambio en BG_{TOT} puede dar como resultado los coeficientes HOA de trasfondo (que también pueden denominarse "coeficientes HOA ambientales") que cambian de trama a trama (aunque, de nuevo, a veces BG_{TOT} puede permanecer constante o igual entre dos o más tramas adyacentes (en el tiempo)). Los cambios a menudo dan como resultado una pérdida de energía para los aspectos del campo de sonido representados por la adición o eliminación de los coeficientes HOA ambientales adicionales y la correspondiente eliminación de coeficientes de, o adición de coeficientes a, los vectores $V[k]$ en primer plano reducidos 55.

[87] Para ilustrar, supongamos que para una trama anterior (indicada como " F_{X-1} "), el número total de coeficientes HOA ambientales (BG_{TOT}) incluye los coeficientes ambientales HOA asociados a los índices de 1, 2, 3 y 4 y el coeficiente HOA ambiental adicional 6. Para una trama actual (indicada como " F_X "), supongamos además que el número total de coeficientes HOA ambientales (BG_{TOT}) incluye los coeficientes HOA ambientales asociados a los índices de 1, 2, 3 y 4 y el coeficiente HOA ambiental adicional 5. El número total de coeficientes HOA ambientales (BG_{TOT}) de la trama previa (F_{X-1}) difiere, por lo tanto, del número total de coeficientes HOA ambientales (BG_{TOT}) de la trama actual (F_X) al reemplazar el coeficiente HOA ambiental adicional asociado al índice 6 por el coeficiente ambiental HOA adicional asociado al índice 5. El vector V de la trama anterior (F_{X-1}) incluye cualquier elemento al que no corresponda uno entre el número total de coeficientes HOA ambientales (BG_{TOT}) de la trama anterior F_{X-1} . Como tal, el vector V puede incluir los elementos 5 y 7 hasta 25 para una representación de cuarto orden del campo de sonido, que se puede indicar como $V[5, 7:25]$. El vector V de la trama actual (F_X) incluye cualquier elemento al que no corresponda uno entre el número total de coeficientes HOA ambientales (BG_{TOT}) de la trama actual (F_X), lo que puede indicarse como $V[6:25]$ para una representación de cuarto orden del campo de sonido.

[88] En la publicación n° WO 2014/194099, el dispositivo de codificación de audio señala $V[5, 7:25]$ para la trama F_{X-1} y $V[6:25]$ para la trama F_X . El dispositivo de codificación de audio también puede especificar que el coeficiente HOA ambiental adicional, asociado al índice 6, debe desaparecer de la reconstrucción de los coeficientes HOA 11' para la trama previa (F_{X-1}), mientras que el coeficiente HOA ambiental adicional, asociado al índice 5, ha de aparecer para la trama actual (F_X) al reconstruir los coeficientes HOA 11'. La transición de los coeficientes HOA ambientales adicionales, asociados al índice 6, provenientes de la reconstrucción en el dispositivo de decodificación de audio durante la trama previa (F_{X-1}), puede reducir la energía total dado que el coeficiente HOA ambiental adicional, asociado al índice 6, representa alguna parte de la energía total del campo de sonido. La reducción de energía puede manifestarse como una distorsión audible del audio.

[89] Del mismo modo, la introducción del coeficiente HOA ambiental adicional, asociado al índice 5, puede, al aparecer durante la trama actual (F_X), dar como resultado una cierta pérdida de energía al reconstruir los coeficientes HOA 11' en el dispositivo de decodificación de audio. La pérdida de energía ocurre porque el coeficiente HOA ambiental adicional, asociado al índice 5, aparece usando, como ejemplo, una operación de fundido inicial lineal que atenúa el coeficiente HOA ambiental adicional, asociado al índice 5, y por lo tanto resta a la energía total. De nuevo, la reducción de energía puede manifestarse como una distorsión de audio.

[90] De acuerdo a diversos aspectos de las técnicas descritas en esta divulgación, la unidad de análisis de campo de sonido 44 puede determinar además cuándo los coeficientes HOA ambientales cambian de trama a trama, y generar un indicador u otro elemento sintáctico indicativo del cambio para el coeficiente HOA ambiental, en términos de ser utilizado para representar los componentes ambientales del campo de sonido (donde el cambio también puede denominarse una "transición" del coeficiente HOA ambiental o como una "transición" del coeficiente HOA ambiental). En particular, la unidad de reducción de coeficientes 46 puede generar el indicador (que puede indicarse como un indicador TransiciónCoefAmb o un indicador TransiciónÍndCoefAmb), proporcionando el indicador a la unidad 42 de generación de flujos de bits para que el indicador pueda incluirse en el flujo de bits 21 (posiblemente como parte de la información de canal lateral).

[91] La unidad de reducción de coeficientes 46 puede, además de especificar el indicador de transición de coeficientes ambientales, también modificar cómo se generan los vectores $V[k]$ en primer plano reducidos 55. En un ejemplo, al determinar que uno de los coeficientes HOA ambientales está en transición durante la trama actual, la unidad de reducción de coeficientes 46 puede especificar un coeficiente vectorial (que también puede denominarse "elemento vectorial" o "elemento") para cada uno de los vectores V de los vectores $V[k]$ en primer plano reducidos 55, que corresponde al coeficiente HOA ambiental en transición. De nuevo, el coeficiente HOA ambiental en transición puede agregar o eliminar del número total BG_{TOT} de coeficientes de trasfondo. Por lo

tanto, el cambio resultante en el número total de coeficientes de trasfondo afecta si el coeficiente HOA ambiental se incluye o no en el flujo de bits, y si el elemento correspondiente de los vectores V se incluye para los vectores V especificados en el flujo de bits en las modalidades de configuración segunda y tercera, descritas anteriormente.

5
 [92] Para ilustrar lo anterior con respecto al ejemplo de las tramas anteriores y actuales (F_{x-1} y F_x), la unidad de reducción de coeficientes 46 puede modificarse a partir de lo especificado en la publicación n° WO 2014/194099 para señalar información redundante en términos de los elementos enviados para el vector V durante las tramas anteriores y actuales (F_{x-1} y F_x). La unidad de reducción de coeficientes 46 puede especificar los elementos vectoriales ($V[5:25]$) para la trama previa F_{x-1} , de modo que el dispositivo de decodificación de audio 24 sea capaz de incorporar el elemento 6 del vector V a la vez que se desvanece el coeficiente HOA ambiental asociado al índice 6. La unidad de reducción de coeficientes 46 puede no especificar ningún elemento sintáctico que indique la transición de los elementos del vector V que están en transición, ya que está implícito a partir de la modalidad de codificación de los vectores V y la información de transición especificada para los coeficientes HOA ambientales. Para la trama actual (F_x), la unidad de reducción de coeficientes 46 puede asimismo especificar el vector V como $V[5:25]$, dado que el dispositivo de decodificación de audio 24 puede utilizar el 5° elemento del vector V en una operación de eliminación para compensar la aparición del coeficiente HOA ambiental asociado al índice 5. La operación de fundido es, en los ejemplos anteriores, complementaria, para el elemento del vector V , con la del coeficiente HOA ambiental, para mantener un nivel de energía uniforme y evitar la introducción de las distorsiones de audio. Si bien se describe como complementario o proporcionando de otra manera una energía uniforme entre transiciones, las técnicas pueden admitir otras formas cualesquiera de operaciones de transición que se usen para evitar o reducir la introducción de distorsiones de audio debido a cambios en la energía.

25
 [93] En otro ejemplo, la unidad de reducción de coeficientes 46 no puede alterar cómo se generan los vectores V de los vectores $V[k]$ en primer plano reducidos 55. Como tal, el indicador de transición se señala en la información de canal lateral. En este ejemplo, el dispositivo de decodificación de audio puede utilizar un vector V de trama anterior o posterior que incluya el coeficiente correspondiente al coeficiente HOA ambiental que está en transición. Este ejemplo puede requerir una funcionalidad adicional en el decodificador (por ejemplo, un mecanismo de anticipación que anticipa las tramas posteriores para copiar el coeficiente de los vectores V de la trama siguiente para su uso en la trama actual cuando se está realizando la transición de un coeficiente HOA ambiental al BG_{TOT}).

35
 [94] A este respecto, las técnicas pueden permitir que el dispositivo de codificación de audio 20 determine cuándo un coeficiente ambisónico ambiental de orden superior 47', que describe un componente ambiental de un campo de sonido, está en transición, en términos de ser utilizado para describir el componente ambiental del campo de sonido. Cuando se hace referencia al componente ambiental del campo de sonido, que se está usando o no, debería entenderse que el dispositivo de codificación de audio 20 puede seleccionar los coeficientes HOA ambientales 47 a usar para reconstruir el campo de sonido en el dispositivo de decodificación de audio 24. Si bien el coeficiente HOA ambiental puede representar algún aspecto del componente de trasfondo o, en otras palabras, del componente ambiental del campo de sonido, el dispositivo de codificación de audio 20 puede determinar que uno o más de los coeficientes HOA ambientales 47 no proporcionan suficiente información relevante para el componente ambiental del campo de sonido, por lo que los bits no se han de utilizar para especificar los uno o más entre los coeficientes HOA ambientales 47 en el flujo de bits 21. El dispositivo de codificación de audio 20 puede identificar algún subconjunto de un conjunto más grande de los coeficientes HOA ambientales 47 que se usan para representar el componente ambiental o aspecto del campo de sonido para cada trama, como ejemplo, para lograr una tasa de bits deseada 41. En cualquier caso, el dispositivo de codificación de audio 20 también puede identificar, en el flujo de bits 21 que incluye el coeficiente ambisónico de orden superior ambiental 47, que el coeficiente ambisónico de orden superior ambiental 47 está en transición.

50
 [95] En estos y otros ejemplos, el dispositivo de codificación de audio 20 puede, al determinar cuándo el coeficiente ambisónico de orden superior ambiental 47' está en transición, determinar que el coeficiente ambisónico de orden superior ambiental 47' no se usa para describir el componente ambiental del campo de sonido. Cuando se identifica que el coeficiente ambisónico ambiental de orden superior 47' está en transición, el dispositivo de codificación de audio 20 puede especificar un indicador TransiciónCoefAmb que indica que el coeficiente ambisónico de orden superior está en transición.

60
 [96] En estos y otros ejemplos, el dispositivo de codificación de audio 20 puede, al determinar cuándo el coeficiente ambisónico de orden superior ambiental 47' está en transición, determinar que el coeficiente ambisónico de orden superior ambiental 47' no se usa para describir el componente ambiental del campo de sonido.

65
 [97] En respuesta a la determinación de que el coeficiente ambisónico de orden superior ambiental 47' no se va a usar, el dispositivo de codificación de audio 20 puede generar una señal basada en vectores, representativa de uno o más componentes distintos del campo de sonido que incluye un elemento de un vector

(por ejemplo, los vectores $V[k]$ en primer plano reducidos 55_o o, en otras palabras, los vectores en primer plano reducidos 55_k), correspondientes al coeficiente ambisónico de orden superior ambiental $47'$. El vector 55_k puede describir aspectos espaciales de un componente distinto del campo de sonido. El vector 55_k también puede haber sido descompuesto a partir de coeficientes ambisónicos de orden superior 11 , descriptivos del campo de sonido de la manera descrita anteriormente.

[98] En estos y otros ejemplos, el dispositivo de codificación de audio 20 puede, cuando determina cuando el coeficiente ambisónico de orden superior $47'$ está en transición, determinar que los coeficientes ambisónicos ambientales de orden superior $47'$ se usen para describir el componente ambiental del campo de sonido.

[99] En estos y otros ejemplos, el dispositivo de codificación de audio 20 puede, al determinar cuándo el coeficiente ambisónico de orden superior ambiental $47'$ está en transición, determinar que el coeficiente ambisónico de orden superior ambiental $47'$ se usa para describir el componente ambiental del campo de sonido. El dispositivo de codificación de audio 20 puede, al identificar que el coeficiente ambisónico de orden superior $47'$ está en transición, especificar también un elemento sintáctico que indica que el coeficiente ambisónico de orden superior $47'$ está en transición.

[100] En estos y otros ejemplos, el dispositivo de codificación de audio 20 puede, al determinar cuándo el coeficiente ambisónico de orden superior ambiental $47'$ está en transición, determinar que el coeficiente ambisónico de orden superior ambiental $47'$ se usa para describir el componente ambiental del campo de sonido. El dispositivo de codificación de audio 20 puede, en respuesta a la determinación de que se va a utilizar el coeficiente ambisónico de orden superior ambiental $47'$, generar una señal basada en vectores, representativa de uno o más componentes distintos del campo de sonido, que incluye un elemento de un vector 55_k , correspondiente al coeficiente ambisónico de orden superior ambiental $47'$. El vector 55_k puede describir aspectos espaciales de un componente distinto del campo de sonido y puede haber sido descompuesto a partir de coeficientes ambisónicos de orden superior descriptivos del campo de sonido.

[101] En algunos ejemplos, la unidad de generación de flujos de bits 42 genera los flujos de bits 21 para incluir tramas de reproducción inmediata (IPF) para, por ejemplo, compensar el retraso del arranque del decodificador. En algunos casos, el flujo de bits 21 puede emplearse junto con normas de transmisión por flujo de Internet, tales como el Flujo de Transmisión Adaptativo Dinámico sobre HTTP (DASH) o la Entrega de Archivos sobre Transporte Unidireccional (FLUTE). El DASH se describe en el documento ISO / IEC 23009-1, "Information Technology - Dynamic adaptive streaming over HTTP (DASH)" ["Tecnología de la información - Flujo de transmisión adaptativo dinámico por HTTP (DASH)"], abril de 2012. La FLUTE se describe en el documento IETF RFC 6726, "FLUTE - File Delivery over Unidirectional Transport" ["FLUTE - Entrega de archivos sobre transporte unidireccional"], noviembre de 2012. Las normas de transmisión por flujo de Internet, tales como FLUTE y DASH, mencionados anteriormente, compensan la pérdida / degradación de tramas y se adaptan al ancho de banda del enlace de transporte de red, permitiendo la reproducción instantánea en los puntos de acceso al flujo (SAP) designados, así como conmutando la reproducción entre representaciones del flujo que difieren en la tasa de bits y / o las herramientas habilitadas en cualquier SAP del flujo. En otras palabras, el dispositivo de codificación de audio 20 puede codificar tramas de tal manera como para cambiar desde una primera representación de contenido (por ejemplo, especificada en una primera tasa de bits) a una segunda representación diferente del contenido (por ejemplo, especificada en una segunda tasa de bits, superior o inferior). El dispositivo de decodificación de audio 24 puede recibir la trama y decodificar independientemente la trama para conmutar desde la primera representación del contenido a la segunda representación del contenido. El dispositivo de decodificación de audio 24 puede seguir decodificando la trama posterior para obtener la segunda representación del contenido.

[102] En el caso de reproducción / conmutación instantánea, el pre-rodaje para una trama de flujo no se ha decodificado para establecer el estado interno requerido para decodificar correctamente la trama, la unidad de generación de flujos de bits 42 puede codificar el flujo de bits 21 para incluir las Tramas de Reproducción Inmediata (IPF), como se describe a continuación con más detalle con respecto a la FIG. 71.

[103] La FIG. 4 es un diagrama de bloques que ilustra el dispositivo de decodificación de audio 24 de la FIG. 2 en más detalle. Como se muestra en el ejemplo de la FIG. 4, el dispositivo de decodificación de audio 24 puede incluir una unidad de extracción 72 , una unidad de reconstrucción basada en la direccionalidad 90 y una unidad de reconstrucción basada en vectores 92 . Aunque se describe a continuación, en la publicación de solicitud de patente internacional número WO 2014/194099, titulada "INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD" ["INTERPOLACIÓN PARA REPRESENTACIONES DESCOMPUESTAS DE UN CAMPO DE SONIDO"] presentada el 29 de mayo de 2014, está disponible más información sobre el dispositivo de decodificación de audio 24 y los diversos aspectos de descompresión, o decodificación de otro modo, de los coeficientes HOA.

[104] La unidad de extracción 72 puede representar una unidad configurada para recibir el flujo de bits 21 y extraer las diversas versiones codificadas (por ejemplo, una versión codificada basada en la direccionalidad o

una versión codificada basada en vector) de los coeficientes HOA 11. La unidad de extracción 72 puede determinar, a partir del elemento sintáctico indicado anteriormente (por ejemplo, el elemento sintáctico TipoCanal 269 mostrado en los ejemplos de las FIGs. 7D y 7E) si los coeficientes HOA 11 fueron codificados mediante las diversas versiones. Cuando se había realizado una codificación basada en la dirección, la unidad de extracción 72 puede extraer la versión basada en la dirección de los coeficientes HOA 11 y los elementos sintácticos asociados a la versión codificada (que se indica como información basada en dirección 91 en el ejemplo de la FIG. 4), pasando la información basada en la dirección 91 a la unidad de reconstrucción basada en la dirección 90. La unidad de reconstrucción basada en la dirección 90 puede representar una unidad configurada para reconstruir los coeficientes HOA en forma de coeficientes HOA 11', basándose en la información basada en la dirección 91. El flujo de bits y la disposición de elementos sintácticos dentro del flujo de bits se describen a continuación con más detalle con respecto al ejemplo de las FIGs. 7A a 7J.

[105] Cuando el elemento sintáctico indica que los coeficientes HOA 11 fueron codificados utilizando una síntesis basada en vectores, la unidad de extracción 72 puede extraer los vectores $V[k]$ codificados de primer plano 57, los coeficientes HOA ambientales codificados 59 y las nFG señales codificadas 61. La unidad de extracción 72 puede pasar los vectores codificados $V[k]$ de primer plano 57 a la unidad de descuantización 74 y los coeficientes HOA ambientales codificados 59, junto con las nFG señales codificadas 61, a la unidad de decodificación psicoacústica 80.

[106] Para extraer los vectores de primer plano codificados $V[k]$ 57, los coeficientes HOA ambientales codificados 59 y las nFG señales codificadas 61, la unidad de extracción 72 puede obtener los vectores de primer plano codificados $V[k]$ 57 (que también pueden ser mencionados como la información de canal lateral 57). La información de canal lateral 57 puede incluir el elemento sintáctico indicado como longitudVectVCodificado. La unidad de extracción 72 puede analizar sintácticamente la longitudVectVCodificado a partir de la información de canal lateral 57. La unidad de extracción 72 puede estar configurada para funcionar en una cualquiera de las modalidades de configuración descritas anteriormente, basándose en el elemento sintáctico longitudVectVCodificado.

[107] La unidad de extracción 72 funciona entonces de acuerdo a una cualquiera de las modalidades de configuración para analizar sintácticamente una forma comprimida de los vectores $V[k]$ de primer plano reducidos 55_k a partir de la información de canal lateral 57. Como se ha indicado anteriormente con respecto a la unidad de generación de flujos de bits 42 del dispositivo de codificación de audio 20 mostrado en el ejemplo de la FIG. 4, un indicador u otro elemento sintáctico puede especificarse en el flujo de bits, indicativo de una transición en los coeficientes HOA ambientales 47 de trama a trama o, posiblemente, de multi-trama a multi-trama. La unidad de extracción 72 puede analizar sintácticamente el elemento sintáctico que indica si un coeficiente HOA ambiental está en transición. Como se muestra adicionalmente en el ejemplo de la FIG. 4, la unidad de extracción 72 puede incluir una unidad de descompresión V 755 (que se muestra como "unidad de descomposición V 755" en el ejemplo de la FIG. 4). La unidad de descompresión V 755 recibe la información de canal lateral del flujo de bits 21 y el elemento sintáctico indicado como longitudVectVCodificado. La unidad de extracción 72 puede analizar sintácticamente el elemento sintáctico longitudVectVCodificado del flujo de bits 21 (y, por ejemplo, de la cabecera de unidad de acceso incluida dentro del flujo de bits 21). La unidad de descompresión V 755 incluye una unidad de configuración de modalidad 756 ("unidad de config. de modalidad 756") y una unidad de análisis sintáctico 758 configurable para funcionar de acuerdo a una cualquiera de las modalidades de configuración 760.

[108] La unidad de extracción 72 puede proporcionar el elemento sintáctico longitudVectVCodificado a la unidad de configuración de modalidad 756. La unidad de extracción 42 también puede extraer un valor para las variables de estado utilizables por la unidad de análisis sintáctico 758.

[109] La unidad de configuración de modalidad 756 puede seleccionar una modalidad de análisis sintáctico 760 basándose en el elemento sintáctico indicativo de una transición de un coeficiente HOA ambiental. Las modalidades de análisis sintáctico 760 pueden, en este ejemplo, especificar ciertos valores para configurar la unidad de análisis sintáctico 758. Los valores adicionales pueden referirse a los valores de las variables denominadas "ModalidadTransiciónCoefAmb" y "CoefAmbFueIncorporado". Los valores mantienen estado con respecto al estado de transición de CanallInfoHoaAmbAdic, como se especifica en la siguiente tabla:

Sintaxis de CanallInfoHoaAmbAdic (i) (A continuación)

Sintaxis de CanallInfoHoaAmbAdic (i)

Sintaxis	Nº de bits	Mnemónico
CanallInfoHOAAmbAdic (i)		
{		
si (IndicadorIndependencia) {		
CoefAmbFueIncorporado [i];	1	Bslbf
TransiciónCoefAmb;	1	Bslbf
ÍndCoefAmb [i] = ÍndCoefAmbCodificado + 1 +	AsignarAmb mBits	Uimsbf
MínNúmDeCoefsParaAmbHOA;		
}		
en caso contrario {		
si (TransiciónCoefAmb) {	1	Bslbf
si (CoefAmbFueIncorporado [i] == 0) {		
ModalidadTransiciónCoefAmb [i] = 1;		
CoefAmbFueIncorporado [i] = 1;		
ÍndCoefAmb [i] = ÍndCoefAmbCodificado + 1 +	AsignarAmb mBits	Uimsbf
MínNúmDeCoefsParaAmbHOA;		
}		
en caso contrario {		
ModalidadTransiciónCoefAmb [i] = 2;		
CoefAmbFueIncorporado [i] = 0;		
}		
}		
en caso contrario {		
ModalidadTransiciónCoefAmb [i] = 0;		
}		
}		
}		
NOTA:		
El ÍndCoefAmbCodificado de la trama anterior se usa en las siguientes condiciones		
si (TransiciónCoefAmb && CoefAmbFueIncorporado [i])		
si (TransiciónCoefAmb == 0)		
La variable CoefAmbFueIncorporado es un alternador e indica si este canal HOA adicional ya se ha incorporado o no. Si CoefAmbFueIncorporado == 1, se debería entender que la siguiente transición es una eliminación en el ejemplo anterior.		
ModalidadTransiciónCoefAmb:		
0: Sin transición (Coeficiente HOA Ambiental Adicional continuo)		
1: Incorporación del Coeficiente HOA Ambiental Adicional		
2: Eliminación del Coeficiente HOA Ambiental adicional		

- 5
- 10
- 15
- 20
- 25
- [110]** En la Tabla CanallInfoHoaAmbAdic anterior, la unidad de configuración de modalidad 756 puede determinar si el valor de IndicadorIndependencia para una trama HOA es verdadero. Un IndicadorIndependencia con un valor verdadero indica que la trama HOA es una trama de reproducción inmediata (IPF).
- [111]** Si el valor de IndicadorIndependencia para la trama HOA es falso, la unidad de configuración de modalidad 756 determina si el indicador TransiciónCoefAmb está fijado en uno. El indicador TransiciónCoefAmb puede representar un bit indicativo de una transición de un coeficiente ambisónico ambiental de orden superior. Aunque se describe como un bit, el indicador TransiciónCoefAmb puede, en algunos ejemplos, incluir uno o más bits. El término "bit", como se usa en el presente documento, ha de entenderse que se refiere a uno o más bits y no debería limitarse a un solo bit, a menos que se indique explícitamente lo contrario.
- [112]** Cuando el indicador TransiciónCoefAmb se fija en uno, la unidad de configuración de modalidad 756 determina si otra variable (o, en otras palabras, elemento sintáctico), CoefAmbFueIncorporado [i], es igual a cero. La variable CoefAmbFueIncorporado [i] es una formación de *i* elementos, uno para cada uno de los CanalesInfoHOAAmbAdic, que indica si el *i*-ésimo CanallInfoHOAAmbAdic se incorporó previamente. Cuando el *i*-ésimo CanallInfoHOAAmbAdic no fue incorporado previamente (lo que significa que el *i*-ésimo CanallInfoHOAAmbAdic es igual a cero), la unidad de configuración de modalidad 756 puede fijar la ModalidadTransiciónCoefAmb para el *i*-ésimo CanallInfoHOAAmbAdic en uno y al mismo tiempo fijar CoefAmbFueIncorporado para el *i*-ésimo CanallInfoHOAAmbAdic en uno. Cuando el *i*-ésimo

CanallInfoHOAAmbAdic fue previamente incorporado (lo que significa que el *i*-ésimo CanallInfoHOAAmbAdic no es igual a cero), la unidad de configuración de modalidad 756 puede fijar la ModalidadTransiciónCoefAmb para el *i*-ésimo CanallInfoHOAAmbAdic en dos y fijar el CoefAmbFueIncorporado para el *i*-ésimo CanallInfoHOAAmbAdic en cero.

5 **[113]** La combinación de los elementos sintácticos CoefAmbFueIncorporado y ModalidadTransiciónCoefAmb puede representar información de estado de transición. La información de estado de transición puede, dado que cada uno de los elementos sintácticos CoefAmbFueIncorporado y ModalidadTransiciónCoefAmb tiene un solo bit, definir hasta cuatro estados. La anterior tabla ejemplar de sintaxis indica que la información de estado de transición indica uno de tres estados. Los tres estados pueden incluir un estado sin transición, un estado de fundido inicial y un estado de fundido final. Aunque se describe en esta divulgación como incluyendo dos bits para indicar uno de tres estados, la información de estado de transición puede ser un solo bit cuando la información de estado de transición indica menos de tres estados. Además, la información de estado de transición puede incluir más de dos bits en ejemplos donde la información de estado de transición indica uno entre cinco o más estados.

10 **[114]** Cuando el indicador TransiciónCoefAmb es igual a cero, la unidad de configuración de modalidad 756 puede configurar la ModalidadTransiciónCoefAmb para el *i*-ésimo CanallInfoHOAAmbAdic en cero. Como se señala en la Tabla anterior, cuando la ModalidadTransiciónCoefAmb es igual a los siguientes valores, se puede realizar la acción correspondiente indicada a continuación:

- 0: Sin transición (coeficiente HOA Ambiental Adicional continuo);
- 1: Incorporación del Coeficiente HOA Ambiental adicional; y
- 2: Eliminación del Coeficiente HOA Ambiental adicional.

20 **[115]** Si el valor de IndicadorIndependencia para la trama HOA es verdadero, la unidad de extracción 72 puede extraer información de transición 757 para el canal HOA ambiental adicional a partir de una estructura sintáctica asociada dentro del flujo de bits 21. Debido a que las IPF son, por definición, decodificables independientemente, la información de transición 757 para la IPF puede proporcionarse junto con la IPF en el flujo de bits, por ejemplo, tal como la información de estado 814 descrita anteriormente. Por lo tanto, la unidad de extracción 72 puede extraer el valor para la variable CoefAmbFueIncorporado [i] para el *i*-ésimo CanallInfoHOAAmbAdic para el que la estructura sintáctica está proporcionando información de transición 757. De esta manera, la unidad de configuración de modalidad 756 puede determinar las modalidades 760 para que el *i*-ésimo CanallInfoHOAAmbAdic sea aplicado por el dispositivo de decodificación de audio 24 en el *i*-ésimo CanallInfoHOAAmbAdic.

30 **[116]** Sin embargo, la sintaxis anterior puede modificarse ligeramente para reemplazar los elementos sintácticos independientes de CoefAmbFueIncorporado [i] y TransiciónCoefAmb con un elemento sintáctico EstadoTransiciónCoefAmb [i] de dos bits y un elemento sintáctico TransiciónIndCoefAmb de un bit. Por lo tanto, la tabla de sintaxis anterior puede reemplazarse por la siguiente tabla de sintaxis:

Sintaxis de CanallInfoHoaAmbAdic (i)

Sintaxis	Nº de bits	Mnemónico
CanallInfoHOAAmbAdic (i)		
{		
si (IndicadorIndependenciaHoa) {		
EstadoTransiciónCoefAmb [i];	2	Uimsbf
ÍndCoefAmb [i] = ÍndCoefAmbCodificado + 1 +	AsignarAmb mBits	Uimsbf
MínNúmDeCoefsParaAmbHOA;		
Idx}		
en caso contrario {		
si (TransiciónIndCoefAmb == 1) {	1	Bslbf
si (EstadoTransiciónCoefAmb [i] > 1) {		
EstadoTransiciónCoefAmb [i] = 1;		
ÍndCoefAmb [i] = ÍndCoefAmbCodificado + 1 +	AsignarAmb mBits	Uimsbf
MínNúmDeCoefsParaAmbHOA;		
}		
en caso contrario {		
EstadoTransiciónCoefAmb [i] = 2;		
}		
}		
en caso contrario {		
EstadoTransiciónCoefAmb [i] = 0;		
}		
}		

Sintaxis	Nº de bits	Mnemónico
}		
<p>NOTA:</p> <p>El ÍndCoefAmb de la trama anterior se usa en las siguientes condiciones ejemplares</p> <p> si (EstadoTransiciónÍndCoefAmb == 0)</p> <p> si (EstadoTransiciónÍndCoefAmb == 2)</p> <p>EstadoTransiciónCoefAmb:</p> <p> 0: Sin transición (Coeficiente HOA Ambiental Adicional continuo)</p> <p> 1: Incorporación del Coeficiente HOA Ambiental Adicional</p> <p> 2: Eliminación del Coeficiente HOA Ambiental adicional</p> <p> 3: Valor inicial</p>		

[117] En la anterior tabla ejemplar de sintaxis, el dispositivo de codificación de audio 20 señala explícitamente el elemento sintáctico EstadoTransiciónCoefAmb cuando el elemento sintáctico IndicadorIndependenciaHOA se fija en un valor de uno. Cuando se señala el elemento sintáctico EstadoTransiciónCoefAmb, el dispositivo de codificación de audio 20 señala el estado actual del correspondiente coeficiente HOA ambiental. De lo contrario, cuando el elemento sintáctico IndicadorIndependenciaHOA se fija en un valor de cero, el dispositivo de codificación de audio 20 no señala el EstadoTransiciónCoefAmb, sino que, en cambio, señala el elemento sintáctico TransiciónÍndCoefAmb, que indica si hay una transición en el correspondiente coeficiente HOA ambiental.

[118] Cuando el elemento sintáctico IndicadorIndependenciaHOA se fija en un valor de cero, la unidad de extracción 72 puede mantener el EstadoTransiciónCoefAmb para el coeficiente correspondiente entre los coeficientes HOA ambientales. La unidad de extracción 72 puede actualizar el elemento sintáctico EstadoTransiciónCoefAmb basándose en el TransiciónÍndCoefAmb. Por ejemplo, cuando el elemento sintáctico EstadoTransiciónCoefAmb se fija en 0 (es decir, sin transición) y el elemento sintáctico TransiciónÍndCoefAmb se fija en 0, la unidad de extracción 72 puede determinar que no se ha producido ningún cambio y, por lo tanto, que no es necesario ningún cambio en el elemento sintáctico EstadoTransiciónCoefAmb. Cuando el elemento sintáctico EstadoTransiciónCoefAmb se fija en 0 (es decir, sin transición) y el elemento sintáctico TransiciónÍndCoefAmb se fija en 1, la unidad de extracción 72 puede determinar que el correspondiente coeficiente HOA ambiental ha de eliminarse y fija el elemento sintáctico EstadoTransiciónCoefAmb en un valor de 2. Cuando el elemento sintáctico EstadoTransiciónCoefAmb se fija en 2 (es decir, el coeficiente HOA ambiental correspondiente fue eliminado) y el elemento sintáctico TransiciónÍndCoefAmb se fija en 1, la unidad de extracción 72 puede determinar que el correspondiente coeficiente HOA ambiental ha de incorporarse. y fija el elemento sintáctico EstadoTransiciónCoefAmb en un valor de 1.

[119] De manera similar al indicador TransiciónCoefAmb, el elemento sintáctico TransiciónÍndCoefAmb puede representar un bit indicativo de una transición de un coeficiente ambisónico de orden superior ambiental. Aunque se describe como un bit, el elemento sintáctico TransiciónÍndCoefAmb puede, en algunos ejemplos, incluir uno o más bits. De nuevo, el término "bit", tal como se usa en el presente documento, debería entenderse como que se refiere a uno o más bits y no debería limitarse a solo un único bit, a menos que se indique explícitamente lo contrario.

[120] Además, el elemento sintáctico EstadoTransiciónCoefAmb [i] puede representar información de estado de transición. La información de estado de transición puede, dado que el elemento sintáctico EstadoTransiciónCoefAmb [i] tiene dos bits, indicar uno de los cuatro estados. La anterior tabla ejemplar de sintaxis indica que la información de estado de transición indica uno de tres estados. Los tres estados pueden incluir un estado sin transición, un estado de fundido inicial y un estado de fundido final. De nuevo, aunque se describe en esta divulgación como incluyendo dos bits para indicar uno de tres estados, la información de estado de transición puede ser un solo bit cuando la información de estado de transición indica menos de tres estados. Además, la información de estado de transición puede incluir más de dos bits en ejemplos donde la información de estado de transición indica uno entre cinco o más estados.

[121] La unidad de extracción 72 también puede funcionar de acuerdo a la instrucción de conmutación que se presenta en el pseudo-código siguiente con la sintaxis que se presenta en la siguiente tabla de sintaxis para DatosVectorV:

```
conmutador longitudVectVCodificado {
    caso 0: // longitud vectorial completa
        LongitudVectV = NúmDeCoefsHoa;
```

```

    para (m=0; m< LongitudVectV; ++m){
        IdentCoefVecV [m] = m;
    }
corte;
5 caso 1: // longitud mínima del vector
    LongitudVectV = NúmDeCoefsHoa - MínNúmDeCoefsParaAmbHOA -
    NúmDeCanalesHoaAdicCont;
    para (i = 0; i < NúmDeCodificadoresAdicionales; ++ i) {
10         si (ModalidadTransiciónCoefAmb [i] == 0) {
            ContÍndCoefAmb [i] = ÍndCoefAmb [i];}
        en caso contrario {ContÍndCoefAmb [i] = -1; }
    }
    para (m=0; m< LongitudVectV; ++m){
15         si (esmiembro (m + MínNúmDeCoefsParaAmbHOA + 1,
            ContÍndCoefAmb) == 0) {
            IdentCoefVecV[m] = m + MínNúmDeCoefsParaAmbHOA;
        }
    }
corte;
20 caso 2: // Se eliminó MínNúmDeCoefsParaAmbHOA (el estado en el software ref de
    RM1)
    LongitudVectV = NúmDeCoefsHoa - MínNúmDeCoefsParaAmbHOA;
    para (m=0; m< LongitudVectV; ++m){
25         IdentCoefVecV[m] = m + MínNúmDeCoefsParaAmbHOA;
    }
corte;
}

```

[122] El caso 0 en el pseudo-código anterior representa un pseudo-código para recuperar todos los elementos del vector V cuando se selecciona la modalidad de codificación. El caso 1 representa un pseudo-código para recuperar el vector V después de haber sido reducido de la manera descrita anteriormente. El caso 1 ocurre cuando se envían tanto los N_{BG} como los adicionales coeficientes HOA ambientales, lo que da como resultado que no se envíen tanto los elementos correspondientes de los vectores V. El caso 2 representa un pseudo-código para recuperar los vectores V cuando se envían los elementos del vector V correspondientes a los coeficientes HOA ambientales adicionales (redundantemente) pero no los elementos del vector V correspondientes a los N_{BG} coeficientes HOA ambientales.

[123] El dispositivo de codificación de audio 20 puede especificar el flujo de bits 21 cuando el dispositivo de decodificación de audio 24 está configurado para funcionar de acuerdo al Caso 2. El dispositivo de codificación de audio 20 puede señalar el Caso 2 al seleccionar señalar explícitamente los elementos del vector V en el flujo de bits 21 durante una transición de un coeficiente HOA ambiental. El dispositivo de codificación de audio 20 puede elegir enviar explícitamente el elemento de vector V redundante para permitir el fundido inicial y el fundido final del elemento de vector V en función de la transición del coeficiente HOA ambiental, como se analiza en más detalle a continuación con respecto a la FIG. 8.

[124] El dispositivo de codificación de audio 20 puede seleccionar el Caso 1 cuando elige configurar el decodificador 24 para realizar una búsqueda anticipada para recuperar los elementos del vector V de una trama posterior en el tiempo (o una búsqueda retroactiva para recuperar los elementos del vector V de una trama anterior en el tiempo). En otras palabras, la unidad de extracción 72 del dispositivo de decodificación de audio 24 puede configurarse para realizar el Caso 1 cuando el dispositivo de codificación de audio 20 elige no enviar el elemento redundante del vector V y en cambio puede configurar la unidad de extracción 72 del dispositivo de decodificación de audio 24 para realizar las operaciones de búsqueda anticipada o búsqueda retroactiva para reutilizar un elemento del vector V de una trama diferente. El dispositivo de decodificación de audio 24 puede entonces realizar la operación de fundido inicial / final usando el elemento del vector V señalado implícitamente (que puede referirse al elemento del vector V reutilizado a partir de una trama anterior o posterior).

[125] La unidad de configuración de modalidad 756 puede seleccionar una de las modalidades 760 que configura la manera adecuada mediante la cual analizar sintácticamente el flujo de bits 21 para recuperar los vectores V [k] codificados en primer plano 57. La unidad de configuración de modalidad 756 puede configurar la unidad de análisis sintáctico 758 con la seleccionada entre las modalidades 760, que a continuación puede analizar sintácticamente el flujo de bits 21 para recuperar el vector codificado en primer plano V [k] 57. La unidad de análisis sintáctico 758 puede emitir luego los vectores V [k] en primer plano codificados 57.

Sintaxis de DatosVectorV (i) (A continuación)

5

Sintaxis de DatosVectorV (i)

Sintaxis	Nº de bits	Mnemónico
<pre> DatosVectorV(i) { si (NbitsQ(k)[i] == 5){ para (m=0; m< LongitudVectV; ++m){ VecV[i][IdentCoefVecV[m]](k) = (ValVec //128,0)- } } en caso contrario, si (NbitsQ(k)[i] >= 6){ para (m=0; m< LongitudVectV; ++m){ indHuff = SeleccionHuff (IdCoefVecV [m], IndicadorP [i], IndicadorCb [i]); cid = / DecodHuff (NbitsQ [i], indHuff, valHuff); valA [i] [m] = 0.0; si (cid > 0){ valA[i][m] = sgn = (valSgn * 2)-1; si (cid > 1){ valA[i][m] = sgn * (2,0^(cid-1) + valAdicEnt); } } VcV [i] [IdentCoefVecV [m]] (k) = valA [i] [m] * (2 ^ (16-NbitsQ (k) [i]) * valA [i] [m]) / 2 ^ 15; si (IndicadorP(k)[i] ==1){ VecV[i][IdCoefVecV[m]](k)+= VecV[i][IdCoefVecV[m]](k-1) } } } } </pre>	<p>1,0: 8</p> <p>dinámico</p> <p>1</p> <p>cid-1</p>	<p>uimsbf</p> <p>decodHuff</p> <p>bslbf</p> <p>uimsbf</p>

10 **[126]** Después de esta instrucción de conmutación, la decisión de realizar una descuantización uniforme puede ser controlada por el elemento sintáctico NbitsQ (o, como se ha indicado anteriormente, el elemento sintáctico *nbits*) que, cuando es igual a 5, se realiza una descuantización escalar uniforme de 8 bits. Por el contrario, un valor de NbitsQ mayor o igual a 6 puede dar como resultado la aplicación de la decodificación de Huffman. El valor *cid* mencionado anteriormente puede ser igual a los dos bits menos significativos del valor NbitsQ. La modalidad de predicción analizada anteriormente se indica como IndicadorP en la tabla de sintaxis anterior, mientras que el bit de información HT se indica como IndicadorCb en la tabla de sintaxis anterior. La sintaxis restante especifica cómo ocurre la decodificación de una manera esencialmente similar a la descrita anteriormente.

20 **[127]** La unidad de reconstrucción basada en vectores 92 representa una unidad configurada para realizar operaciones recíprocas a las descritas anteriormente con respecto a la unidad de descomposición basada en vectores 27, según se ilustra en la FIG. 3 para reconstruir los coeficientes HOA 11'. La unidad de reconstrucción basada en vectores 92 puede incluir una unidad de descuantización 74, una unidad de interpolación espacio-temporal 76, una unidad de formulación de primer plano 78, una unidad de decodificación psicoacústica 80, una unidad de fundido 770 y una unidad de formulación de coeficientes HOA 82.

25 **[128]** La unidad de descuantización 74 puede representar una unidad configurada para funcionar de manera recíproca a la unidad de cuantización 52 que se muestra en el ejemplo de la FIG. 3, descuantizando los vectores $V[k]$ de primer plano codificados 57 para generar vectores $V[k]$ de primer plano reducidos 55_k . La unidad de descuantización 74 puede, en algunos ejemplos, realizar una forma de decodificación por entropía y de descuantización escalar de una manera recíproca a la descrita anteriormente con respecto a la unidad de cuantización 52. La unidad de descuantización 74 puede remitir los vectores $V[k]$ de primer plano reducidos 55_k a la unidad de interpolación espacio-temporal 76.

30 **[129]** La unidad de decodificación psicoacústica 80 puede funcionar de manera recíproca a la unidad codificadora de audio psicoacústica 40 que se muestra en el ejemplo de la FIG. 3 para decodificar los coeficientes HOA ambientales codificados 59 y las nFG señales codificadas 61 y, de este modo, generar coeficientes HOA ambientales compensados en energía 47' y las nFG señales interpoladas 49' (que también se pueden denominar nFG objetos de audio interpolados 49'). La unidad de decodificación psicoacústica 80 puede pasar los coeficientes HOA ambientales compensados en energía 47' a la unidad de fundido 770 y las nFG señales 49' a la unidad de formulación de primer plano 78.

40

[130] La unidad de interpolación espacio-temporal 76 puede funcionar de una manera similar a la descrita anteriormente con respecto a la unidad de interpolación espacio-temporal 50. La unidad de interpolación espacio-temporal 76 puede recibir los vectores $V[k]$ reducidos de primer plano 55_k y realizar la interpolación espacio-temporal con respecto a los vectores $V[k]$ de primer plano 55_k y los vectores $V[k-1]$ reducidos de primer plano 55_{k-1} para generar vectores $V[k]$ de primer plano interpolados 55_k . La unidad de interpolación espacio-temporal 76 puede remitir los vectores $V[k]$ de primer plano interpolados 55_k a la unidad de fundido 770.

[131] La unidad de extracción 72 también puede emitir una señal 757 indicativa de cuando uno de los coeficientes HOA ambientales está en transición hacia la unidad de fundido 770, que puede entonces determinar cuál de los SHC_{BG} 47' (donde los SHC_{BG} 47' también se pueden indicar como "canales de HOA ambientales 47'" o "coeficientes HOA ambientales 47'") y los elementos de los vectores $V[k]$ de primer plano interpolados 55_k han de incorporarse o eliminarse. En algunos ejemplos, la unidad de fundido 770 puede funcionar opuesta con respecto a cada uno de los coeficientes HOA ambientales 47' y los elementos de los vectores $V[k]$ en primer plano interpolados 55_k . Es decir, la unidad de fundido 770 puede realizar un fundido inicial o un fundido final, o tanto un fundido inicial como un fundido final, con respecto al correspondiente coeficiente HOA ambiental 47', mientras realiza un fundido inicial o un fundido final, o tanto un fundido inicial como un fundido final, con respecto al correspondiente de los elementos de los vectores $V[k]$ en primer plano interpolados 55_k . La unidad de fundido 770 puede emitir coeficientes HOA ambientales ajustados 47" a la unidad de formulación de coeficientes HOA 82 y los vectores ajustados $V[k]$ en primer plano 55_k a la unidad de formulación de primer plano 78. A este respecto, la unidad de fundido 770 representa una unidad configurada para realizar una operación de fundido con respecto a diversos aspectos de los coeficientes HOA o sus derivados, por ejemplo, en forma de los coeficientes HOA ambientales 47' y los elementos de los vectores $V[\lambda]$ interpolados de primer plano 55_k .

[132] En otras palabras, el elemento Beck asociado a un coeficiente HOA transmitido adicionalmente puede no tener que ser transmitido. Para las tramas en las que un coeficiente HOA adicional es transitorio (lo que significa incorporado o eliminado), el elemento $VecV$ se transmite para evitar los agujeros de energía en el campo de sonido HOA reconstruido.

[133] En estos y otros ejemplos, el dispositivo de decodificación de audio 24 puede, al determinar cuándo un coeficiente ambisónico ambiental de orden superior (tal como el coeficiente ambisónico ambiental de orden superior 47') está en transición, obtener un indicador $TransiciónCoefAmb$ desde un flujo de bits (tal como el flujo de bits 21 en el ejemplo de la FIG. 4) que también incluye el coeficiente ambisónico ambiental de orden superior 47'. El indicador $TransiciónCoefAmb$ indica que el coeficiente ambisónico de orden superior está en transición.

[134] En estos y otros ejemplos, el dispositivo de decodificación de audio 24 puede, al determinar cuándo el coeficiente ambisónico ambiental de orden superior 47' está en transición, determinar que el coeficiente ambisónico ambiental de orden superior 47' no se usa para describir el componente ambiental del campo de sonido. En respuesta a la determinación de que no se usa el coeficiente ambisónico de orden superior ambiental 47', el dispositivo de decodificación de audio 24 puede obtener una señal basada en vectores, representativa de uno o más componentes distintos del campo de sonido que incluye un elemento de un vector correspondiente al coeficiente ambisónico ambiental de orden superior 47'. El vector puede referirse a uno de los vectores $V[k]$ en primer plano reducidos 55_k , y como tal se puede denominar vector 55_k . El vector 55_k puede describir aspectos espaciales de un componente distinto del campo de sonido y puede haber sido descompuesto a partir de coeficientes ambisónicos de orden superior 11, descriptivos del campo de sonido. El dispositivo de decodificación de audio 24 puede realizar además una operación de fundido inicial con respecto al elemento del vector 55_k correspondiente al coeficiente ambisónico ambiental de orden superior 47' para incorporar el elemento del vector. El dispositivo de decodificación de audio 24 puede realizar la operación de fundido inicial para añadir el elemento del vector 55_k aumentando linealmente una ganancia del elemento del vector 55_k durante la trama, como se describe con más detalle con respecto al ejemplo de la FIG. 8.

[135] En estos y otros ejemplos, el dispositivo de decodificación de audio 24 puede, al determinar cuándo el coeficiente ambisónico ambiental de orden superior 47' está en transición, determinar que el coeficiente ambisónico ambiental de orden superior 47' no se usa para describir el componente ambiental del campo de sonido. En respuesta a la determinación de que no se usan los coeficientes ambisónicos ambientales de orden superior, el dispositivo de decodificación de audio 24 puede obtener una señal basada en vectores, representativa de uno o más componentes distintos del campo de sonido que incluye un elemento de un vector 55_k correspondiente al coeficiente ambisónico ambiental de orden superior 47'. El vector 55_k puede, como se ha indicado anteriormente, describir aspectos espaciales de un componente distinto del campo de sonido y haber sido descompuesto a partir de coeficientes ambisónicos de orden superior 11, descriptivos del campo de sonido. El dispositivo de decodificación de audio 24 también puede realizar una operación de fundido inicial con respecto al elemento del vector 55_k correspondiente al coeficiente ambisónico de orden superior ambiental 47' para incorporar el elemento del vector 55_k . El dispositivo de decodificación de audio 24 puede realizar además una operación de fundido final con respecto al coeficiente ambisónico de orden superior ambiental 47' para eliminar el coeficiente ambisónico ambiental de orden superior 47'.

- 5 **[136]** En estos y otros ejemplos, el dispositivo de decodificación de audio 24 puede, al determinar cuándo el coeficiente ambisónico ambiental de orden superior 47' está en transición, determinar que el coeficiente ambisónico de orden superior ambiental se usa para describir el componente ambiental del campo de sonido. En respuesta a la determinación de que se va a utilizar el coeficiente ambisónico de orden superior ambiental, el dispositivo de decodificación de audio 24 puede obtener una señal basada en vectores, representativa de uno o más componentes distintos del campo de sonido, que incluye un elemento de un vector 55_k correspondiente al coeficiente ambisónico ambiental de orden superior 47'. Nuevamente, el vector 55_k puede describir aspectos espaciales de un componente distinto del campo de sonido y haber sido descompuesto a partir de coeficientes ambisónicos de orden superior 11, descriptivos del campo de sonido. El dispositivo de decodificación de audio 24 puede realizar una operación de fundido final con respecto al elemento del vector 55_k correspondiente al coeficiente ambisónico de orden superior ambiental 47' para eliminar el elemento del vector.
- 10
- 15 **[137]** En estos y otros ejemplos, el dispositivo de decodificación de audio 24 puede, al determinar cuándo el coeficiente ambisónico ambiental de orden superior 47' está en transición, determinar que el coeficiente ambisónico ambiental de orden superior 47' se usa para describir el componente ambiental del campo de sonido. En respuesta a la determinación de que se usa el coeficiente ambisónico de orden superior ambiental 47', el dispositivo de decodificación de audio 24 puede obtener una señal basada en vectores, representativa de uno o más componentes distintos del campo de sonido, que incluye un elemento de un vector 55_k . correspondiente al coeficiente ambisónico ambiental de orden superior. El vector 55_k puede, de nuevo, describir aspectos espaciales de un componente distinto del campo de sonido y haber sido descompuesto a partir de coeficientes ambisónicos de orden superior, descriptivos del campo de sonido. El dispositivo de decodificación de audio 24 también puede realizar una operación de fundido final con respecto al elemento del vector 55_k correspondiente al coeficiente ambisónico de orden superior ambiental 47' para eliminar el elemento del vector 55_k . El dispositivo de decodificación de audio 24 puede realizar además una operación de fundido inicial con respecto al canal ambisónico ambiental de orden superior 47' para incorporar el canal ambisónico ambiental de orden superior 47'.
- 20
- 25
- 30 **[138]** En estos y otros ejemplos, el dispositivo de decodificación de audio 24 puede, al obtener la señal basada en vectores que incluye el elemento del vector 55_k correspondiente al coeficiente ambisónico ambiental de orden superior 47', determinar el elemento del vector 55_k a partir de la trama actual, una trama posterior a la trama actual o una trama anterior a la trama actual, en la que se realiza la operación de fundido con respecto al elemento del vector 55_k .
- 35
- [139]** En estos y otros ejemplos, el dispositivo de decodificación de audio 24 puede obtener un objeto de audio correspondiente al vector 55_k , y generar un objeto de audio ajustado espacialmente como una función del objeto de audio y del vector 55_k . El objeto de audio puede referirse a uno de los objetos de audio 49', que también pueden denominarse las nFG señales interpoladas 49'.
- 40
- [140]** La unidad de formulación del primer plano 78 puede representar una unidad configurada para realizar la multiplicación matricial con respecto a los vectores $V[k]$ ajustados de primer plano 55_k y las nFG señales interpoladas 49' para generar los coeficientes HOA de primer plano 65. La unidad de formulación de primer plano 78 puede realizar una multiplicación matricial de las nFG señales interpoladas 49' por los vectores $V[k]$ ajustados de primer plano 55_k .
- 45
- [141]** La unidad de formulación de coeficientes HOA 82 puede representar una unidad configurada para combinar los coeficientes HOA de primer plano 65 con los canales HOA ambientales 47', para obtener los coeficientes HOA 11', donde la notación de prima ' refleja que los coeficientes HOA 11' pueden ser similares a, pero no los mismos que, los coeficientes HOA 11. Las diferencias entre los coeficientes HOA 11 y 11' pueden resultar de pérdidas debidas a la transmisión por un medio de transmisión con pérdidas, cuantización u otras operaciones con pérdidas.
- 50
- [142]** La FIG. 5A es un diagrama de flujo que ilustra una operación ejemplar de un dispositivo de codificación de audio, tal como el dispositivo de codificación de audio 20 que se muestra en la FIG. 3, en la realización de diversos aspectos de las técnicas de síntesis basadas en vectores, descritas en esta divulgación. Inicialmente, el dispositivo de codificación de audio 20 recibe los coeficientes HOA 11 (106). El dispositivo de codificación de audio 20 puede invocar la unidad de LIT 30, que puede aplicar una LIT con respecto a los coeficientes HOA para emitir los coeficientes HOA transformados (por ejemplo, en el caso de la SVD, los coeficientes HOA transformados pueden comprender los vectores $US[k]$ 33 y los vectores $V[k]$ 35 (107).
- 55
- 60
- [143]** El dispositivo de codificación de audio 20 puede invocar, a continuación, la unidad de cálculo de parámetros 32 para realizar el análisis descrito anteriormente con respecto a cualquier combinación de los vectores $US[k]$ 33, los vectores $US[k-1]$ 33, los vectores $V[k]$ y/o los vectores $V[k-1]$ 35, para identificar varios parámetros de la manera descrita anteriormente. Es decir, la unidad de cálculo de parámetros 32 puede
- 65

determinar al menos un parámetro basándose en un análisis de los coeficientes HOA transformados 33/35 (108).

[144] El dispositivo de codificación de audio 20 puede entonces invocar la unidad de reordenamiento 34, que puede reordenar los coeficientes HOA transformados (que, de nuevo en el contexto de la SVD, pueden referirse a los vectores $US[k]$ 33 y a los vectores $V[k]$ 35 basándose en el parámetro para generar los coeficientes HOA transformados reordenados 33'/35' (o, en otras palabras, los vectores $US[k]$ 33' y los vectores $V[k]$ 35'), como se ha descrito anteriormente (109). El dispositivo de codificación de audio 20 puede, durante cualquiera de las operaciones anteriores o operaciones subsiguientes, invocar también la unidad de análisis del campo sonoro 44. La unidad de análisis del campo sonoro 44 puede, como se ha descrito anteriormente, realizar un análisis de campo sonoro con respecto a los coeficientes HOA transformados 11 y/o los coeficientes HOA transformados 33/35 para determinar el número total de canales de primer plano (nFG) 45, el orden del campo sonoro de trasfondo (N_{BG}) y el número ($nBGa$) y los índices (i) de canales adicionales HOA de BG a enviar (que pueden indicarse colectivamente como información de canal de trasfondo 43 en el ejemplo de la FIG. 3) (109).

[145] El dispositivo de codificación de audio 20 también puede invocar la unidad de selección de trasfondo 48. La unidad de selección de trasfondo 48 puede determinar coeficientes HOA de trasfondo o HOA ambientales 47 basándose en la información de canal de trasfondo 43 (110). El dispositivo de codificación de audio 20 puede invocar adicionalmente la unidad de selección del primer plano 36, la cual puede seleccionar los vectores $US[k]$ reordenados 33' y los vectores $V[k]$ reordenados 35 que representan componentes en primer plano o distintos del campo sonoro basándose en nFG 45 (que puede representar uno o más índices que identifican los vectores de primer plano) (112).

[146] El dispositivo de codificación de audio 20 puede invocar la unidad de compensación de energía 38. La unidad de compensación de energía 38 puede realizar compensación de energía con respecto a los coeficientes HOA ambientales 47 para compensar la pérdida de energía debida a la eliminación de varios de los coeficientes HOA por la unidad de selección de trasfondo 48 (114) y generar así coeficientes HOA ambientales compensados en energía 47'.

[147] El dispositivo de codificación de audio 20 puede también invocar la unidad de interpolación espacio-temporal 50. La unidad de interpolación espacio-temporal 50 puede realizar interpolación espacio-temporal con respecto a los coeficientes HOA transformados reordenados 33'/35' para obtener las señales de primer plano interpoladas 49' (que también pueden denominarse las "nFG señales interpoladas 49'") y la información direccional restante en primer plano 53 (que también se puede mencionar como los "vectores $V[k]$ 53") (116). El dispositivo de codificación de audio 20 puede entonces invocar la unidad de reducción de coeficientes 46. La unidad de reducción de coeficientes 46 puede realizar una reducción de los coeficientes con respecto a los restantes vectores $V[k]$ de primer plano, basándose en la información de canal de trasfondo 43, para obtener información direccional reducida en primer plano 55 (que también se puede mencionar como vectores de primer plano reducidos $V[k]$ 55) (118).

[148] El dispositivo de codificación de audio 20 puede entonces invocar la unidad de cuantización 52 para comprimir, de la manera descrita anteriormente, los vectores $V[k]$ de primer plano reducidos 55 y generar los vectores $V[k]$ de primer plano codificados 57 (120).

[149] El dispositivo de codificación de audio 20 también puede invocar la unidad codificadora de audio psicoacústico 40. La unidad codificadora de audio psicoacústico 40 puede codificar psicoacústicamente cada vector de los coeficientes HOA ambientales compensados en energía 47' y las nFG señales interpoladas 49' para generar coeficientes HOA ambientales codificados 59 y nFG señales codificadas 61. El dispositivo de codificación de audio puede entonces invocar la unidad de generación de flujos de bits 42. La unidad de generación del flujo de bits 42 puede generar el flujo de bits 21 basándose en la información direccional codificada en primer plano 57, los coeficientes HOA ambientales codificados 59, las nFG señales codificadas 61 y la información de canal de trasfondo 43.

[150] La FIG. 5B es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de codificación de audio en la realización de las técnicas de transición descritas en esta divulgación. El dispositivo de codificación de audio 20 puede representar un ejemplo de un dispositivo de codificación de audio configurado para realizar las técnicas de transición descritas en esta divulgación. En particular, la unidad de generación de flujos de bits 42 puede mantener información de estado de transición (como se describe con más detalle a continuación con respecto a la FIG. 8) para cada coeficiente HOA ambiental (incluidos los coeficientes HOA ambientales adicionales). La información de estado de transición puede indicar si cada uno de los coeficientes HOA ambientales se encuentran actualmente en uno de los tres estados. Los tres estados pueden incluir un estado de fundido inicial, un estado sin cambios y un estado de fundido final. Mantener la información de estado de transición puede permitir que la unidad de generación de flujos de bits 42 reduzca la sobrecarga de bits en cuanto a que uno o más elementos sintácticos puedan obtenerse basándose en la información de estado de transición mantenida en el dispositivo de decodificación de audio 24.

[151] La unidad de generación de flujos de bits 42 puede determinar además cuándo uno de los coeficientes HOA ambientales especificados en uno de los canales de transporte (tales como el expuesto a continuación con respecto a las FIGs. 7D y 7E) está en transición (302). La unidad de generación de flujos de bits 42 puede determinar cuándo el coeficiente HOA está en transición basándose en el nFG 45 y la información de canal de trasfondo 43. La unidad de generación de flujos de bits 42 puede actualizar la información de estado de transición para aquel de los coeficientes HOA para el que se determina que está en transición (304). Basándose en la información de estado de transición actualizada, la unidad de generación de flujos de bits 42 puede obtener un bit indicativo de cuándo el coeficiente HOA ambiental está en transición (306). La unidad de generación de flujos de bits 42 puede producir el flujo de bits 21 para incluir el bit indicativo de cuándo uno de los coeficientes HOA está en transición (308).

[152] Aunque se describe como realizado por la unidad de generación de flujos de bits 42, las técnicas anteriores pueden realizarse mediante cualquier combinación de unidades 44, 48, 46 y 42. Por ejemplo, la unidad de análisis de campo de sonido 44 puede mantener la información de estado de transición para cada uno de los coeficientes HOA ambientales en base a la información de canal de trasfondo 43. La unidad de análisis de campo de sonido 44 puede obtener el bit indicativo de la transición basándose en la información de estado de transición y proporcionar este bit a la unidad de generación de flujos de bits 42. La unidad de generación de flujos de bits 42 puede producir entonces el flujo de bits 21 para incluir el bit indicativo de la transición.

[153] Como otro ejemplo, la unidad de selección de trasfondo 48 puede mantener la información de estado de transición basándose en la información de canal de trasfondo 43 y obtener el bit indicativo de la transición basándose en la información de estado de transición. La unidad de generación de flujos de bits 42 puede obtener el bit indicativo de la transición desde la unidad de selección de trasfondo 48 y producir el flujo de bits 21 para incluir el bit indicativo de la transición.

[154] Como otro ejemplo más, la unidad de reducción de coeficientes 46 puede mantener la información de estado de transición basándose en la información de canal de trasfondo 43 y obtener el bit indicativo de la transición basándose en la información de estado de transición. La unidad de generación de flujos de bits 42 puede obtener el bit indicativo de la transición desde la unidad de reducción de coeficientes 46 y producir el flujo de bits 21 para incluir el bit indicativo de la transición.

[155] La FIG. 6A es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de decodificación de audio, tal como el dispositivo de decodificación de audio 24 que se muestra en la FIG. 4, en la realización de diversos aspectos de las técnicas descritas en esta divulgación. Inicialmente, el dispositivo decodificador de audio 24 puede recibir el flujo de bits 21 (130). Al recibir el flujo de bits, el dispositivo de decodificación de audio 24 puede invocar la unidad de extracción 72. Suponiendo, a efectos de análisis, que el flujo de bits 21 indica que se va a realizar una reconstrucción basada en vectores, el dispositivo de extracción 72 puede analizar sintácticamente el flujo de bits para recuperar la información indicada anteriormente, pasando la información a la unidad de reconstrucción basada en vectores 92.

[156] En otras palabras, la unidad de extracción 72 puede extraer la información direccional codificada en primer plano 57 (que, de nuevo, también se puede mencionar como los vectores $V[k]$ de primer plano codificados 57), los coeficientes HOA ambientales codificados 59 y las señales de primer plano codificadas (que también se pueden denominar las nFG señales de primer plano codificadas 59 o los objetos de audio de primer plano codificados 59) del flujo de bits 21, de la manera descrita anteriormente (132).

[157] El dispositivo de decodificación de audio 24 puede invocar adicionalmente la unidad de descuantización 74. La unidad de descuantización 74 puede decodificar por entropía y descuantizar la información direccional codificada en primer plano 57 para obtener información direccional reducida en primer plano 55_k (136). El dispositivo de decodificación de audio 24 también puede invocar la unidad de decodificación psicoacústica 80. La unidad de codificación de audio psicoacústico 80 puede decodificar los coeficientes HOA ambientales codificados 59 y las señales de primer plano codificadas 61 para obtener coeficientes HOA ambientales compensados en energía $47'$ y las señales de primer plano interpoladas $49'$ (138). La unidad de decodificación psicoacústica 80 puede pasar los coeficientes HOA ambientales compensados en energía $47'$ a la unidad de fundido 770 y las nFG señales $49'$ a la unidad de formulación de primer plano 78.

[158] El dispositivo de decodificación de audio 24 puede invocar a continuación la unidad de interpolación espacio-temporal 76. La unidad de interpolación espacio-temporal 76 puede recibir la información direccional de primer plano reordenada $55_k'$ y realizar la interpolación espacio-temporal con respecto a la información direccional reducida de primer plano $55_k/55_{k-1}$ para generar la información direccional de primer plano interpolada $55_k''$ (140). La unidad de interpolación espacio-temporal 76 puede remitir los vectores $V[k]$ de primer plano interpolados $55_k''$ a la unidad de fundido 770.

[159] El dispositivo de decodificación de audio 24 puede invocar la unidad de fundido 770. La unidad de

fundido 770 puede recibir u obtener de otro modo elementos sintácticos (por ejemplo, desde la unidad de extracción 72) indicativos de cuándo los coeficientes HOA ambientales compensados de energía 47' están en transición (por ejemplo, el elemento sintáctico TransiciónCoefAmb). La unidad de fundido 770 puede, basándose en los elementos sintácticos de transición y la información de estado de transición mantenida, incorporar o eliminar los coeficientes HOA ambientales compensados en energía 47', emitiendo coeficientes HOA ambientales ajustados 47" a la unidad de formulación de coeficientes HOA 82. La unidad de fundido 770 puede también, sobre la base de los elementos sintácticos y la información de estado de transición mantenida, incorporar o eliminar los correspondientes uno o más elementos de los vectores de primer plano interpolados $V[k]$ 55_k", emitiendo los vectores $V[k]$ ajustados en primer plano 55_k" a la unidad de formulación de primer plano 78 (142).

[160] El dispositivo de decodificación de audio 24 puede invocar la unidad de formulación del primer plano 78. La unidad de formulación del primer plano 78 puede realizar la multiplicación matricial de las nFG señales 49' por la información direccional de primer plano ajustada 55_k" para obtener los coeficientes HOA de primer plano 65 (144). El dispositivo de decodificación de audio 24 también puede invocar la unidad de formulación de coeficientes HOA 82. La unidad de formulación de coeficientes HOA 82 puede añadir los coeficientes HOA de primer plano 65 a los coeficientes HOA ambientales 47' para obtener los coeficientes HOA 11' (146).

[161] La FIG. 6B es un diagrama de flujo que ilustra el funcionamiento ejemplar de un dispositivo de decodificación de audio en la realización de las técnicas de transición descritas en esta divulgación. El dispositivo decodificador de audio 24 que se muestra en el ejemplo de la FIG. 4 puede representar un ejemplo de un dispositivo de decodificación de audio configurado para realizar las técnicas de transición descritas en esta divulgación.

[162] En particular, la unidad de fundido 770 puede obtener un bit (en forma de indicación 757, donde la indicación 757 puede representar un elemento sintáctico de TransiciónCoefAmb), indicativo de cuándo uno de los coeficientes HOA ambientales 47' está en transición (352). La unidad de fundido 770 puede mantener la información de estado de transición que se describe a continuación con más detalle a continuación con respecto al ejemplo de la FIG. 8, basado en el bit indicativo de la transición (354). La información de estado de transición puede indicar si cada uno de los coeficientes HOA ambientales está actualmente en uno de tres estados. Los tres estados pueden incluir un estado de fundido inicial, un estado sin cambios y un estado de fundido final.

[163] La unidad de fundido 770 puede mantener la información de estado de transición, al menos en parte, actualizando la información de estado de transición basándose en la indicación 757 de que uno de los coeficientes HOA ambientales 47' está en transición. Por ejemplo, la unidad de fundido 770 puede mantener información de estado de transición para uno de los coeficientes HOA ambientales 47' que indica que uno de los coeficientes HOA ambientales 47' está en un estado de transición sin cambio. Al obtener una indicación de que uno de los coeficientes HOA ambientales 47' está en transición, la unidad de fundido 770 puede actualizar la información de estado de transición para uno de los coeficientes HOA ambientales 47', para indicar que uno de los coeficientes HOA ambientales 47' debe ser eliminado. Como otro ejemplo, la unidad de fundido 770 puede mantener información de estado de transición para uno de los coeficientes HOA ambientales 47' que indica que uno de los coeficientes HOA ambientales 47' se ha eliminado. Tras obtener una indicación de que uno de los coeficientes HOA ambientales 47' está en transición, la unidad de fundido 770 puede actualizar la información de estado de transición para uno de los coeficientes HOA ambientales 47', para indicar que uno de los coeficientes HOA ambientales 47' debe ser incorporado. La unidad de fundido 770 puede entonces realizar la transición basándose en la información de estado de transición actualizada, de la manera descrita anteriormente con respecto a la FIG. 4, y más adelante con mayor detalle con respecto a la FIG. 8 (356).

[164] Las FIGs. 7A a 7J son diagramas que ilustran partes de la información del flujo de bits o del canal lateral, que puede especificar los componentes espaciales comprimidos con más detalle. En el ejemplo de la FIG. 7A, una parte 250 incluye un campo identificador de reproductor ("Identificador de reproductor") 251 y un campo ConfigDecodificadorHOA 252 (que también se puede denominar campo ConfigHOA 252). El campo Identificador de reproductor 251 puede representar un campo que almacena un Identificador del reproductor que se ha utilizado para la mezcla del contenido HOA. El campo ConfigDecodificadorHOA 252 puede representar un campo configurado para almacenar información para inicializar el decodificador espacial HOA, tal como el dispositivo de decodificación de audio 24 mostrado en el ejemplo de la FIG. 4.

[165] El campo ConfigDecodificadorHOA 252 incluye además un campo de información direccional ("información de dirección") 253, un campo TiempoInterpolaciónEspacialCodificada 254, un campo ProcedimientoInterpolaciónEspacial 255, un campo LongitudVectVCodificado 256 y un campo de información de ganancia 257. El campo de información direccional 253 puede representar un campo que almacena información para configurar el decodificador de síntesis de base direccional. El campo TiempoInterpolaciónEspacialCodificada 254 puede representar un campo que almacena un tiempo de la interpolación espacio-temporal de las señales basadas en vectores. El campo ProcedimientoInterpolaciónEspacial 255 puede representar un campo que almacena una indicación del tipo de

interpolación aplicada durante la interpolación espacio-temporal de las señales basadas en vectores. El campo LongitudVectVCodificado 256 puede representar un campo que almacena una longitud del vector de datos transmitido, utilizado para sintetizar las señales basadas en vectores. El campo de información de ganancia 257 representa un campo que almacena información indicativa de una corrección de ganancia aplicada a las señales.

[166] En el ejemplo de la FIG. 7B, la parte 258A representa una parte del canal de información lateral, en donde la parte 258A incluye una cabecera de trama 259 que incluye un campo del número de octetos 260 y un campo nbits 261. El campo del número de octetos 260 puede representar un campo para expresar el número de octetos incluidos en la trama para especificar los componentes espaciales v1 a vn que incluyen los ceros para el campo de alineación de octetos 264. El campo nbits 261 representa un campo que puede especificar el valor *nbits* identificado para su uso en la descompresión de los componentes espaciales v1 a vn.

[167] Como se muestra adicionalmente en el ejemplo de la FIG. 7B, la parte 258A puede incluir sub-flujos de bits para v1 a vn, cada uno de los cuales incluye un campo de modalidad de predicción 262, un campo de información de tabla de Huffman 263 y un correspondiente componente entre los componentes espaciales comprimidos v1 a vn. El campo de modalidad de predicción 262 puede representar un campo para almacenar una indicación de si la predicción fue realizada con respecto al componente correspondiente entre los componentes espaciales comprimidos v1 a vn. El campo de información de tabla de Huffman 263 representa un campo para indicar, al menos en parte, qué tabla de Huffman se va a utilizar para decodificar diversos aspectos del componente correspondiente entre los componentes espaciales comprimidos v1 a vn.

[168] A este respecto, las técnicas pueden permitir que el dispositivo de codificación de audio 20 obtenga un flujo de bits que comprende una versión comprimida de un componente espacial de un campo sonoro, el componente espacial generado realizando una síntesis basada en vectores con respecto a una pluralidad de coeficientes armónicos esféricos.

[169] La FIG. 7C es un diagrama que ilustra una parte 250 del flujo de bits 21. La parte 250 que se muestra en el ejemplo de la FIG. 7C incluye un campo OrdenHOA (que no se ha mostrado en el ejemplo de la FIG. 7F con fines de facilitar la ilustración), un campo MínOrdenHoaAmb (que de nuevo no se ha mostrado en el ejemplo de la FIG. 10 con fines de facilitar la ilustración), el campo de información direccional 253, el campo TiempoInterpolaciónEspacialCodificada 254, el campo ProcedimientoInterpolaciónEspacial 255, el campo LongitudVectVCodificado 256 y el campo de información de ganancia 257. Como se muestra en el ejemplo de la FIG. 7C, el campo TiempoInterpolaciónEspacialCodificada 253 puede comprender un campo de tres bits, el campo ProcedimientoInterpolaciónEspacial 255 puede comprender un campo de un bit y el campo LongitudVectVCodificado 256 puede comprender un campo de dos bits. La FIG. 7D es un diagrama que ilustra las tramas ejemplares 249C y 249R especificadas de acuerdo a diversos aspectos de las técnicas descritas en esta divulgación. Como se muestra en el ejemplo de la FIG. 7D, la trama 249A incluye los campos DatosInfoLateralCanal (CSID) 154A a 154D, los campos DatosCorrecciónGananciaHOA (HOAGCD), los campos DatosVectorV 156 y 156B y los campos InfoPredicciónHOA. El campo CSID 154A incluye un elemento sintáctico ("unidadC") 267, un elemento sintáctico ("bb") 266 y un elemento sintáctico ("ba") 265, junto con un elemento sintáctico TipoCanal 269, cada uno de los cuales está configurado en los valores correspondientes 01, 1, 0 y 01, que se muestran en el ejemplo de la FIG. 7D. El campo CSID 154B incluye los campos unitC 267, bb 266 y ba265, junto con el TipoCanal 269, cada uno de los cuales está configurado en los valores correspondientes 01, 1, 0 y 01 que se muestran en el ejemplo de la FIG. 7D. Cada uno de los campos CSID 154C y 154D incluye el campo TipoCanal 269 que tiene un valor de 3 (11_2). Cada uno de los campos CSID 154A a 154C corresponde al canal respectivo de los canales de transporte 1, 2, 3 y 4. En efecto, cada campo CSID 154A a 154C indica si una carga útil correspondiente son señales basadas en la dirección (cuando el TipoCanal correspondiente es igual a cero), señales basadas en vectores (cuando el TipoCanal correspondiente es igual a uno), un coeficiente HOA ambiental adicional (cuando el TipoCanal correspondiente es igual a dos), o está vacía (cuando el TipoCanal es igual a tres).

[170] En el ejemplo de la FIG. 7D, la trama 249Q incluye dos señales basadas en vectores (dado el TipoCanal 269 igual a 1 en los campos CSID 154A y 154B) y dos vacías (dado el TipoCanal 269 igual a 3 en los campos CSID 154C y 154D). Dada la parte de ConfigHOA anterior 250, el dispositivo de decodificación de audio 24 puede determinar que todos los 16 elementos del vector V están codificados. Por tanto, los DatosVectorV 156A y 156B incluyen, cada uno, todos los 16 elementos vectoriales, cada uno de ellos uniformemente cuantizado con 8 bits. El número y los índices de los elementos de DatosVectorV codificados están especificados por el parámetro longitudVectVCodificado = 0. Además, el esquema de codificación se señala mediante NbitsQ = 5 en el campo CSID para el canal de transporte correspondiente.

[171] Las tramas 249Q y 249R también incluyen un indicador de independencia HOA ("IndicadorIndependenciaHOA") 860. El indicador de independencia HOA 860 representa un campo que especifica si la trama es una trama de emisión inmediata. Cuando el valor del campo 860 se fija en uno, las tramas 249Q y / o 249R se pueden decodificar de forma independiente sin referencia a otras tramas (es decir, no puede requerirse ninguna predicción para decodificar la trama). Cuando el valor del campo 860 se fija en

cero, las tramas 249Q y / o 249R pueden no ser decodificables independientemente (es decir, que la predicción de varios valores descritos anteriormente puede predecirse a partir de otras tramas). Sin embargo, como se muestra en el ejemplo de la FIG. 7D, la trama 249Q no incluye un campo InfoPredicciónHOA. En consecuencia, el campo InfoPredicciónHOA puede representar un campo optativo en el flujo de bits.

[172] La FIG. 7E es un diagrama que ilustra tramas ejemplares 249S y 249T especificadas de acuerdo a diversos aspectos de las técnicas descritas en esta divulgación. La trama 249S puede ser similar a la trama 249Q, excepto en que la trama 249S puede representar un ejemplo en el que el indicador de independencia HOA 860 se fija en cero y la predicción ocurre con respecto a la parte unitaria del elemento sintáctico Nbits para el número de transporte 2 que se reutiliza desde la trama anterior (que se supone que sea 5 en el ejemplo de la FIG. 7E). La trama 249T también puede ser similar a la trama 249Q, excepto en que la trama 249T tiene un valor de uno para el indicador de independencia HOA 860. En este ejemplo, se supone que la parte unitaria del valor Q de Nbits podría haberse reutilizado a partir de la trama anterior, como en el ejemplo de la trama 249S. Sin embargo, debido a que el indicador de independencia HOA (que también puede indicarse como un elemento sintáctico) se fija en uno, el dispositivo de codificación de audio 20 especifica todo el elemento sintáctico Nbits 261 para el segundo canal de transporte, de modo que la trama 249S pueda decodificarse independientemente sin referencia a valores previos (por ejemplo, la parte unitaria del campo Nbits 261 de la trama anterior).

[173] Además, como el indicador de independencia HOA se fija en uno (lo que significa que la trama 249T es decodificable independientemente sin referencia a tramas anteriores), el dispositivo de codificación de audio 20 puede no señalar el indicador de predicción utilizado para la cuantización escalar, ya que no se permite ninguna predicción para tramas decodificables independientemente (lo que puede representar otra forma de referirse a las "tramas de emisión inmediata" espuestas en esta divulgación). Cuando el elemento sintáctico del indicador de independencia HOA 860 se fija en uno, en otras palabras, el dispositivo de codificación de audio 20 no necesita señalar el indicador de predicción ya que el dispositivo de decodificación de audio 24 puede determinar, basándose en el valor del elemento sintáctico del indicador de independencia HOA 860, que la predicción con fines de cuantización escalar ha sido desactivada.

[174] La FIG. 7F es un diagrama que ilustra un segundo flujo de bits ejemplar 248K y la parte correspondiente de ConfigHOA 250K, que se han generado para corresponder al caso 1 en el seudo-código anterior. En el ejemplo de la FIG. 7F, la parte de ConfigHOA 250K incluye un elemento sintáctico LongitudVectVCodificado 256, configurado para indicar que todos los elementos de un vector V están codificados, excepto los elementos sintácticos 1 al MínNúmDeCoefsParaAmbHOA y los elementos especificados en un elemento sintáctico CanalHoaAmbAdicCont (se supone que es uno en este ejemplo). La parte de ConfigHOA 250K también incluye un elemento sintáctico ProcedimientoInterpolaciónEspacial 255, fijado para indicar que la función de interpolación de la interpolación espacio-temporal es un coseno elevado. La parte de ConfigHOA 250K además incluye un TiempoInterpolaciónEspacialCodificada 254 para indicar una duración de la muestra interpolada de 256.

[175] La parte de ConfigHOA 250N incluye además un elemento sintáctico MínOrdenHoaAmb 150 configurado para indicar que el orden de MínimoHOA del contenido de HOA ambiental es uno, donde el dispositivo de decodificación de audio 24 puede obtener un elemento sintáctico MínNúmDeCoefsParaAmbHOA igual a $(1+1)^2$, o cuatro. El dispositivo de decodificación de audio 24 puede obtener también un elemento sintáctico N°MáxDeCoefsAmbActivosAdic, configurado como una diferencia entre el elemento sintáctico NúmDeCoefHoa y el MínNúmDeCoefsParaAmbHOA, que se supone en este ejemplo igual a $16-4$, o 12. El dispositivo de decodificación de audio 24 también puede obtener un elemento sintáctico AsignarAmbmBits, configurado como $\text{ceil}(\log_2(\text{MáxN}^\circ\text{DeCoefsAmbActivosAdic})) = \text{ceil}(\log_2(12)) = 4$. La parte de ConfigHOA 250K incluye un elemento sintáctico OrdenHoa 152 configurado para indicar que el orden HOA del contenido es igual a tres (o, en otras palabras, $N = 3$), donde el dispositivo de decodificación de audio 24 puede obtener un NúmDeCoefsHoa que es igual a $(N+1)^2$, o 16.

[176] Como se muestra adicionalmente en el ejemplo de la FIG. 7F, la parte 248K incluye una trama de audio USAC-3D en la que dos tramas HOA 249G y 249H se almacenan en una carga útil de extensión USAC, dado que dos tramas de audio se almacenan dentro de una trama USAC-3D cuando la replicación de banda espectral (SBR) está habilitada. El dispositivo de decodificación de audio 24 puede obtener una serie de canales de transporte flexibles como una función de un elemento sintáctico númCanalesTransporteHOA y un elemento sintáctico MínNúmDeCoefsParaAmbHOA. En los siguientes ejemplos, se supone que el elemento sintáctico númCanalesTransporteHOA es igual a 7 y el elemento sintáctico MínNúmDeCoefsParaAmbHOA es igual a cuatro, donde el número de canales de transporte flexibles es igual al elemento sintáctico númCanalesTransporteHOA menos el elemento sintáctico MínNúmDeCoefsParaAmbHOA (o tres).

[177] La FIG. 7G es un diagrama que ilustra las tramas 249G y 249H con más detalle. Como se muestra en el ejemplo de la FIG. 7G, la trama 249G incluye los campos CSID 154 a 154C y los campos DatosVectorV 156. El campo CSID 154 incluye el ÍndCoefAmbCodif 246, el TransiciónÍndCoefAmb 247 (donde el doble asterisco (**)) indica que, para el canal de transporte flexible n° 1, aquí se supone que el estado interno del decodificador

es EstadoTransiciónÍndCoefAmb = 2, lo que da como resultado que el campo de bits ÍndCoefAmbCodif se señalice o se especifique de otro modo en el flujo de bits) y el TipoCanal 269 (que es igual a dos, señalizando que la carga útil correspondiente es un coeficiente HOA ambiental adicional). El dispositivo de decodificación de audio 24 puede obtener el ÍndCoefAmb como igual al ÍndCoefAmbCodificado + 1 + MínNúmDeCoefsParaAmbHOA, o 5 en este ejemplo. El campo CSID 154B incluye unidad C 267, bb 266 y ba 265, junto con el TipoCanal 269, cada uno de los cuales está fijado en los valores correspondientes 01, 1, 0 y 01 que se muestran en el ejemplo de la FIG. 10K(ii). El campo CSID 154C incluye el campo TipoCanal 269 que tiene un valor de 3.

[178] En el ejemplo de la FIG. 7G, la trama 249G incluye una única señal basada en vectores (dado el TipoCanal 269 igual a 1 en los campos CSID 154B) y una vacía (dado el TipoCanal 269 igual a 3 en los campos CSID 154C). Dada la parte de ConfigHOA anterior 250K, el dispositivo de decodificación de audio 24 puede determinar que se codifiquen 11 elementos del vector V (donde 12 se obtiene como $\text{OrdenHOA} + 1$)² - (MínNúmDeCoefsParaAmbHOA) - (CanHoaAmbAdicCont) = 16-4-1 = 11). Por lo tanto, los DatosVectorV 156 incluyen todos los 11 elementos vectoriales, cada uno de ellos uniformemente cuantizado con 8 bits. Como se indica en la nota al pie 1, el número y los índices de los elementos de DatosVectorV codificados se especifican mediante el parámetro LongitudVectVCodificado = 0. Además, como se indica en la nota a pie de página 2, el esquema de codificación se indica mediante NbitsQ = 5 en el campo CSID para el canal de transporte correspondiente.

[179] En la trama 249P, el campo CSID 154 incluye una TransiciónÍndCoefAmb 247 que indica que no ha ocurrido ninguna transición y por lo tanto el ÍndCoefAmbCodificado 246 puede estar implícito a partir de la trama anterior y no necesita ser señalado o especificado de nuevo de otro modo. El campo CSID 154B y 154C de la trama 249H es el mismo que el de la trama 249G y, por lo tanto, al igual que la trama 249G, la trama 249H incluye un único campo DatosVectorV 156, el cual incluye 10 elementos vectoriales, cada uno de ellos cuantizado uniformemente con 8 bits. El dispositivo de codificación de audio 20 solo especifica 10 elementos vectoriales porque el coeficiente HOA ambiental especificado en el canal de transporte número uno ya no está en transición y, como resultado, el número de CanHoaAmbAdicCont es igual a dos. En consecuencia, el dispositivo de codificación de audio 20 determina que el número de elementos del vector V a especificar es $(\text{OrdenHOA} + 1)^2 - (\text{MínNúmDeCoefsParaAmbHOA}) - (\text{CanHoaAmbAdicCont}) = 16-4-2 = 10$.

[180] Aunque el ejemplo de las FIGs. 7F y 7G representan el flujo de bits 21, estructurado de acuerdo a una de las modalidades codificadas para el vector V, otros diversos ejemplos del flujo de bits 21 pueden estructurarse de acuerdo a las otras modalidades de codificación para el vector V. Los ejemplos adicionales se exponen en más detalle con respecto a la publicación anteriormente indicada, n° WO 2014/194099.

[181] La FIG. 7H es un diagrama que ilustra un ejemplo alternativo de la trama 249H en la que el IndicadorIndependenciaHOA se fija en uno de acuerdo a diversos aspectos de las técnicas descritas en esta divulgación. La trama alternativa de 249H se indica como la trama 249H'. Cuando el elemento sintáctico IndicadorIndependenciaHOA 860 se fija en uno, la trama 249H' puede representar una trama de emisión inmediata (IPF), como se analiza en más detalle a continuación. Como resultado, el dispositivo de codificación de audio 20 puede especificar elementos sintácticos adicionales en el CAMPO CSID 154A y 154C. Los elementos sintácticos adicionales pueden proporcionar información de estado mantenida por el dispositivo de decodificación de audio 24 basándose en elementos sintácticos pasados. Sin embargo, en el contexto de la IPF 249H', el dispositivo de decodificación de audio 24 puede no tener la información de estado. Como resultado, el dispositivo de codificación de audio 20 especifica el elemento sintáctico EstadoTransiciónCoefAmb 400 en el campo CSID 154A y 154C, para permitir que el dispositivo de decodificación de audio 24 comprenda la transición actual que está siendo señalizada por el elemento sintáctico TransiciónÍndCoefAmb 247 de cada uno de los campos CSID 154A y 154C.

[182] La FIG. 7I es un diagrama que ilustra tramas ejemplares para uno o más canales de al menos un flujo de bits de acuerdo a las técnicas descritas en este documento. El flujo de bits 808 incluye las tramas 810A a 810E que pueden incluir, cada una, uno o más canales, y el flujo de bits 808 puede representar cualquier combinación de flujos de bits 21 modificados de acuerdo a las técnicas descritas en este documento, con el fin de incluir las IPF. Las tramas 810A a 810E pueden incluirse dentro de las respectivas unidades de acceso y, alternativamente, pueden denominarse "unidades de acceso 810A a 810E".

[183] En el ejemplo ilustrado, una Trama de Reproducción Inmediata (IPF) 816 incluye una trama independiente 810E así como información de estado de las tramas anteriores 810B, 810C y 810D, representadas en la IPF 816 como información de estado 812. Es decir, la información de estado 812 puede incluir un estado mantenido por una máquina de estados 402 a partir del procesamiento de tramas previas 810B, 810C y 810D, representadas en la IPF 816. La información de estado 812 puede estar codificada dentro de la IPF 816 usando una extensión de carga útil dentro del flujo de bits 808. La información de estado 812 puede compensar el retardo del arranque del decodificador para configurar internamente el estado del decodificador para permitir una decodificación correcta de la trama independiente 810E. La información de estado 812 puede, por esta razón, ser mencionada, alternativamente y colectivamente, como "pre-despliegue"

para la trama independiente 810E. En varios ejemplos, el decodificador puede usar más o menos tramas para compensar el retraso del arranque del decodificador, lo que determina la cantidad de información de estado 812 para una trama. La trama independiente 810E es independiente porque las tramas 810E se pueden decodificar independientemente. Como resultado, se puede hacer referencia a la trama 810E como "trama decodificable independientemente 810". La trama independiente 810E puede, como resultado, constituir un punto de acceso al flujo para el flujo de bits 808.

[184] La información de estado 812 puede incluir además los elementos sintácticos ConfigHOA que pueden enviarse al comienzo del flujo de bits 808. La información de estado 812 puede, por ejemplo, describir la tasa de bits del flujo de bits 808 u otra información utilizable para la conmutación del flujo de bits la adaptación de la tasa de bits. Otro ejemplo de lo que puede incluir una parte de la información de estado 814 son los elementos sintácticos ConfigHOA mostrados en el ejemplo de la FIG. 7C. A este respecto, la IPF 816 puede representar una trama sin estado, que no puede, por decirlo de alguna manera, tener ningún recuerdo del pasado. La trama independiente 810E puede, en otras palabras, representar una trama sin estado, que puede decodificarse independientemente de cualquier estado previo (ya que el estado se proporciona en términos de la información de estado 812).

[185] El dispositivo de codificación de audio 20 puede, al seleccionar la trama 810E para que sea una trama independiente, realizar un proceso de transición de la trama 810E desde una trama decodificable dependientemente a una trama decodificable independientemente. El proceso puede implicar especificar información de estado 812 que incluye la información de estado de transición en la trama, la información de estado que permite decodificar y reproducir el flujo de bits de los datos de audio codificados de la trama, sin referencia a tramas anteriores del flujo de bits.

[186] Un decodificador, tal como el decodificador 24, puede acceder aleatoriamente al flujo de bits 808 en la IPF 816 y, al decodificar la información de estado 812 para inicializar los estados del decodificador y las memorias intermedias (por ejemplo, de la máquina de estados del lado del decodificador 402), decodificar la trama independiente 810E para emitir una versión comprimida de los coeficientes HOA. Los ejemplos de la información de estado 812 pueden incluir los elementos sintácticos especificados en la siguiente tabla:

Elemento sintáctico afectado por el IndicadorIndependenciaHOA	Sintaxis descrita en Norma	Finalidad
NbitsQ	Sintaxis de DatosInfoLateralCanal	Cuantización del vector V
IndicadorP	Sintaxis de DatosInfoLateralCanal	Codificación de Huffman de vector V
EstadoTransiciónCoefAmb	Sintaxis de CanallInfoHoaAmbAdic	Señalización de HOA adicional
ExpAmpPrevCorrGanancia	Sintaxis de DatosCorrecciónGananciaHOA	Módulo de compensación automática de ganancias

El decodificador 24 puede analizar los elementos sintácticos anteriores de la información de estado 812 para obtener una o más informaciones de estado de cuantización en forma del elemento sintáctico NbitsQ, información de estado de predicción en forma del elemento sintáctico IndicadorP e información de estado de transición en forma del elemento sintáctico EstadoTransiciónCoefAmb. El decodificador 24 puede configurar la máquina de estados 402 con la información de estado analizada sintácticamente 812 para permitir que la trama 810E sea decodificada independientemente. El decodificador 24 puede continuar la decodificación normal de tramas, después de la decodificación de la trama independiente 810E.

[187] De acuerdo a las técnicas descritas en este documento, el dispositivo de codificación de audio 20 puede configurarse para generar la trama independiente 810E de las IPF 816 de forma diferente a otras tramas 810 para permitir la reproducción inmediata en la trama independiente 810E y / o conmutar entre representaciones de audio del mismo contenido que difieren en la tasa de bits y / o en herramientas habilitadas en la trama independiente 810E. Más específicamente, la unidad de generación de flujos de bits 42 puede mantener la información de estado 812 usando la máquina de estados 402. La unidad de generación de flujos de bits 42 puede generar la trama independiente 810E para incluir información de estado 812 utilizada para configurar la máquina de estados 402 para uno o más coeficientes HOA ambientales. La unidad de generación de flujos de bits 42 puede generar adicionalmente o alternativamente la trama independiente 810E para codificar de forma diferente la información de cuantización y / o predicción a fin de, por ejemplo, reducir un tamaño de trama con respecto a las otras tramas, no IPF, del flujo de bits 808. De nuevo, la unidad de generación de flujos de bits 42 puede mantener el estado de cuantización en la forma de la máquina de estados 402. Además, la unidad de generación de flujos de bits 42 puede codificar cada trama de las tramas 810A a 810E para incluir un indicador u otro elemento sintáctico que indique si la trama es una IPF. El elemento sintáctico puede mencionarse en otro lugar en esta divulgación como un *IndicadorIndependencia* o un *IndicadorIndependenciaHOA*.

[188] A este respecto, diversos aspectos de las técnicas pueden habilitar, como un ejemplo, la unidad de

generación de flujos de bits 42 del dispositivo de codificación de audio 20 para especificar, en un flujo de bits (tal como el flujo de bits 21) que incluye un coeficiente ambisónico de orden superior (tal como uno de los coeficientes ambisónicos ambientales de orden superior 47', información de transición 757 (como parte de la información de estado 812, por ejemplo) para una trama independiente (tal como la trama independiente 810E en el ejemplo de la FIG. 7I) para el coeficiente ambisónico de orden superior 47'. La trama independiente 810E puede incluir información de referencia adicional (que puede referirse a la información de estado 812) para permitir que la trama independiente sea decodificada e inmediatamente reproducida sin referencia a tramas anteriores (por ejemplo, las tramas 810A a 810D) del coeficiente ambisónico de orden superior 47'. Si bien se describe como reproducido de manera inmediata o instantánea, el término inmediatamente o instantáneamente se refiere a lo reproducido de manera casi inmediata, subsiguiente o casi instantánea y no pretende referirse a definiciones literales de "inmediatamente" o "instantáneamente". Además, el uso de los términos es con fines de adoptar el lenguaje utilizado por varias normas, tanto actuales como emergentes.

[189] En estos y en otros casos, la información de transición 757 especifica si el coeficiente ambisónico de orden superior 47' está eliminado. Como se ha indicado anteriormente, la información de transición 757 puede identificar si el coeficiente ambisónico de orden superior 47' está siendo eliminado o incorporado y, como tal, si el coeficiente ambisónico de orden superior 47' se usa para representar diversos aspectos del campo de sonido. En algunos casos, la unidad de generación de flujos de bits 42 especifica la información de transición 757 como varios elementos sintácticos. En estos y en otros casos, la información de transición 757 comprende un indicador CoefAmbFueIncorporado o un elemento sintáctico EstadoTransiciónCoefAmb para el coeficiente ambisónico de orden superior 47', para especificar si el coeficiente ambisónico de orden superior 47' ha de eliminarse para una transición. En estos y otros casos, la información de transición especifica que el coeficiente ambisónico de orden superior 47' está en transición.

[190] En estos y otros casos, la información de transición 757 comprende un indicador TransiciónÍndCoefAmb para especificar que el coeficiente ambisónico de orden superior 47' está en transición.

[191] En estos y otros casos, la unidad de generación de flujos de bits 42 puede configurarse adicionalmente para generar una señal basada en vectores, representativa de uno o más componentes distintos del campo de sonido que incluye un elemento de un vector (tal como uno de los vectores $V[k]$ en primer plano reducidos 55) correspondientes al coeficiente ambisónico de orden superior 47'. El vector 55 puede describir aspectos espaciales de un componente distinto del campo de sonido y puede haber sido descompuesto a partir de coeficientes ambisónicos de orden superior 11 descriptivos del campo de sonido, en donde la trama comprende la señal basada en vectores.

[192] En estos y otros ejemplos, la unidad de generación de flujos de bits 42 puede configurarse además para emitir la trama mediante un protocolo de transmisión por flujo.

[193] Varios aspectos de las técnicas también pueden, en algún ejemplo, permitir que la unidad de generación de flujos de bits 42 especifique, en un flujo de bits 21 que incluye un coeficiente ambisónico de orden superior 47', si una trama para el coeficiente ambisónico de orden superior 47' es una trama independiente (por ejemplo, especificando el elemento sintáctico *IndicadorIndependenciaHOA*) que incluye información de referencia adicional (por ejemplo, la información de estado 812) para permitir que la trama sea decodificada e inmediatamente reproducida sin referencia a las tramas previas 810A a 810D del coeficiente ambisónico de orden superior 47'. La unidad de generación de flujos de bits 42 también puede especificar, en la secuencia de bits 21 y solo cuando la trama no es una trama independiente, información de predicción (por ejemplo, el elemento sintáctico *IndicadorP*) para la trama, para decodificar la trama con referencia a una trama previa del coeficiente ambisónico de orden superior 47'.

[194] En estos y otros ejemplos, la unidad de generación de flujos de bits 42 está configurada además para especificar, en el flujo de bits 21 y cuando la trama es una trama independiente, información de cuantización (por ejemplo, el elemento sintáctico *NbitsQ*), la suficiente para permitir que la trama sea decodificada e inmediatamente reproducida sin referencia a la información de cuantización para las tramas previas del coeficiente ambisónico de orden superior 47'. La unidad de generación de flujos de bits 42 también puede especificar, en el flujo de bits 21 y si la trama no es una trama independiente, información de cuantización para la trama, que es insuficiente para permitir que la trama se decodifique e inmediatamente se reproduzca sin referencia a información de cuantización para tramas anteriores del coeficiente ambisónico de orden superior 47'.

[195] En estos y otros ejemplos, la información de cuantización para la trama incluye un elemento sintáctico *Nbits* para la trama, suficiente para permitir que la trama sea decodificada e inmediatamente reproducida sin referencia a la información de cuantización para las tramas previas del canal ambisónico de orden superior.

[196] En estos y otros ejemplos, la unidad de generación de flujos de bits 42 está configurada además para generar una señal basada en vectores, representativa de uno o más componentes distintos del campo de sonido que incluye un elemento de un vector (tal como el vector 55) correspondiente al coeficiente ambisónico

de orden superior 47', describiendo el vector los aspectos espaciales de un componente distinto del campo de sonido y habiendo sido descompuesto a partir de coeficientes ambisónicos de orden superior 11 descriptivos del campo de sonido. La trama, en este ejemplo, comprende la señal basada en vectores.

5 **[197]** En estos y otros ejemplos, la unidad de generación de flujos de bits 42 está configurada además para emitir la trama mediante un protocolo de transmisión por flujo.

10 **[198]** Diversos aspectos de las técnicas también pueden, en algún ejemplo, permitir que la unidad de generación de flujos de bits 42 especifique, en un flujo de bits 21 que incluye un coeficiente ambisónico de orden superior 47', que una trama para el coeficiente ambisónico de orden superior 47' es una trama independiente que incluye información de referencia adicional para permitir que la trama sea decodificada e inmediatamente reproducida sin referencia a tramas previas del coeficiente ambisónico de orden superior 47'.

15 **[199]** En estos y otros ejemplos, la unidad de generación de flujos de bits 42 está configurada para, al especificar que la trama para el coeficiente ambisónico de orden superior 47' es una trama independiente 810E, señalar en el flujo de bits 21, un elemento sintáctico IndicadorIndependencia que indica la trama es una trama independiente 810E.

20 **[200]** Además, diversos aspectos de las técnicas pueden permitir que el dispositivo de decodificación de audio 24 se configure para obtener, usando un flujo de bits 21 que incluye un coeficiente ambisónico de orden superior 47, información de transición (tal como la información de transición 757 mostrada en el ejemplo de la FIG. 4) para una trama independiente para el coeficiente ambisónico de orden superior 47'. La trama independiente puede incluir información de estado 812 para permitir que la trama independiente sea decodificada y reproducida sin referencia a tramas previas del coeficiente ambisónico de orden superior 47'.

25 **[201]** En estos y en otros casos, la información de transición 757 especifica si el coeficiente ambisónico de orden superior 47' ha de eliminarse para una transición.

30 **[202]** En estos y otros casos, la información de transición 757 comprende un indicador CoefAmbFueIncorporado para el canal ambisónico de orden superior, para especificar si el coeficiente ambisónico de orden superior 47' ha de eliminarse para una transición.

35 **[203]** En estos y en otros casos, el dispositivo de decodificación de audio 24 puede configurarse para determinar que la información de transición 757 especifica que el coeficiente ambisónico de orden superior 47' ha de eliminarse para una transición. El dispositivo de decodificación de audio 24 también puede configurarse para, en respuesta a la determinación de que la información de transición 757 especifica que el coeficiente ambisónico de orden superior 47' ha de eliminarse para una transición, realizar una operación de fundido final con respecto al coeficiente ambisónico de orden superior 47'.

40 **[204]** En estos y otros casos, la información de transición 757 especifica que el coeficiente ambisónico de orden superior 47' está en transición.

45 **[205]** En estos y otros casos, la información de transición 757 comprende un indicador TransiciónCoefAmb para especificar que el coeficiente ambisónico de orden superior 47' está en transición.

50 **[206]** En estos y otros casos, el dispositivo de decodificación de audio 24 puede configurarse para obtener una señal basada en vectores, representativa de uno o más componentes distintos del campo de sonido que incluye un elemento de un vector 55_k correspondiente al coeficiente ambisónico de orden superior 47'. El vector 55_k puede, como se ha indicado anteriormente, describir aspectos espaciales de un componente distinto del campo de sonido y puede haber sido descompuesto a partir de coeficientes ambisónicos de orden superior 11, descriptivos del campo de sonido. El dispositivo de decodificación de audio 24 también puede configurarse para determinar que la información de transición 757 especifica que el coeficiente ambisónico de orden superior 47' ha de eliminarse. El dispositivo de decodificación de audio 24 también puede configurarse para, en respuesta a la determinación de que la información de transición 757 especifica que el coeficiente ambisónico de orden superior 47 ha de eliminarse para una transición, realizar una operación de fundido final con respecto al elemento del vector 55_k correspondiente al canal ambisónico de orden superior 47, para eliminar el elemento del vector 55_k utilizando la trama o una trama posterior para el coeficiente ambisónico de orden superior 47'.

60 **[207]** En estos y otros casos, el dispositivo de decodificación de audio 24 puede configurarse para emitir la trama mediante un protocolo de transmisión por flujo.

65 **[208]** Varios aspectos de las técnicas también pueden permitir que el dispositivo de decodificación de audio 24 se configure para determinar, usando un flujo de bits 21 que incluye un coeficiente ambisónico de orden superior 47', si una trama para el coeficiente ambisónico de orden superior 47' es una trama independiente que incluye información de referencia adicional (por ejemplo, la información de estado 812) para permitir que la

trama sea decodificada y reproducida sin referencia a las tramas previas 810A a 810D del coeficiente ambisónico de orden superior 47'. El dispositivo de decodificación de audio 24 también puede configurarse para obtener, a partir del flujo de bits 21 y solo en respuesta a la determinación de que la trama no es una trama independiente, información de predicción (por ejemplo, a partir de la información de estado 812) para la trama, para decodificar la trama con referencia a una trama anterior para el coeficiente ambisónico de orden superior 47'.

[209] En estos y otros casos, el dispositivo de decodificación de audio 24 puede configurarse para obtener una señal basada en vectores, representativa de uno o más componentes distintos del campo de sonido que incluye un elemento de un vector 55_k correspondiente al coeficiente ambisónico de orden superior 47'. El vector 55_k puede describir aspectos espaciales de un componente distinto del campo de sonido y puede haber sido descompuesto a partir de coeficientes ambisónicos de orden superior 11, descriptivos del campo de sonido. El dispositivo de decodificación de audio 24 también puede configurarse para decodificar la señal basada en vectores usando la información de predicción.

[210] En estos y otros casos, el dispositivo de decodificación de audio 24 puede configurarse para obtener, usando el flujo de bits 21 y si la trama es una trama independiente, información de cuantización (por ejemplo, a partir de la información de estado 812) para la trama, suficiente para habilitar la trama para ser decodificada y reproducida sin referencia a la información de cuantización para tramas previas. El dispositivo de decodificación de audio 24 también puede configurarse para obtener, usando el flujo de bits 21 y si la trama no es una trama independiente, información de cuantización para la trama, que es insuficiente para permitir que la trama sea decodificada y reproducida sin referencia a la información de cuantización para las tramas anteriores. El dispositivo de decodificación de audio 24 también puede configurarse para decodificar la trama usando la información de cuantización.

[211] En estos y otros casos, la información de cuantización para la trama incluye un elemento sintáctico Nbits para la trama, suficiente para permitir que la trama sea decodificada y reproducida sin referencia a la información de cuantización para las tramas anteriores.

[212] En estos y otros casos, el dispositivo de decodificación de audio 24 puede configurarse para emitir la trama mediante un protocolo de transmisión por flujo.

[213] Diversos aspectos de las técnicas pueden permitir adicionalmente que el dispositivo de decodificación de audio 24 se configure para determinar, usando un flujo de bits 21 que incluye un coeficiente ambisónico de orden superior 47', que una trama para el coeficiente ambisónico de orden superior 47' es una trama independiente que incluye información de referencia adicional (por ejemplo, la información de estado 812), para permitir que la trama sea decodificada y reproducida sin referencia a tramas previas.

[214] En estos y otros casos, cuando se determina que la trama para el canal ambisónico de orden superior es una trama independiente, el dispositivo de decodificación de audio 24 puede obtener, usando el flujo de bits 21, un elemento sintáctico IndicadorIndependencia que indica que la trama es una trama independiente.

[215] La FIG. 7J es un diagrama que ilustra tramas ejemplares para uno o más canales de al menos un flujo de bits, de acuerdo a las técnicas descritas en el presente documento. El flujo de bits 450 incluye las tramas 810A a 810H que pueden incluir, cada una, uno o más canales. El flujo de bits 450 puede representar cualquier combinación de flujos de bits 21 mostrados en los ejemplos de las FIGs. 7A a 7H). El flujo de bits 450 puede ser esencialmente similar al flujo de bits 808, excepto porque el flujo de bits 450 no incluye las IPF. Como resultado, el dispositivo de decodificación de audio 24 mantiene información de estado, actualizando la información de estado para determinar cómo decodificar la trama actual k. El dispositivo de decodificación de audio 24 puede utilizar información de estado de la configuración 814 y de las tramas 810B a 810D. La diferencia entre la trama 810E y la IPF 816 es que la trama 810E no incluye la información de estado precedente, mientras que la IPF 816 incluye la información de estado precedente.

[216] En otras palabras, el dispositivo de codificación de audio 20 puede incluir, dentro de la unidad de generación de flujos de bits 42, por ejemplo, la máquina de estados 402 que mantiene información de estado para codificar cada una de las tramas 810A a 810E, en cuanto a que la unidad de generación de flujos de bits 42 puede especificar elementos sintácticos para cada una de las tramas 810A a 810E basándose en la máquina de estados 402.

[217] El dispositivo de decodificación de audio 24 puede incluir asimismo, dentro de la unidad de extracción de flujos de bits 72, por ejemplo, una máquina de estados 402 similar que emite elementos sintácticos (algunos de los cuales no están especificados explícitamente en el flujo de bits 21) basándose en la máquina de estados 402. La máquina de estados 402 del dispositivo de decodificación de audio 24 puede funcionar de una manera similar a la de la máquina de estados 402 del dispositivo de codificación de audio 20. Como tal, la máquina de estados 402 del dispositivo de decodificación de audio 24 puede mantener información de estado, actualizando la información de estado basándose en la configuración 814 y, en el ejemplo de la FIG. 7J, la decodificación de

las tramas 810B a 810D. En base a la información de estado, la unidad de extracción de flujos de bits 72 puede extraer la trama 810E basándose en la información de estado mantenida por la máquina de estados 402. La información de estado puede proporcionar una cantidad de elementos sintácticos implícitos que el dispositivo de codificación de audio 20 puede utilizar cuando decodifica los diversos canales de transporte de la trama 810E.

[218] La FIG. 8 es un diagrama de flujo que ilustra los canales de audio 800A a 800E, a los cuales un dispositivo de codificación de audio, tal como el dispositivo de codificación de audio 24 que se muestra en la FIG. 4, puede aplicar las técnicas descritas en esta divulgación. Como se muestra en el ejemplo de la FIG. 8, el canal de trasfondo 800A representa los coeficientes HOA ambientales que son el cuarto de los $(n + 1)^2$ posibles coeficientes HOA. Los canales de primer plano 800B y 800D representan un primer vector V y un segundo vector V , respectivamente. El canal de fondo 800C representa coeficientes HOA ambientales que son el segundo de los $(n + 1)^2$ posibles coeficientes HOA. El canal de fondo 800E representa coeficientes HOA ambientales que son el quinto de los $(n + 1)^2$ posibles coeficientes HOA.

[219] Como se muestra adicionalmente en el ejemplo de la FIG. 8, el coeficiente HOA ambiental 4 en el canal de trasfondo 800A experimenta un período de transición (desaparece) durante la trama 13, mientras que los elementos de un vector en el canal de primer plano 800D se incorporan durante la trama 14 para reemplazar el coeficiente HOA ambiental 4 en el canal de trasfondo 800A durante la decodificación del flujo de bits. La referencia al término "reemplazar" en el contexto de uno de los canales 800A a 800E que reemplaza a otro de los canales 800A a 800E, se refiere al ejemplo en el que el dispositivo de codificación de audio 20 genera el flujo de bits 21 para tener canales de transporte flexibles.

[220] Para ilustrar, cada una de las tres filas en la FIG. 8 puede representar un canal de transporte. Cada uno de los canales de transporte puede denominarse canal de trasfondo o canal de primer plano, según el tipo de datos de audio codificados que el canal de transporte está especificando actualmente. Por ejemplo, cuando el canal de transporte está especificando uno de los mínimos coeficientes HOA ambientales o un coeficiente HOA ambiental adicional, el canal de transporte puede denominarse canal de trasfondo. Cuando el canal de transporte está especificando un vector V , el canal de transporte puede denominarse canal de primer plano. El canal de transporte puede, por lo tanto, referirse a canales tanto de trasfondo como de primer plano. El canal de primer plano 800D puede, a este respecto, describirse como reemplazando el canal de trasfondo 800A en la trama 14 del primer canal de transporte. El canal de trasfondo 800E también se puede describir como reemplazando el canal de trasfondo 800C en la trama 13 en el tercer canal de transporte. Aunque se describe con respecto a tres canales de transporte, el flujo de bits 21 puede incluir cualquier número de canales de transporte, incluyendo entre cero canales de transporte y dos, tres o incluso más canales de transporte. Por consiguiente, las técnicas no deberían estar limitadas a este respecto.

[221] En particular, el ejemplo de la FIG. 8 también muestra generalmente los elementos del vector del cambio de canal de primer plano 800B en las tramas 12, 13 y 14, tal como se describe con más detalle a continuación, y la longitud del vector cambia durante las tramas. El coeficiente HOA ambiental 2 en el canal de trasfondo 800C experimenta una transición durante la trama 12. El canal de trasfondo 800E del coeficiente HOA ambiental 5 experimenta una transición (se incorpora) durante la trama 13 para reemplazar el coeficiente HOA ambiental 2 en el canal de trasfondo 800C durante la decodificación del flujo de bits.

[222] Durante los periodos de transición descritos anteriormente, el dispositivo de codificación de audio 20 puede especificar el indicador TransiciónCoefAmb 757 en el flujo de bits con un valor de uno para cada uno de los canales 800A, 800C, 800D y 800E, para indicar que cada uno de los canales ambientales respectivos 800A, 800C y 800E están en transición en las tramas respectivas 13, 12 y 13. Dado el estado anterior de la ModalidadTransiciónCoefAmb, el dispositivo de codificación de audio 20 puede por lo tanto proporcionar el indicador TransiciónCoefAmb 757 al dispositivo de decodificación de audio 24 a fin de indicar que el coeficiente respectivo está en transición de salida (o, en otras palabras, siendo eliminado) del flujo de bits o en transición de entrada (o, en otras palabras, siendo incorporado) al flujo de bits.

[223] El dispositivo de decodificación de audio 24 puede funcionar entonces como se ha expuesto anteriormente para identificar los canales 800 en el flujo de bits y realizar la operación de incorporación o eliminación progresiva, tal como se analiza a continuación en más detalle.

[224] Además, como resultado del fundido inicial y del fundido final de los diversos canales ambientales 800A, 800C y 800E, en cierta cuantización vectorial, el dispositivo codificador de audio 20 puede especificar el vector V en los canales de primer plano 800B y 800D usando una número reducido de elementos, tal como se ha descrito anteriormente con respecto al dispositivo de codificación de audio 20 mostrado en el ejemplo de la FIG. 3. El dispositivo de decodificación de audio 24 puede funcionar con respecto a cuatro modalidades de reconstrucción diferentes, una de las cuales puede implicar la reducción de los elementos del vector V cuando la energía de ese elemento se ha incorporado al coeficiente HOA ambiental subyacente. Lo anterior puede estar representado generalmente por el siguiente pseudo-código:

```

%% memoria intermedia de relleno desde la trama de audio
memIntVecPrimerPlano (:, canalesTransporteConComponentesDistintos) =
tramaAudio (:, canalesTransporteConComponentesDistintos);
%% 1. Reconstruir componentes distintos recientemente introducidos (si los hay)
5 si ~ estávacío (nuevosCanalesTransporteConComponentesDistintos)
    memIntInterpVecPrimerPlano =
        memIntVecPrimerPlano (1: longitudInterp,
            nuevosCanalesTransporteConComponentesDistintos) *
            memIntV (nuevosCanalesTransporteConComponentesDistintos, :);
10 fin
%% 2. reconstruir componentes continuos distintos (si los hay) y aplicar interpolación
espacio-temporal
si ~ estávacío (canalesComunesTransporteConComponentesDistintos)
    para indCanUi =
15         canalesTransporteConComponentesDistintos (
            canalesComunesTransporteConComponentesDistintos)
            viejoHOA = memIntVecPrimerPlano (1: longitudInterp, indCanUi) *
                memIntV_tramaPrev (indCanUi, :);
            nuevoHOA = memIntVecPrimerPlano (1: longitudInterp, indCanUi) * memIntV
20             (indCanUi, :);
            memIntInterpVecPrimerPlano = memIntInterpVecPrimerPlano + (viejoHOA. *
                eliminaciónCruzada) +
                (nuevoHOA. * incorporaciónCruzada);
    fin
25 fin
tramaHoaReconstruida (indInicial: indInicial + longitudInterp-1, :) =
memIntInterpVecPrimerPlano;
tramaHoaReconstruida (indInicial + longitudInterp: indFinal,:) =
memIntVecPrimerPlano (longitudInterp + 1: fin,
30 canalesTransporteConComponentesDistintos) * memIntV
(canalesTransporteConComponentesDistintos, :);
% comprobar si hay coeficientes HOA ambientales de transición presentes en la trama,
aplicando fundido inicial / fundido final
si ~ estávacío (canalesTransporteConHoaIncorporado)
35     para canalTransiciónUi =
        indCoefAmb (canalesTransporteConHoaIncorporado)
            tramaHoaReconstruida (:, canalTransiciónUi) =
                tramaHoaReconstruida (:, canalTransiciónUi). *
                    ventanaEliminaciónCuandoIncorporaciónCanalHoa;
40     fin
fin
si ~ estávacío (canalesTransporteConEliminaciónHoa)
    para canalTransiciónUi =
        indCoefAmb (canalesTransporteConEliminaciónHoa)
45         tramaHoaReconstruida (:, canalTransiciónUi) =
            tramaHoaReconstruida (:, canalTransiciónUi). * =
                ventanaIncorporaciónCuandoCanalHoaEliminación;
    fin
50 fin
%% 3. agregar coeficientes HOA ambientes por omisión
tramaHoaReconstruida (:, 1: estadoDescompresión.NúmMínDeCoefsParaHOAAmbiental) =
tramaAudio (:, NúmDeCodificadoresPerceptualesAdicionales + 1: end);
%% 4. agregar coeficientes HOA ambientales dependientes de la trama
tramaHoaReconstruida (:, canalesHoaAmbAdic) =
55 tramaHoaReconstruida (:, canalesHoaAmbAdic) +
tramaAudio (:, canalesTransporteConHoaAmbientalAdic);

```

[225] El pseudo-código anterior tiene cuatro secciones diferentes o modalidades de reconstrucción de funcionamiento, indicadas por comentarios (que comienzan con el signo de porcentaje ("%")) seguido por el número 1 a 4. La primera sección para la primera modalidad de reconstrucción proporciona pseudo-código para reconstruir componentes distintos recientemente introducidos cuando están presentes. La segunda sección para la segunda modalidad de reconstrucción proporciona pseudo-código para reconstruir componentes continuos distintos cuando están presentes y para aplicar interpolación espacio-temporal. En la sección dos del pseudo-código, hay operaciones de incorporación cruzada y eliminación cruzada, realizadas en la memoria intermedia de interpolación de vectores V en primer plano (memIntInterpVecPrimerPlano) para incorporar nuevos coeficientes HOA y eliminar viejos coeficientes HOA, congruentes con varios aspectos de las técnicas

descritas en esta divulgación. La tercera sección para la tercera modalidad de reconstrucción proporciona un pseudo-código para agregar coeficientes HOA ambientales predeterminados. La cuarta sección para la cuarta modalidad de reconstrucción proporciona un pseudo-código para añadir coeficientes HOA dependientes de la trama, congruentes con diversos aspectos de las técnicas descritas en esta divulgación.

[226] En otras palabras, para reducir el número de elementos del vector V transmitidos, solo pueden transmitirse los elementos del campo de sonido HOA que no estén codificados como coeficientes HOA ambientales. En algunos casos, el número total o los coeficientes HOA reales de los componentes ambientales pueden ser dinámicos para tener en cuenta los cambios en el campo de sonido codificado. Sin embargo, para las veces en que un canal de trasfondo, que incluye los coeficientes HOA ambientales, se incorpora o se elimina, puede haber una distorsión notable debido al cambio en la energía.

[227] Por ejemplo, con referencia a la FIG. 8, en la trama 10 y 11 hay dos canales de trasfondo 800A y 800C y un canal de primer plano 800B. En las tramas 10 y 11, el vector V especificado en el canal de primer plano 800B puede no incluir los coeficientes de conversión a sonido multicanal para los coeficientes HOA ambientales 47' especificados en los canales de trasfondo 800A y 800C porque los coeficientes HOA ambientales 47' especificados en los canales de trasfondo 800A y 800C se pueden codificar directamente. En la trama 12, el coeficiente HOA ambiental 47' especificado en el canal de trasfondo 800C se está eliminando en este ejemplo. En otras palabras, el dispositivo de decodificación de audio 24 puede eliminar el coeficiente HOA ambiental 47' especificado en el canal de trasfondo 800C usando cualquier tipo de fundido, tal como el fundido lineal mostrado en la FIG. 8. Es decir, aunque se muestra como un fundido inicial lineal, el dispositivo de decodificación de audio 24 puede realizar cualquier forma de operaciones de fundido inicial, que incluyen operaciones de fundido inicial no lineales (por ejemplo, una operación de fundido inicial exponencial). En la trama 13, el coeficiente HOA ambiental 47' especificado en el canal de trasfondo 800A está, en este ejemplo, siendo eliminado y el coeficiente HOA ambiental 47' especificado en el canal de trasfondo 800E está, en este ejemplo, siendo incorporado. El flujo de bits 21 puede señalar los sucesos cuando un coeficiente HOA ambiental 47' especificado en un canal de trasfondo se elimina o incorpora, como se ha descrito anteriormente. El dispositivo de decodificación de audio 24 puede realizar de manera similar cualquier forma de operación de fundido final que incluya la operación de fundido inicial lineal mostrada en el ejemplo de la FIG. 8 y operaciones de fundido final no lineales.

[228] En el ejemplo de la FIG. 8, el dispositivo de codificación de audio 20 puede mantener información de estado que indique un estado de transición para cada coeficiente HOA ambiental especificado en uno de los tres canales de transporte mostrados en la FIG. 8, como se ha descrito anteriormente. Para el canal de trasfondo 800A, el dispositivo de codificación de audio 20 puede mantener el elemento sintáctico `CoefAmbFuelIncorporado [i]` ("FuelIncorporado [i]") (que también se puede indicar como un elemento de estado), el elemento sintáctico `ModalidadTransiciónCoefAmb [i]` ("ModalidadTransición [i]") (que también se puede indicar como un elemento de estado) y un elemento sintáctico `TransiciónCoefAmb` ("Transición"). Los elementos de estado `FuelIncorporado [i]` y `ModalidadTransición [i]` pueden indicar un estado dado del coeficiente HOA ambiental especificado en el canal 800A. Hay tres estados de transición, como se ha esbozado anteriormente en la tabla de sintaxis `CanallInfoHOAAmbAdic (i)`. El primer estado de transición es sin transición, que se representa mediante el elemento de estado `ModalidadTransiciónCoefAmb [i]` que se fija en cero (0). El segundo estado de transición es el fundido inicial de un coeficiente HOA ambiental adicional, que se representa mediante el elemento de estado `ModalidadTransiciónCoefAmb [i]` que se fija en uno (1). El tercer estado de transición es el fundido final del coeficiente HOA ambiental adicional, que se representa mediante el elemento de estado `ModalidadTransiciónCoefAmb [i]` que se fija en dos (2). El dispositivo de codificación de audio 20 usa el elemento de estado `FuelIncorporado [i]` para actualizar el elemento de estado `ModalidadTransición [i]`, de nuevo, como se ha esbozado anteriormente en la tabla de sintaxis `CanallInfoHOAAmbAdic (i)`.

[229] El dispositivo de decodificación de audio 24 puede asimismo mantener el elemento sintáctico `CoefAmbFuelIncorporado [i]` ("FuelIncorporado [i]") (que también se puede indicar como un elemento de estado), el elemento sintáctico `ModalidadTransiciónCoefAmb [i]` ("ModalidadTransición [i]") (que también se puede indicar como un elemento de estado) y un elemento sintáctico `TransiciónCoefAmb` ("Transition"). De nuevo, los elementos de estado `FuelIncorporado [i]` y `ModalidadTransición [i]` pueden indicar un estado dado del coeficiente HOA ambiental especificado en el canal 800A. La máquina de estados 402 (según se representa en la FIG. 7J) en el dispositivo de decodificación de audio 24 puede asimismo configurarse en uno de los tres estados de transición, como se ha esbozado anteriormente en las tablas ejemplares de sintaxis `CanallInfoHOAAmbAdic (i)`. De nuevo, el primer estado de transición es sin transición, que se representa mediante el elemento de estado `ModalidadTransiciónCoefAmb [i]` que se fija en cero (0). El segundo estado de transición es el fundido inicial de un coeficiente HOA ambiental adicional, que se representa mediante el elemento de estado `ModalidadTransiciónCoefAmb [i]` que se fija en uno (1). El tercer estado de transición es el fundido final del coeficiente HOA ambiental adicional, que se representa mediante el elemento de estado `ModalidadTransiciónCoefAmb [i]` que se fija en dos (2). El dispositivo de decodificación de audio 24 usa el elemento de estado `FuelIncorporado [i]` para actualizar de nuevo el elemento de estado `ModalidadTransición [i]`, como se ha esbozado anteriormente en la tabla de sintaxis `CanallInfoHOAAmbAdic (i)`.

[230] Con referencia de nuevo al canal de trasfondo 800A, el dispositivo de codificación de audio 20 puede mantener información de estado (por ejemplo, la información de estado 812 mostrada en el ejemplo de la FIG. 7J), en la trama 10, indicando que el elemento de estado Fuelncorporado [i] está fijado en uno y el elemento de estado ModalidadTransición [i] está fijado en cero, donde i indica el índice asignado al coeficiente HOA ambiental. El dispositivo de codificación de audio 20 puede mantener la información de estado 812 con el fin de determinar los elementos sintácticos (TransiciónCoefAmb y, para las tramas de reproducción inmediata, Fuelncorporado [i] o la alternativa TransiciónÍndCoefAmb y, para las tramas de reproducción inmediata, EstadoTransiciónCoefAmb [i]), que son enviados para permitir que el dispositivo de decodificación de audio 24 realice las operaciones de fundido inicial o fundido final con respecto a los coeficientes HOA ambientales y los elementos del vector V de los canales en primer plano. Aunque se describe como que mantiene la información de estado 812 con el fin de generar y especificar los elementos sintácticos adecuados, el dispositivo de codificación de audio 20 también puede realizar las técnicas para hacer efectivamente la transición de los elementos, eliminando potencialmente de ese modo la realización de una operación adicional en el dispositivo decodificador de audio 24, y facilitar una decodificación más eficaz (en términos de eficacia energética, ciclos de procesador, etc.).

[231] El dispositivo de codificación de audio 20 puede entonces determinar si fue especificado el mismo coeficiente HOA 4 en la trama 9 anterior (no mostrado en el ejemplo de la FIG. 8). Cuando se especifica, el dispositivo de codificación de audio 20 puede especificar el elemento sintáctico de transición en el flujo de bits 21 con un valor cero. El dispositivo de codificación de audio 20 también puede mantener la información de estado 812 para el canal 800C, que es la misma que la especificada para el canal 800A. Como resultado de especificar dos coeficientes HOA ambientales 47' que tienen un índice 2 y 4 mediante los canales 800C y 800A, el dispositivo de codificación de audio 20 puede especificar un vector V ("VecV") que tiene un total de 23 elementos (para el orden N = 4, que es $(4 + 1)^2 - 2$, o 25-2, para determinar los 23 elementos). El dispositivo de codificación de audio 20 puede especificar los elementos [1, 3, 5:25], omitiendo los elementos que corresponden a los coeficientes HOA ambientales 47' que tienen un índice de 2 y 4. Dado que no se producen transiciones hasta la trama 12, el dispositivo de codificación de audio 20 mantiene la misma información de estado para los canales 800A y 800C durante la trama 11.

[232] El dispositivo de decodificación de audio 24 puede mantener similarmente información de estado (por ejemplo, la información de estado 812 mostrada en el ejemplo de la FIG. 7J), en la trama 10, que indica que el elemento de estado Fuelncorporado [i] está fijado en uno y el elemento de estado ModalidadTransición [i] está fijado en cero. El dispositivo de decodificación de audio 24 puede mantener la información de estado 812 con el fin de minimizar la transición adecuada en función de los elementos sintácticos (TransiciónCoefAmb) que se envían en el flujo de bits 21. En otras palabras, el dispositivo de decodificación de audio 24 puede invocar a la máquina de estados 402 para actualizar la información de estado 812 en base a los elementos sintácticos especificados en el flujo de bits 21. La máquina de estados 812 puede pasar de uno de los tres estados de transición indicados anteriormente a otro de los tres estados, basándose en los elementos sintácticos, como se ha descrito con más detalle anteriormente con respecto a las tablas ejemplares de sintaxis CanalInfoHOAAmbAdic (i). En otras palabras, según el valor del elemento sintáctico TransiciónCoefAmb señalado en el flujo de bits y la información de estado 812, la máquina de estados 402 del dispositivo de decodificación de audio 24 puede conmutar entre los estados sin transición, fundido final y fundido inicial, como se describe a continuación con respecto a las tramas ejemplares 12, 13 y 14.

[233] El dispositivo de decodificación de audio 24 puede por lo tanto obtener los coeficientes HOA ambientales 47' que tienen un índice de 4 mediante el canal de trasfondo 800A en las tramas 10 y 11. El dispositivo de decodificación de audio 24 también puede obtener el coeficiente HOA ambiental 47' que tiene un índice de 2 mediante el canal de trasfondo 800C en las tramas 10 y 11. El dispositivo de decodificación de audio 24 puede obtener, durante la trama 10 y para cada uno de los coeficientes HOA ambientales 47' que tienen un índice de 2 y 4, una indicación indicativa de si los coeficientes HOA ambientales 47' que tienen un índice de 2 y 4 están en transición durante la trama 10. La máquina de estados 402 del dispositivo de decodificación de audio 24 puede mantener además la información de estado 812 para el coeficiente HOA ambiental 47' que tiene un índice de 2 en forma de los elementos de estado Fuelncorporado [2] y ModalidadTransición [2]. La máquina de estados 402 del dispositivo de decodificación de audio 24 puede mantener además la información de estado 812 para el coeficiente HOA ambiental 47' que tiene un índice de 4 en forma de los elementos de estado Fuelncorporado [4] y ModalidadTransición [4]. Dado que la información de estado para los coeficientes HOA ambientales 47' que tienen el índice de 2 y 4 indica que los coeficientes 47' están en un estado sin transición y, basándose en la indicación de Transición que indica que los coeficientes HOA ambientales 47' que tienen un índice de 2 y 4 no están en transición durante cualquiera de las tramas 10 u 11, el dispositivo de decodificación de audio 24 puede determinar que el vector reducido 55_k especificado en el canal de primer plano 800B incluye elementos vectoriales [1, 3, 5:23] y omite los elementos que corresponden a los coeficientes HOA ambientales 47' que tienen un índice de 2 y 4 para ambas tramas 10 y 11. El dispositivo de decodificación de audio 24 puede obtener entonces el vector reducido 55_k del flujo de bits 21 para las tramas 10 y 11 mediante, como un ejemplo, el análisis sintáctico correcto de los 23 elementos del vector reducido 55_k .

[234] En la trama 12, el dispositivo de codificación de audio 20 determina que el coeficiente HOA ambiental que tiene un índice de 2 transportado por el canal 800C debe ser eliminado. Como tal, el dispositivo de codificación de audio 20 puede especificar un elemento sintáctico de transición en el flujo de bits 21 para el canal 800C con un valor de uno (que indica la transición). El dispositivo de codificación de audio 20 puede actualizar los elementos de estado interno FueIncorporado [2] y ModalidadTransición [2] para el canal 800C para que sean cero y dos, respectivamente. Como resultado del cambio de estado, de ninguna transición a fundido final, el dispositivo de codificación de audio 20 puede añadir un elemento de vector V al vector V especificado en el canal de primer plano 800B, correspondiente al coeficiente HOA ambiental 47' que tiene un índice de 2.

[235] El dispositivo de decodificación de audio 24 puede invocar a la máquina de estados 402 para actualizar la información de estado 812 para el canal 800C. La máquina de estados 402 puede actualizar los elementos de estado interno FueIncorporado [2] y ModalidadTransición [2] para el canal 800C, para que sean cero y dos, respectivamente. En base a la información de estado actualizada 812, el dispositivo de decodificación de audio 24 puede determinar que el coeficiente HOA ambiental 47' que tiene un índice de 2 sea eliminado durante la trama 12. El dispositivo de decodificación de audio 24 puede determinar además que el vector reducido 55_k para la trama 12 incluye un elemento adicional correspondiente a los coeficientes HOA ambientales 47' que tienen un índice de 2. El dispositivo de decodificación de audio 24 puede entonces incrementar el número de elementos vectoriales para el vector reducido 55_k especificado en el canal de primer plano 800B, para reflejar el elemento vectorial adicional (lo que se indica en el ejemplo de la FIG. 8 como que los elementos de VecV son iguales a 24 en la trama 12). El dispositivo de decodificación de audio 24 puede entonces obtener el vector reducido 55_k especificado mediante el canal de primer plano 800B, basándose en el número actualizado de elementos vectoriales. El dispositivo de decodificación de audio 24, después de obtener el vector reducido 55_k , puede incorporar el elemento adicional de VecV 2 (indicado como "Vec-V [2]") durante la trama 12. En la trama 13, el dispositivo de codificación de audio 20 indica dos transiciones, una para señalar que el coeficiente HOA 4 está en transición, o siendo eliminado, y otra para indicar que el coeficiente HOA 5 está en transición, o siendo incorporado, al canal 800C. Si bien el canal en realidad no cambia, a los efectos de indicar el cambio en lo que el canal está especificando, el canal se puede indicar como el canal 800E después de la transición.

[236] En otras palabras, el dispositivo de codificación de audio 20 y el dispositivo de decodificación de audio 24 pueden mantener la información de estado por cada canal de transporte. Como tal, el canal de trasfondo 800A y el canal de primer plano 800D son transportados por el mismo de los tres canales de transporte, mientras que los canales de trasfondo 800C y 800E también son transportados por el mismo de los tres canales de transporte. En cualquier caso, el dispositivo de codificación de audio 20 puede mantener información de estado de transición para el canal de trasfondo 800E indicando que los coeficientes HOA ambientales 47' que tienen un índice de 5, y especificados mediante el canal de trasfondo 800E están incorporados (por ejemplo, FueIncorporado [5] = 1) y que la modalidad de transición es el fundido inicial (por ejemplo, ModalidadTransición [5] = 1). El dispositivo de codificación de audio 20 también puede mantener información de estado de transición para el canal 800A indicando que el coeficiente HOA ambiental que tiene un índice de 4 ya no está incorporado (por ejemplo, FueIncorporado [4] = 0) y que la modalidad de transición es la eliminación (por ejemplo, ModalidadTransición [4] = 2).

[237] El dispositivo de decodificación de audio 24 puede mantener nuevamente la información de estado 812 similar a la descrita anteriormente con respecto al dispositivo de codificación de audio 20 y, basándose en la información de estado actualizada, eliminar el coeficiente HOA ambiental 47' que tiene un índice de 4, mientras incorpora el coeficiente HOA ambiental 47' que tiene un índice de 5. En otras palabras, el dispositivo de decodificación de audio 24 puede obtener el elemento sintáctico de Transición para el canal 800A durante la trama 13, indicando que el coeficiente HOA ambiental 47' que tiene un índice 4 está en transición. El dispositivo de decodificación de audio 24 puede invocar a la máquina de estados 402 para procesar el elemento sintáctico Transición para actualizar los elementos sintácticos FueIncorporado [4] y ModalidadTransición [4], para indicar que el coeficiente HOA ambiental 47', que tiene un índice de 4, ya no se incorpora (por ejemplo, FueIncorporado [4] = 0) y que la modalidad de transición es el fundido final (por ejemplo, ModalidadTransición [4] = 2).

[238] El dispositivo de decodificación de audio 24 también puede obtener el elemento sintáctico Transición para el canal 800C durante la trama 13, indica que el coeficiente HOA ambiental 47' que tiene un índice 5 está en transición. El dispositivo de decodificación de audio 24 puede invocar a la máquina de estados 402 para procesar el elemento sintáctico Transición para actualizar los elementos sintácticos FueIncorporado [5] y ModalidadTransición [5], para indicar que el coeficiente HOA ambiental 47' que tiene un índice de 4 se incorpora durante la trama 13 (por ejemplo, FueIncorporado [5] = 1) y que la modalidad de transición es de fundido inicial (por ejemplo, ModalidadTransición [5] = 1). El dispositivo de decodificación de audio 24 puede realizar una operación de fundido final con respecto al coeficiente HOA ambiental 47' que tiene un índice de 4 y una operación de fundido inicial con respecto al coeficiente HOA ambiental 47' que tiene un índice de 5.

[239] Sin embargo, el dispositivo de decodificación de audio 24 puede utilizar un vector V completo (suponiendo nuevamente una representación de cuarto orden) que tenga 25 elementos, de modo que el $\text{Vec}V$ [4] pueda incorporarse y el $\text{Vec}V$ [5] pueda eliminarse. El dispositivo de codificación de audio 20 puede por lo tanto proporcionar un $\text{Vec}V$ en el canal de primer plano 800B que tiene 25 elementos.

[240] Dado que hay tres canales de transporte, dos de los cuales se someten a una transición, siendo el restante de los tres canales de transporte el canal de primer plano 800B, el dispositivo de decodificación de audio 24 puede determinar que el vector reducido 55_k puede, en la situación ejemplar, incluir todos los 24 elementos del vector. Como resultado, el dispositivo de decodificación de audio 24 puede obtener el vector reducido 55_k del flujo de bits 21 que tenga todos los 25 elementos vectoriales. El dispositivo de decodificación de audio 24 puede incorporar luego, durante la trama 13, el elemento vectorial del vector reducido 55_k asociado al coeficiente HOA ambiental $47'$ que tiene un índice de 4 para compensar la pérdida de energía. El dispositivo de decodificación de audio 24 puede eliminar luego, durante la trama 13, el elemento vectorial del vector reducido 55_k asociado al coeficiente HOA ambiental $47'$ que tiene un índice de 5 para compensar la ganancia de energía.

[241] En la trama 14, el dispositivo de codificación de audio 20 puede proporcionar otro vector V que reemplaza el canal de trasfondo 800A en el canal de transporte, que puede especificarse en el canal de primer plano 800D. Dado que no hay transiciones de coeficientes HOA ambientales, el dispositivo de codificación de audio 20 puede especificar los vectores V en los canales de primer plano 800D y 800B con 24 elementos, dado que el elemento correspondiente al coeficiente HOA ambiental $47'$ que tiene un índice de 5 no necesita ser enviado (como resultado de enviar el coeficiente HOA ambiental $47'$ que tiene un índice de 5 en el canal de trasfondo 800E). La trama 14 puede, a este respecto, indicarse como una trama posterior a la trama 13. En la trama 14, el coeficiente HOA ambiental $47'$ se especifica en el canal de trasfondo 800E y no está en transición. Como resultado, el dispositivo de codificación de audio 20 puede retirar el elemento del vector V correspondiente a los coeficientes HOA ambientales $47'$, especificados en el canal de trasfondo 800E a partir del vector reducido 55_k especificado en el canal de primer plano 800B, generando de este modo un vector V actualizado reducido (que tiene 24 elementos en lugar de los 25 elementos en la trama anterior).

[242] El dispositivo de decodificación de audio 24 puede, durante la trama 14, invocar la máquina de estados 402 para actualizar la información de estado 812 para indicar que el coeficiente HOA ambiental $47'$ que tiene un índice de 5, y especificado mediante el canal de trasfondo 800E, no está en transición ("ModalidadTransición" [5] = 0 ") y fue incorporado previamente ("FueIncorporado" [5] = 1"). Como resultado, el dispositivo de decodificación de audio 24 puede determinar que los vectores reducidos 55_k , especificados en el canal de primer plano 800D y 800B, tienen 24 elementos vectoriales (ya que el elemento vectorial asociado al coeficiente HOA ambiental $47'$ que tiene un índice de 5 no está especificado). Sin embargo, el dispositivo de decodificación de audio 24 puede incorporar todos los elementos vectoriales del vector reducido 55_k especificado en el canal de primer plano 800D durante la trama 14, ya que los elementos no se especificaron previamente en el flujo de bits en la trama precedente.

[243] En la trama 15, el dispositivo de codificación de audio 20 y el dispositivo de decodificación de audio 24 mantienen el mismo estado que en la trama 14, dado, una vez más, que no se han producido transiciones.

[244] A este respecto, las técnicas pueden permitir que el dispositivo de codificación de audio 20 se configure para determinar cuándo un coeficiente ambisónico ambiental de orden superior $47'$ (como se especifica, por ejemplo, en el canal de trasfondo 800C) está en transición durante una trama de un flujo de bits 21 (como se ha mostrado por primera vez en las FIGs. 3 y 4 y posteriormente esclarecido en la FIG. 8), representativo de los datos de audio codificados (que pueden referirse a cualquier combinación de los coeficientes HOA ambientales, los objetos de audio en primer plano y los correspondientes vectores V), el coeficiente ambisónico de orden superior ambiental $47'$, al menos en parte, de un componente ambiental de un campo de sonido. El dispositivo de codificación de audio 20 también puede configurarse para identificar un elemento de un vector (tal como uno de los restantes vectores $V[k]$ de primer plano) que está asociado al coeficiente ambisónico de orden superior ambiental $47'$ en transición. El vector 53 puede ser representativo, al menos en parte, de un componente espacial del campo de sonido. El dispositivo de codificación de audio 20 puede además ser configurado para generar, basándose en el vector 53, un vector reducido 55 para incluir el elemento identificado del vector para la trama. Para ilustrar, considere el canal de primer plano 800B en la trama 12, donde el dispositivo de codificación de audio 20 genera el vector reducido 55 para incluir el elemento del vector V correspondiente al coeficiente HOA ambiental 2 especificado en el canal de trasfondo 800C en la trama 12, lo que se indica como $\text{Vec}V$ [2] en el ejemplo de la FIG. 8. El dispositivo de codificación de audio 20 también puede estar configurado para producir el flujo de bits 21 para incluir un bit indicativo del vector reducido y un bit (por ejemplo, una indicación 757 como se representa en la FIG. 4) indicativo de la transición del coeficiente ambisónico ambiental de orden superior $47'$ durante la trama.

[245] En estos y otros casos, el dispositivo de codificación de audio 20 puede configurarse para mantener información del estado de transición basándose en el coeficiente ambisónico de orden superior ambiental en transición. Por ejemplo, el dispositivo de codificación de audio 20 puede incluir la máquina de estados 402

mostrada en el ejemplo de la FIG. 71 que mantiene la información de estado de transición y cualquier otra información de estado 812. El dispositivo de codificación de audio 20 puede estar configurado además para obtener la indicación 757 de la transición basándose en la información de estado de transición.

5 **[246]** En estos y otros casos, la información de estado de transición indica uno entre un estado sin transición, un estado de fundido inicial y un estado de fundido final.

10 **[247]** En estos y otros casos, el dispositivo de codificación de audio 20 puede configurarse para producir el flujo de bits 21 para incluir adicionalmente un bit indicativo de la información de estado 812 que incluye la información de estado de transición en la trama. El bit indicativo de la información de estado 812 puede permitir que la trama sea decodificada sin referencia a tramas previas del flujo de bits 21.

[248] En estos y otros casos, la información de estado 812 incluye información de cuantización.

15 **[249]** En estos y otros casos, la trama se emite mediante un protocolo de transmisión por flujo.

[250] En estos y otros casos, el bit 757 indicativo de la transición especifica si el coeficiente ambisónico de orden superior ha de ser eliminado por un decodificador, tal como el dispositivo de decodificación de audio 24, durante la trama.

20 **[251]** En estos y otros casos, el bit indicativo de la transición especifica si el coeficiente ambisónico de orden superior ha de ser incorporado por un decodificador, tal como el dispositivo de decodificación de audio 24, durante la trama.

25 **[252]** En estos y en otros casos, el dispositivo de codificación de audio 20 puede configurarse para actualizar el vector reducido 55 eliminando un segundo elemento del vector 53 asociado al coeficiente ambisónico ambiental de orden superior 47' que no está en transición durante una trama posterior. Para ilustrar, considere la trama 14 donde el dispositivo de codificación de audio 20 actualiza el vector reducido 55 de la trama 13 para eliminar el elemento del vector reducido 55 de la trama 13 asociada al coeficiente HOA ambiental que tiene un índice de cinco (donde el elemento es indicado como "VecV[5j]"). El dispositivo de codificación de audio 20 puede configurarse adicionalmente para producir el flujo de bits 21, para incluir, durante la trama subsiguiente 14, un bit indicativo del vector reducido actualizado y un bit que indica que el coeficiente ambisónico de orden superior ambiental 47' que tiene un índice de 5 no está en transición.

35 **[253]** En estos y otros casos, el dispositivo 20 de codificación de audio puede configurarse para realizar los aspectos independientes de las técnicas descritas con más detalle anteriormente junto con los aspectos de transición de las técnicas descritas anteriormente.

40 **[254]** Además, los aspectos de transición de las técnicas pueden permitir que el dispositivo de decodificación de audio 24 se configure para obtener, a partir de una trama (por ejemplo, las tramas 10 a 15 en la FIG. 8) de un flujo de bits 21 representativo de los datos de audio codificados, un bit indicativo de un vector reducido. Los datos de audio codificados pueden incluir una versión codificada de los coeficientes HOA 11 o un subproducto de los mismos, es decir, como ejemplo, los coeficientes HOA ambientales codificados 59, las nFG señales codificadas 61, los vectores $V[k]$ en primer plano codificados 57 y cualquier elemento sintáctico acompañante o bits indicativos de cada uno de los precedentes. El vector reducido puede representar, al menos en parte, un componente espacial de un campo de sonido. El vector reducido puede referirse a uno de los vectores $V[k]$ en primer plano reducidos 55_k , descritos anteriormente con respecto al ejemplo de la FIG. 4. El dispositivo de decodificación de audio 24 puede configurarse además para obtener, desde la trama, un bit 757 (mostrado en la FIG. 4 y representado en el ejemplo de la FIG. 8 como el indicador de "Transición") indicativo de una transición de un coeficiente ambisónico 'de orden superior ambiental 47' (como se especifica, por ejemplo, en el canal 800C). El coeficiente ambisónico de orden superior ambiental 47' puede representar, al menos en parte, un componente ambiental de un campo de sonido. El vector reducido puede incluir un elemento vectorial asociado al coeficiente ambisónico ambiental de orden superior en transición, tal como en el ejemplo de la trama 13, donde el canal en primer plano 800B incluye el elemento 5 del vector V asociado al canal de trasfondo 800E. El vector reducido puede referirse a uno de los vectores $V[k]$ en primer plano reducidos 55_k y, como tal, se puede indicar como vector reducido 55_k .

50 **[255]** En estos y otros casos, el dispositivo de decodificación de audio 24 puede configurarse adicionalmente para obtener el bit indicativo del vector reducido 55_k , de acuerdo a la Modalidad 2 descrita anteriormente entre una pluralidad de modalidades (por ejemplo, Modalidad 0, Modalidad 1 y Modalidad 2). La modalidad 2 puede indicar que el vector reducido incluye el elemento vectorial asociado al coeficiente ambisónico de orden superior ambiental en transición.

65 **[256]** En estos y otros casos, la pluralidad de modalidades incluye además la Modalidad 1 descrita anteriormente. La modalidad 1 puede, como se ha descrito anteriormente, indicar que el elemento vectorial asociado al coeficiente ambisónico de orden superior ambiental no está incluido en el vector reducido.

5 [257] En estos y otros casos, el dispositivo de decodificación de audio 24 puede configurarse además para mantener información de estado de transición basándose en el bit 757 indicativo de la transición del coeficiente ambisónico de orden superior ambiental. La unidad de extracción de flujos de bits 72 del dispositivo de decodificación de audio 24 puede incluir la máquina de estados 402 para mantener la información de estado 812 que incluye la información de estado de transición. El dispositivo de decodificación de audio 24 también puede configurarse para determinar si se realiza una operación de fundido inicial o una operación de fundido final con respecto al coeficiente ambisónico ambiental de orden superior 47' del canal 800C, basándose en la información de estado de transición. El dispositivo de decodificación de audio 24 puede configurarse para invocar a la unidad de fundido 770 para llevar a cabo la operación de fundido inicial o la operación de fundido final, con respecto al coeficiente ambisónico ambiental de orden superior 47', basándose en la determinación de si se incorpora o se elimina el coeficiente ambisónico ambiental de orden superior.

15 [258] En estos y otros casos, la información de estado de transición indica uno entre un estado sin transición, un estado de fundido inicial y un estado de fundido final.

20 [259] En estos y otros casos, el dispositivo de decodificación de audio 24 puede configurarse además para obtener la información de estado de transición de un bit indicativo de la información de estado 812. La información de estado 812 puede permitir que la trama sea decodificada sin referencia a tramas previas del flujo de bits.

25 [260] En estos y otros casos, el dispositivo de decodificación de audio 24 puede configurarse además para descuantizar el vector reducido 55_k en base a la información de cuantización incluida en el bit indicativo de la información de estado 812.

[261] En estos y otros casos, la trama se emite mediante un protocolo de transmisión por flujo.

30 [262] En estos y otros casos, la indicación 757 de la transición especifica si el coeficiente ambisónico de orden superior 47' se elimina durante la trama.

[263] En estos y otros casos, la indicación 757 de la transición especifica si el coeficiente ambisónico de orden superior se incorpora durante la trama.

35 [264] En estos y otros casos, el dispositivo de decodificación de audio 24 puede configurarse además para obtener, durante una trama posterior (por ejemplo, la trama 14) del flujo de bits 21, un bit indicativo de un segundo vector reducido (que puede referirse al mismo vector que el especificado para la trama 13 en el canal de primer plano 800C, solo actualizado para reflejar el cambio en los elementos desde la trama 13 a la trama 14 y, por lo tanto, puede mencionarse como un vector reducido actualizado), un bit indicativo del coeficiente ambisónico ambiental de orden superior 47' especificado en el canal de trasfondo 800E en la trama 14, y un bit 757 que indica 757 que el coeficiente ambisónico de orden superior ambiental 47' no está en transición. En este caso, el segundo vector reducido para la trama posterior 14 no incluye un elemento asociado al coeficiente ambisónico de orden superior ambiental 47', por las razones indicadas anteriormente.

45 [265] En estos y en otros casos, la indicación 757 de la transición indica que el coeficiente ambisónico de orden superior ambiental 47' ha de eliminarse (tal como el coeficiente HOA ambiental 2 del canal de trasfondo 800C en la trama 12). En este caso, el dispositivo de decodificación de audio 24 puede configurarse para realizar una operación de fundido final con respecto al coeficiente ambisónico de orden superior ambiental 47' durante la trama 12. El dispositivo de decodificación de audio 24 puede estar configurado para realizar la operación complementaria con respecto al elemento correspondiente del vector reducido 55_k especificado en el canal de primer plano 800B en la trama 12. En otras palabras, el dispositivo de decodificación de audio 24 puede configurarse para realizar una operación de fundido inicial con respecto al elemento vectorial durante la trama 12, para compensar el cambio de energía que se produce como resultado del fundido final del coeficiente ambiental de orden superior ambiental 47'.

55 [266] En estos y en otros casos, la indicación 757 de la transición indica que el coeficiente ambisónico de orden superior ambiental 47' ha de ser eliminado (tal como el coeficiente HOA ambiental 4 del canal de trasfondo 800A en la trama 13). En este caso, el dispositivo de decodificación de audio 24 puede configurarse para realizar una operación de fundido final con respecto al coeficiente ambisónico de orden superior ambiental 47' durante la trama 12. El dispositivo de decodificación de audio 24 puede estar configurado para realizar la operación complementaria con respecto al elemento correspondiente del vector reducido 55_k especificado en el canal de primer plano 800B en la trama 13. En otras palabras, el dispositivo de decodificación de audio 24 puede configurarse para realizar una operación de fundido inicial con respecto al elemento vectorial (VecV [4]) durante la trama 13, para compensar el cambio de energía que se produce como resultado del fundido final del coeficiente ambisónico ambiental de orden superior 47'.

65 [267] En estos y otros casos, la indicación 757 de la transición indica que el coeficiente ambisónico de orden

superior ambiental 47' ha de incorporarse (tal como el coeficiente HOA ambiental 5 especificado en el canal de trasfondo 800E en la trama 13). En este caso, el dispositivo de decodificación de audio 24 puede configurarse para realizar una operación de fundido inicial con respecto al coeficiente ambisónico ambiental de orden superior 47' durante la trama 13. El dispositivo de decodificación de audio 24 puede estar configurado para realizar la operación complementaria con respecto al elemento correspondiente del vector reducido 55_k" especificado en el canal de primer plano 800B en la trama 13. En otras palabras, el dispositivo de decodificación de audio 24 puede configurarse para realizar una operación de fundido final con respecto al elemento vectorial durante la trama 13 para compensar el cambio de energía que se produce como resultado del fundido inicial del coeficiente ambisónico de orden superior ambiental 47'.

[268] En estos y otros casos, el dispositivo de decodificación de audio 24 puede, de forma similar al dispositivo de codificación de audio 20, configurarse para realizar los aspectos independientes de las técnicas descritas con más detalle anteriormente junto con los aspectos de transición de las técnicas descritas anteriormente.

[269] La FIG. 9 es un diagrama que ilustra el fundido final de un coeficiente HOA ambiental adicional, el fundido inicial de una correspondiente contribución reconstruida de los distintos componentes y una suma de los coeficientes HOA y la contribución reconstruida. Tres gráficos 850, 852 y 854 se muestran en el ejemplo de la FIG. 9. El gráfico 850 ilustra un coeficiente HOA ambiental adicional que se elimina en 512 muestras. El gráfico 852 muestra el objeto de audio reconstruido (que ha sido reconstruido usando coeficientes incorporados para el vector V, como se ha descrito anteriormente). El gráfico 854 muestra la suma de los coeficientes HOA y la contribución reconstruida, donde no se introducen distorsiones en este ejemplo (donde las distorsiones podrían referirse a "agujeros" en el campo de sonido debido a una pérdida de energía).

[270] Las técnicas anteriores pueden realizarse con respecto a cualquier cantidad de contextos y ecosistemas de audio diferentes. A continuación se describen varios contextos ejemplares, aunque las técnicas deberían limitarse a los contextos ejemplares. Un ecosistema ejemplar de audio puede incluir contenido de audio, estudios de cine, estudios de música, estudios de audio de juegos, contenido de audio basado en canales, motores de codificación, derivaciones de audio de juegos, motores de codificación / reproducción y sistemas de suministro.

[271] Los estudios de cine, los estudios de música y los estudios de audio de juegos pueden recibir contenido de audio. En algunos ejemplos, el contenido de audio puede representar el resultado de una adquisición. Los estudios de cine pueden emitir contenido de audio basado en canales (por ejemplo, en 2.0, 5.1 y 7.1), tal como utilizando una estación de trabajo de audio digital (DAW). Los estudios de música pueden emitir contenido de audio basado en canales (por ejemplo, en 2.0 y 5.1), tal como utilizando una DAW. En cualquier caso, los motores de codificación pueden recibir y codificar el contenido de audio basado en canales basándose en uno o más códecs (por ejemplo, AAC, AC3, Dolby True HD, Dolby Digital Plus y DTS Master Audio) para emitir por los sistemas de suministro. Los estudios de audio de juegos pueden emitir uno o más fragmentos de audio de juego, por ejemplo, mediante una DAW. Los motores de codificación/reproducción de audios de juego pueden codificar y/o reproducir las derivaciones de audio como contenido de audio basado en canales para su emisión por los sistemas de suministro. Otro ejemplo de contexto en el que se pueden realizar las técnicas comprende un ecosistema de audio que puede incluir grabación difundida de objetos de audio, sistemas de audio profesionales, captura de consumidores en el dispositivo, formato de audio HOA, representación en el dispositivo, audio del consumidor, televisión y accesorios, y sistemas de audio para el automóvil.

[272] Los objetos de audio de grabación difundida, los sistemas de audio profesionales 320 y la captura de consumidores en el dispositivo pueden codificar todos sus resultados utilizando el formato de audio HOA. De esta manera, el contenido de audio puede codificarse utilizando el formato de audio HOA en una representación única que puede reproducirse utilizando la reproducción en dispositivo, el audio de consumo, la televisión y accesorios y los sistemas de audio para automóviles. En otras palabras, la representación única del contenido de audio se puede reproducir en un sistema genérico de reproducción de audio (es decir, en lugar de requerir una configuración particular tal como 5.1, 7.1, etc.), tal como un sistema de reproducción de audio 16.

[273] Otros ejemplos de contexto en los que se pueden realizar las técnicas incluyen un ecosistema de audio que puede incluir elementos de adquisición y elementos de reproducción. Los elementos de adquisición pueden incluir dispositivos de adquisición cableados y/o inalámbricos (por ejemplo, micrófonos Eigen), captura de sonido envolvente en dispositivo y dispositivos móviles (por ejemplo, teléfonos inteligentes y tabletas). En algunos ejemplos, los dispositivos de adquisición cableados y/o inalámbricos pueden estar acoplados al dispositivo móvil mediante uno o más canales de comunicación cableados y/o inalámbricos.

[274] De acuerdo a una o más técnicas de esta divulgación, el dispositivo móvil puede usarse para adquirir un campo sonoro. Por ejemplo, el dispositivo móvil puede adquirir un campo sonoro mediante dispositivos de adquisición cableados y/o inalámbricos y/o la captura de sonido envolvente en el dispositivo (por ejemplo, una pluralidad de micrófonos integrados en el dispositivo móvil). El dispositivo móvil puede entonces codificar el

campo de sonido adquirido en los coeficientes HOA para la reproducción por uno o más de los elementos de reproducción. Por ejemplo, un usuario del dispositivo móvil puede grabar (adquirir un campo sonoro de) un suceso en vivo (por ejemplo, una reunión, una conferencia, una obra de teatro, un concierto, etc.) y codificar la grabación en coeficientes HOA.

5
 [275] El dispositivo móvil también puede utilizar uno o más de los elementos de reproducción para reproducir el campo sonoro codificado por HOA. Por ejemplo, el dispositivo móvil puede decodificar el campo sonoro codificado por HOA y emitir una señal a uno o más de los elementos de reproducción, que hace que los uno o más de los elementos de reproducción recreen el campo sonoro. Como ejemplo, el dispositivo móvil puede utilizar los canales de comunicación inalámbricos y/o inalámbricos 338 para emitir la señal a uno o más altavoces (por ejemplo, matrices de altavoces, barras de sonido, etc.). Como otro ejemplo, el dispositivo móvil puede utilizar soluciones de acoplamiento para emitir la señal a una o más estaciones de acoplamiento y/o uno o más altavoces acoplados (por ejemplo, sistemas de sonido en coches inteligentes y/u hogares). Como otro ejemplo, el dispositivo móvil puede utilizar la reproducción de auriculares para emitir la señal a un conjunto de auriculares, por ejemplo, para crear un sonido binaural realista.

10
 [276] En algunos ejemplos, un dispositivo móvil particular puede adquirir un campo sonoro tridimensional y reproducir el mismo campo sonoro tridimensional en un momento posterior. En algunos ejemplos, el dispositivo móvil puede adquirir un campo sonoro tridimensional, codificar el campo sonoro tridimensional en HOA y transmitir el campo sonoro tridimensional codificado a otros uno o más dispositivos (por ejemplo, otros dispositivos móviles y/u otros dispositivos no móviles) para su reproducción.

15
 [277] Otro contexto más en el que las técnicas se pueden realizar incluye un ecosistema de audio que puede incluir contenido de audio, estudios de juegos, contenido de audio codificado, motores de reproducción y sistemas de suministro. En algunos ejemplos, los estudios de juego pueden incluir una o más DAW que pueden prestar soporte a la edición de señales HOA. Por ejemplo, las una o más DAW pueden incluir acoples HOA y/o herramientas que pueden ser configuradas para funcionar con (por ejemplo, trabajar con) uno o más sistemas de audio de juegos. En algunos ejemplos, los estudios de juego pueden producir nuevos formatos de derivación que den soporte a HOA. En cualquier caso, los estudios de juego pueden emitir contenido de audio codificado a los motores de reproducción que pueden representar un campo sonoro para su reproducción mediante los sistemas de suministro.

20
 [278] Las técnicas también se pueden realizar con respecto a dispositivos ejemplares de adquisición de audio. Por ejemplo, las técnicas se pueden realizar con respecto a un micrófono Eigen que puede incluir una pluralidad de micrófonos que están configurados colectivamente para grabar un campo de sonido tridimensional. En algunos ejemplos, la pluralidad de micrófonos del micrófono Eigen puede estar situada en la superficie de una bola esencialmente esférica con un radio de aproximadamente 4 cm. En algunos ejemplos, el dispositivo de codificación de audio 20 puede estar integrado en el micrófono Eigen a fin de emitir un flujo de bits 21 directamente desde el micrófono.

25
 [279] Otro contexto ejemplar de adquisición de audio puede incluir un camión de producción que puede configurarse para recibir una señal desde uno o más micrófonos, tal como uno o más micrófonos Eigen. El camión de producción también puede incluir un codificador de audio, tal como el codificador de audio 20 de la FIG. 3.

30
 [280] El dispositivo móvil puede incluir también, en algunos casos, una pluralidad de micrófonos que están configurados colectivamente para grabar un campo sonoro tridimensional. En otras palabras, la pluralidad de micrófonos puede tener diversidad X, Y, Z. En algunos ejemplos, el dispositivo móvil 374 puede incluir un micrófono que puede ser girado para proporcionar diversidad X, Y, Z con respecto a otros uno o más micrófonos del dispositivo móvil. El dispositivo móvil también puede incluir un codificador de audio, tal como el codificador de audio 20 de la FIG. 3.

35
 [281] Un dispositivo de captura de vídeo robustecido puede configurarse adicionalmente para grabar un campo sonoro tridimensional. En algunos ejemplos, el dispositivo de captura de vídeo robustecido puede estar unido a un casco de un usuario que participa en una actividad. Por ejemplo, el dispositivo de captura de vídeo robustecido puede estar unido a un casco de un usuario de navegación en aguas bravas. De esta manera, el dispositivo de captura de vídeo robustecido puede capturar un campo sonoro tridimensional que representa la acción alrededor del usuario (por ejemplo, agua que se estrella detrás del usuario, otro navegante que habla delante del usuario, etc.).

40
 [282] Las técnicas también se pueden realizar con respecto a un dispositivo móvil mejorado accesorio, que se puede configurar para grabar un campo de sonido tridimensional. En algunos ejemplos, el dispositivo móvil puede ser similar a los dispositivos móviles expuestos anteriormente, con la adición de uno o más accesorios. Por ejemplo, un micrófono Eigen se puede unir al dispositivo móvil mencionado anteriormente para formar un dispositivo móvil mejorado accesorio. De esta manera, el dispositivo móvil mejorado accesorio puede capturar una versión del campo sonoro tridimensional de calidad superior que utilizando únicamente componentes de

captura de sonido integrados en el dispositivo móvil mejorado accesorio.

5 **[283]** Los dispositivos ejemplares de reproducción de audio que pueden llevar a cabo diversos aspectos de las técnicas descritas en esta divulgación se exponen adicionalmente a continuación. De acuerdo a una o más técnicas de esta divulgación, los altavoces y/o las barras de sonido pueden disponerse en cualquier configuración arbitraria mientras se sigue reproduciendo un campo sonoro tridimensional. Además, en algunos ejemplos, los dispositivos de reproducción de auriculares pueden estar acoplados a un decodificador 24 mediante una conexión por cable o bien inalámbrica. De acuerdo a una o más técnicas de esta divulgación, se puede utilizar una única representación genérica de un campo sonoro para reproducir el campo sonoro sobre cualquier combinación de altavoces, barras de sonido y dispositivos de reproducción de auriculares.

15 **[284]** Un cierto número de diferentes entornos ejemplares de reproducción de audio también pueden ser adecuados para realizar diversos aspectos de las técnicas descritas en esta divulgación. Por ejemplo, un entorno de reproducción de altavoces 5.1, un entorno de reproducción de altavoces 2.0 (por ejemplo, estéreo), un entorno de reproducción de altavoces 9.1 con altavoces frontales de altura completa, un entorno de reproducción de altavoces 22.2, un entorno de reproducción de altavoces 16.0, un entorno de reproducción de altavoces automotrices y un dispositivo móvil con entorno de reproducción por audífono pueden ser entornos adecuados para realizar diversos aspectos de las técnicas descritas en esta divulgación.

20 **[285]** De acuerdo a una o más técnicas de esta divulgación, se puede utilizar una única representación genérica de un campo sonoro para representar el campo sonoro en cualquiera de los entornos de reproducción precedentes. Adicionalmente, las técnicas de esta divulgación permiten a un reproductor reproducir un campo sonoro desde una representación genérica para la reproducción en entornos de reproducción distintos al descrito anteriormente. Por ejemplo, si las consideraciones de diseño prohíben la colocación correcta de los altavoces de acuerdo a un entorno de reproducción de altavoces 7.1 (por ejemplo, si no es posible colocar un altavoz envolvente derecho), las técnicas de esta divulgación permiten que un reproductor compense con los otros 6 altavoces de modo que la reproducción se pueda lograr en un entorno de reproducción de altavoces 6.1.

30 **[286]** Además, un usuario puede observar un juego deportivo mientras utiliza los auriculares. De acuerdo a una o más técnicas de esta divulgación, el campo sonoro tridimensional del juego deportivo puede ser adquirido (por ejemplo, uno o más micrófonos Eigen pueden ser colocados en y / o alrededor del estadio de béisbol), los coeficientes HOA correspondientes al campo sonoro tridimensional pueden obtenerse y transmitirse a un decodificador, el decodificador puede reconstruir el campo sonoro tridimensional basándose en los coeficientes HOA y emitir el campo sonoro tridimensional reconstruido a un reproductor, el reproductor puede obtener una indicación en cuanto al tipo de entorno de reproducción (por ejemplo, los auriculares) y reproducir el campo sonoro tridimensional reconstruido en señales que hacen que los auriculares emitan una representación del campo sonoro tridimensional del juego deportivo.

40 **[287]** En cada uno de los diversos casos descritos anteriormente, debería entenderse que el dispositivo de codificación de audio 20 puede realizar un procedimiento o comprender de otro modo medios para realizar cada etapa del procedimiento, para llevar a cabo el cual el dispositivo de codificación de audio 20 está configurado en algunos casos; estos medios pueden comprender uno o más procesadores. En algunos casos, los uno o más procesadores pueden representar un procesador de propósito especial configurado por medio de instrucciones almacenadas en un medio de almacenamiento no transitorio legible por ordenador. En otras palabras, varios aspectos de las técnicas en cada uno de los conjuntos de ejemplos de codificación pueden proveer un medio de almacenamiento no transitorio legible por ordenador que tiene almacenadas en el mismo instrucciones que, cuando se ejecutan, hacen que los uno o más procesadores realicen el procedimiento, para llevar a cabo el cual el dispositivo de codificación de audio 20 ha sido configurado.

50 **[288]** En uno o más ejemplos, las funciones descritas pueden implementarse en hardware, software, firmware o cualquier combinación de los mismos. Si se implementan en software, las funciones, como una o más instrucciones o código, pueden almacenarse en, y transmitirse por, un medio legible por ordenador, y ejecutarse mediante una unidad de procesamiento basada en hardware. Los medios legibles por ordenador pueden incluir medios de almacenamiento legibles por ordenador, que corresponden a un medio tangible tal como medios de almacenamiento de datos. Los medios de almacenamiento de datos pueden ser medios disponibles cualesquiera a los que se puede acceder desde uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en esta divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

60 **[289]** Análogamente, en cada uno de los diversos casos descritos anteriormente, debería entenderse que el dispositivo de decodificación de audio 24 puede realizar un procedimiento o comprender de otro modo medios para realizar cada etapa del procedimiento para realizar el cual está configurado el dispositivo de decodificación de audio 24. En algunos casos, estos medios pueden comprender uno o más procesadores. En algunos casos, los uno o más procesadores pueden representar un procesador de propósito especial, configurado por medio de instrucciones almacenadas en un medio de almacenamiento no transitorio legible por

ordenador. En otras palabras, varios aspectos de las técnicas en cada uno de los conjuntos de ejemplos de codificación pueden proporcionar un medio de almacenamiento no transitorio legible por ordenador que tiene almacenadas en él instrucciones que, cuando se ejecutan, hacen que uno o más procesadores realicen el procedimiento para realizar el cual el dispositivo de decodificación de audio 24 ha sido configurado.

5
[290] A modo de ejemplo, y no de limitación, tales medios de almacenamiento legibles por ordenador pueden comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento de disco óptico, almacenamiento de disco magnético u otros dispositivos de almacenamiento magnético, memoria flash o cualquier otro medio que pueda usarse para almacenar código de programa deseado en forma de instrucciones o estructuras de datos y al que pueda accederse mediante un ordenador. Sin embargo, debería entenderse que los medios de almacenamiento legibles por ordenador y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales u otros medios transitorios, sino que, en cambio, se orientan a medios de almacenamiento tangibles no transitorios. Los discos, tal como se utilizan en el presente documento, incluyen un disco compacto (CD), un disco láser, un disco óptico, un disco versátil digital (DVD), un disco flexible y un disco Blu-ray donde algunos discos reproducen usualmente los datos magnéticamente, mientras que otros discos reproducen los datos ópticamente con láseres. Las combinaciones de los anteriores deberían incluirse también dentro del alcance de los medios legibles por ordenador.

20
[291] Las instrucciones pueden ser ejecutadas por uno o más procesadores, tales como uno o más procesadores de señales digitales (DSP), microprocesadores de propósito general, circuitos integrados específicos de la aplicación (ASIC), formaciones lógicas programables en el terreno (FPGA) u otros circuitos lógicos integrados o discretos equivalentes. Por consiguiente, el término "procesador", tal como se usa en el presente documento, puede referirse a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento puede proporcionarse dentro de módulos de hardware y/o software dedicados, configurados para la codificación y la decodificación, o incorporarse en un códec combinado. Además, las técnicas podrían implementarse completamente en uno o más circuitos o elementos lógicos.

30
[292] Las técnicas de esta divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, que incluyen un equipo manual inalámbrico, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). Diversos componentes, módulos o unidades se describen en esta divulgación para enfatizar aspectos funcionales de dispositivos configurados para realizar las técnicas divulgadas, pero no requieren necesariamente su realización mediante diferentes unidades de hardware. En cambio, como se ha descrito anteriormente, diversas unidades pueden combinarse en una unidad de hardware de códec o proporcionarse por medio de un grupo de unidades de hardware interoperativas, que incluyen uno o más procesadores como los descritos anteriormente, conjuntamente con software y/o firmware adecuados.

40
[293] Se han descrito diversos aspectos de las técnicas. Estos y otros aspectos están dentro del alcance de las reivindicaciones siguientes.

REIVINDICACIONES

1. Un procedimiento para producir, mediante un dispositivo de codificación de audio, un flujo de bits (21) de datos de audio codificados, que comprende:
- 5 determinar (302) cuándo un coeficiente ambisónico de orden superior ambiental (47') está en transición durante una trama, siendo el coeficiente ambisónico de orden superior ambiental (47') representativo, al menos en parte, de un componente ambiental de un campo de sonido;
- 10 mantener (304) información de estado de transición basándose en el coeficiente ambisónico de orden superior ambiental (47'), la información de estado de transición que indica, para el coeficiente ambisónico ambiental de orden superior (47'), uno entre: un estado sin transición, un estado de fundido inicial y un estado de fundido final;
- 15 obtener (306), basándose en la información de estado de transición mantenida, un bit (757) indicativo de la transición del coeficiente ambisónico de orden superior ambiental durante la trama;
- 20 identificar un elemento de un vector (53) que está asociado al coeficiente ambisónico de orden superior ambiental (47') en transición, siendo el vector representativo, al menos en parte, de un componente espacial del campo de sonido;
- 25 generar, basándose en el vector (53), un vector reducido (55) para incluir el elemento identificado del vector para la trama; y
- 30 producir (308) el flujo de bits (21) para incluir un bit indicativo del vector reducido y el bit (757) indicativo de la transición del coeficiente ambisónico de orden superior ambiental (47') durante la trama y un bit indicativo de información de estado (812) que incluye la información de estado de transición en la trama, permitiendo el bit indicativo de la información de estado (812) decodificar el flujo de bits de los datos de audio codificados de la trama sin referencia a tramas previas del flujo de bits (21).
2. El procedimiento de la reivindicación 1, en el que la información de estado (812) incluye información de cuantización.
- 35 3. El procedimiento de la reivindicación 1, en el que la trama se emite mediante un protocolo de transmisión por flujo.
4. Un dispositivo de codificación de audio, configurado para producir un flujo de bits (21) de datos de audio codificados, comprendiendo el dispositivo de codificación de audio una memoria configurada para almacenar el flujo de bits (21) y uno o más procesadores configurados para:
- 40 determinar (302) cuándo un coeficiente ambisónico ambiental de orden superior (47') está en transición durante una trama, siendo el coeficiente ambisónico ambiental de orden superior (47') representativo, al menos en parte, de un componente ambiental de un campo de sonido;
- 45 mantener (304) información de estado de transición basándose en el coeficiente ambisónico de orden superior ambiental (47'), la información de estado de transición que indica, para el coeficiente ambisónico de orden superior ambiental (47'), uno entre: un estado sin transición, un estado de fundido inicial y un estado de fundido final;
- 50 obtener (306), basándose en la información de estado de transición mantenida, un bit (757) indicativo de la transición del coeficiente ambisónico de orden superior ambiental durante la trama;
- 55 identificar un elemento de un vector (53) que está asociado al coeficiente ambisónico de orden superior ambiental (47') en transición, siendo el vector representativo, al menos en parte, de un componente espacial del campo de sonido;
- 60 generar, basándose en el vector (53), un vector reducido para incluir el elemento identificado del vector para la trama; y
- 65 producir (308) el flujo de bits (21) para incluir un bit indicativo del vector reducido y el bit (757) indicativo de la transición del coeficiente ambisónico ambiental de orden superior (47') durante la trama y un bit indicativo de información de estado (812) que incluye la información de estado de transición en la trama, permitiendo el bit indicativo de la información de estado (812) decodificar el flujo de bits de los datos de audio codificados de la trama sin referencia a tramas previas del flujo de bits (21).

5. El dispositivo de codificación de audio de la reivindicación 4, en el que el bit indicativo de la información de estado (812) incluye información de cuantización.
- 5 6. El dispositivo de codificación de audio de la reivindicación 4, en el que la trama se emite mediante un protocolo de transmisión por flujo.
7. Un sistema que comprende el dispositivo de codificación de audio de cualquiera de las reivindicaciones 4 a 6 y una pluralidad de micrófonos para grabar un campo de sonido tridimensional.
- 10 8. Un procedimiento para decodificar, mediante un dispositivo de decodificación de audio, un flujo de bits (21) de datos de audio codificados, comprendiendo el procedimiento:
- 15 obtener, en un decodificador y a partir de una trama del flujo de bits (21), información de estado de transición de un bit indicativo de información de estado (812), indicando la información de estado de transición, para un coeficiente ambisónico de orden superior ambiental (47') representativo, al menos en parte, de un componente ambiental de un campo de sonido, uno entre: un estado sin transición, un estado de fundido inicial y un estado de fundido final, permitiendo el bit indicativo de la información de estado (812) decodificar el flujo de bits de los datos de audio codificados de la trama sin referencia a tramas previas del flujo de bits;
- 20 obtener (352), desde la trama del flujo de bits (21), un bit (757) indicativo de una transición del coeficiente ambisónico de orden superior ambiental (47') durante la trama;
- 25 mantener (354) la información de estado de transición basándose en el bit (757) indicativo de la transición del coeficiente ambisónico de orden superior ambiental (47');
- 30 obtener, de la trama, un bit indicativo de un vector reducido representativo, al menos en parte, de un componente espacial del campo de sonido, en función de si el coeficiente ambisónico de orden superior ambiental (47') está en transición, en donde el vector reducido incluye un elemento vectorial asociado al coeficiente ambisónico de orden superior ambiental (47') en transición;
- 35 determinar si se realiza una operación de fundido inicial o una operación de fundido final con respecto al coeficiente ambisónico de orden superior ambiental (47') basándose en la información de estado de transición; y
- 40 realizar (356) la operación de fundido inicial o la operación de fundido final, con respecto al coeficiente ambisónico ambiental de orden superior (47'), basándose en la determinación de si se realiza el fundido inicial o el fundido final del coeficiente ambisónico de orden superior ambiental (47').
- 45 9. El procedimiento de la reivindicación 8, en el que la obtención del bit indicativo del vector reducido comprende obtener un bit indicativo del vector reducido de acuerdo a una primera modalidad entre una pluralidad de modalidades, indicando la primera modalidad que el vector reducido incluye el elemento vectorial asociado al coeficiente ambisónico ambiental de orden superior (47') en transición, en donde la pluralidad de modalidades incluye además una segunda modalidad que indica que el elemento vectorial asociado al coeficiente ambisónico de orden superior ambiental (47') no está incluido en el vector reducido.
- 50 10. El procedimiento de la reivindicación 8, que comprende además descuantizar el vector reducido en función de la información de cuantización incluida en el bit indicativo de la información de estado (812).
- 55 11. El procedimiento de la reivindicación 8, que comprende además decodificar la trama para conmutar desde una primera representación de contenido a una segunda representación del contenido, en donde la segunda representación es diferente a la primera representación.
- 60 12. El procedimiento de la reivindicación 8, que comprende además:
- obtener, durante una trama posterior, un bit indicativo de un segundo vector reducido, un bit indicativo del coeficiente ambisónico ambiental de orden superior y un bit que indica que el coeficiente ambisónico ambiental de orden superior no está en transición,
- 65 en el que el segundo vector reducido para la trama posterior no incluye un elemento asociado al coeficiente ambisónico de orden superior ambiental para la trama siguiente.
13. Un dispositivo de decodificación de audio configurado para decodificar un flujo de bits (21) de datos de audio codificados, comprendiendo el dispositivo de decodificación de audio:

una memoria configurada para almacenar una trama del flujo de bits de datos de audio codificados;

uno o más procesadores configurados para:

- 5 obtener, desde la trama, información de estado de transición de un bit indicativo de información de estado (812), indicando la información de estado de transición un coeficiente ambisónico ambiental de orden superior (47') representativo, al menos en parte, de un componente ambiental de un campo de sonido, uno entre: un estado sin transición, un estado de fundido inicial y un estado de fundido final, permitiendo el bit indicativo de la información de estado (812) decodificar el flujo de bits de los datos de audio codificados de la trama sin referencia a tramas previas del flujo de bits;
- 10 obtener (352), de la trama, un bit indicativo de una transición del coeficiente ambisónico de orden superior ambiental (47') durante la trama;
- 15 mantener (354) la información del estado de transición basándose en el bit obtenido (757), indicativo de la transición del coeficiente ambisónico de orden superior ambiental (47');
- 20 obtener, de la trama, un bit indicativo de un vector reducido, representativo, al menos en parte, de un componente espacial del campo de sonido en función de si el coeficiente ambisónico de alto orden ambiental (47') está en transición, en donde el vector reducido incluye un elemento vectorial asociado al coeficiente ambisónico de orden superior ambiental (47') en transición;
- 25 determinar si se realiza una operación de fundido inicial o una operación de fundido final con respecto al coeficiente ambisónico de orden superior ambiental basándose en la información de estado de transición; y
- 30 realizar (356) la operación de fundido inicial o la operación de fundido final, con respecto al coeficiente ambisónico ambiental de orden superior, basándose en la determinación de si se realiza el fundido inicial o el fundido final del coeficiente ambisónico ambiental de orden superior.
- 35 **14.** El dispositivo de decodificación de audio de la reivindicación 13, en el que los uno o más procesadores están configurados para obtener el bit indicativo del vector reducido de acuerdo a una primera modalidad entre una pluralidad de modalidades, indicando la primera modalidad que el vector reducido incluye el elemento vectorial asociado al coeficiente ambisónico ambiental de orden superior en transición.
- 40 **15.** El dispositivo de decodificación de audio de la reivindicación 14, en el que la pluralidad de modalidades incluye además una segunda modalidad que indica que el elemento vectorial asociado al coeficiente ambisónico de orden superior ambiental no está incluido en el vector reducido.
- 45 **16.** El dispositivo de decodificación de audio de la reivindicación 13, en el que los uno o más procesadores están configurados además para descuantizar el vector reducido basándose en la información de cuantización incluida en el bit indicativo de la información de estado (812).
- 50 **17.** El dispositivo de decodificación de audio de la reivindicación 13, en el que los uno o más procesadores están configurados además para decodificar la trama para conmutar desde una primera representación de contenido a una segunda representación del contenido, en donde la segunda representación es diferente a la primera representación.
- 55 **18.** El dispositivo de decodificación de audio de la reivindicación 13, en el que el bit (757) indicativo de la transición indica si el coeficiente ambisónico de orden superior ambiental se elimina durante la trama.
- 19.** El dispositivo de decodificación de audio de la reivindicación 13, en el que el bit (757) indicativo de la transición indica si el coeficiente ambisónico de orden superior ambiental se incorpora durante la trama.
- 60 **20.** El dispositivo de decodificación de audio de la reivindicación 13, en el que los uno o más procesadores están configurados además para obtener, durante una trama posterior, un bit indicativo de un segundo vector reducido, un bit indicativo del coeficiente ambisónico ambiental de orden superior y un bit que indica que el coeficiente ambisónico ambiental de orden superior no está en transición, en el que el segundo vector reducido para la trama posterior no incluye un elemento asociado al coeficiente ambisónico de orden superior ambiental para la trama siguiente.
- 65 **21.** El dispositivo de decodificación de audio de la reivindicación 13, en el que los uno o más procesadores están configurados además para realizar una operación de fundido final con respecto al coeficiente ambisónico ambiental de orden superior durante la trama, y realizar una operación de fundido inicial con respecto al elemento vectorial durante la trama para compensar el cambio de energía que se produce

como resultado de la eliminación del coeficiente ambisónico ambiental de orden superior.

- 5 **22.** El dispositivo de decodificación de audio de la reivindicación 13, en el que los uno o más procesadores están configurados además para realizar una operación de fundido inicial con respecto al coeficiente ambisónico ambiental de orden superior durante la trama, y realizar una operación de fundido final con respecto al elemento vectorial durante la trama para compensar el cambio de energía que ocurre como resultado de la incorporación del coeficiente ambisónico ambiental de orden superior.
- 10 **23.** Un sistema que comprende el dispositivo de decodificación de audio de cualquiera de las reivindicaciones 13 a 22 y una pluralidad de altavoces para emitir una representación de un campo de sonido tridimensional.
- 15 **24.** Un medio de almacenamiento no transitorio legible por ordenador que tiene instrucciones almacenadas en el mismo que, al ejecutarse, hacen que uno o más procesadores de un dispositivo realicen un procedimiento según cualquiera de las reivindicaciones 1 a 3 o cualquiera de las reivindicaciones 8 a 12.

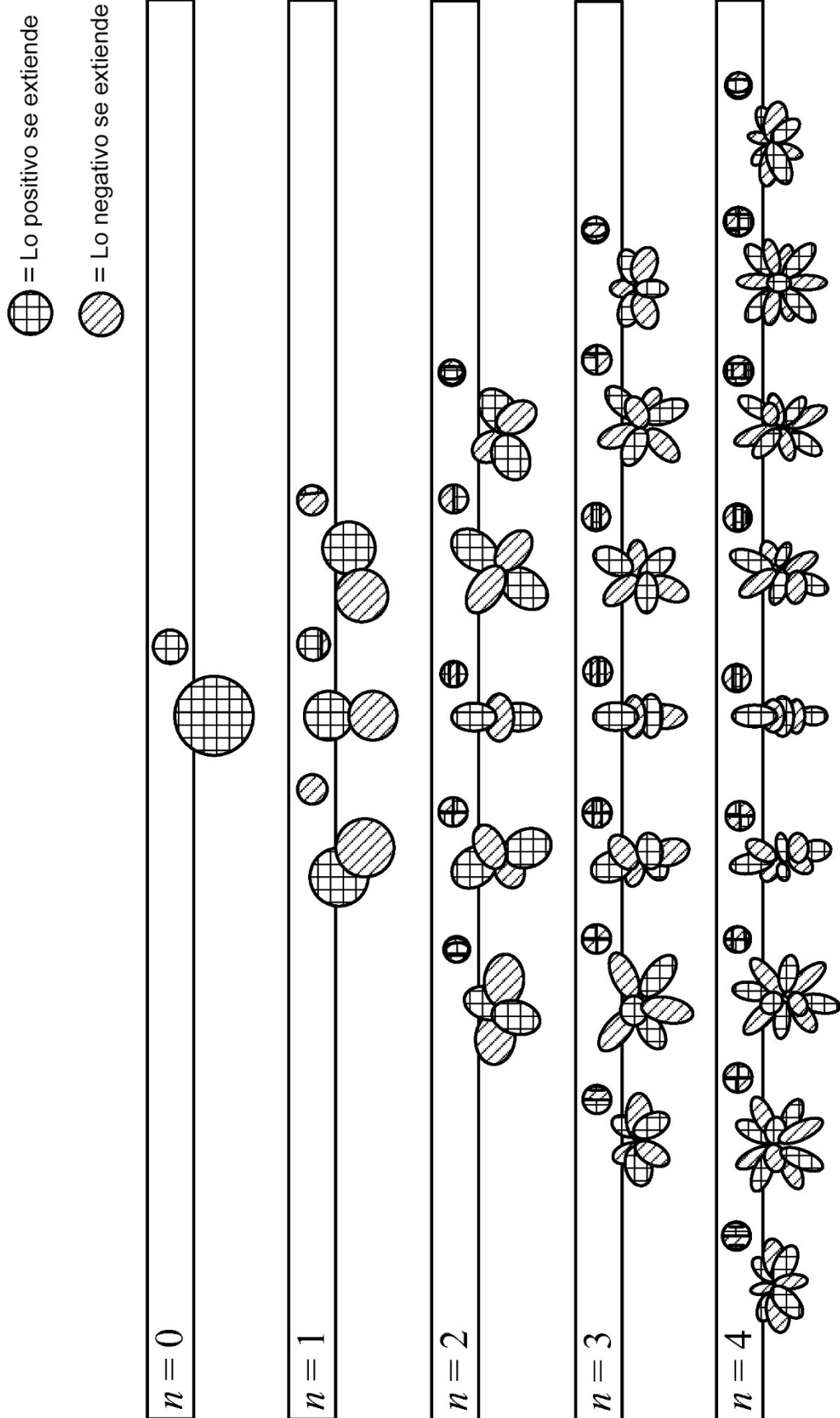


FIG. 1

10 ↗

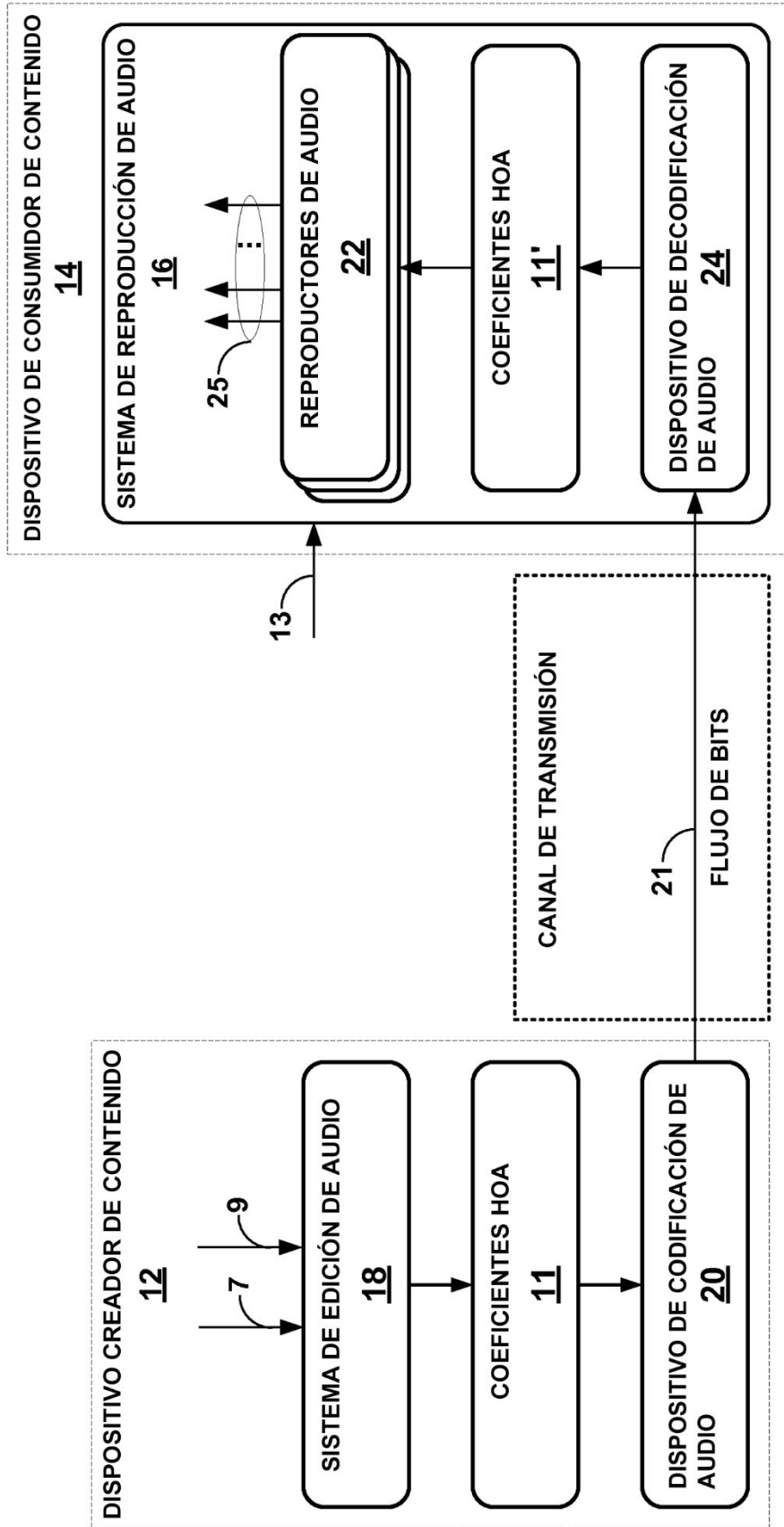


FIG. 2

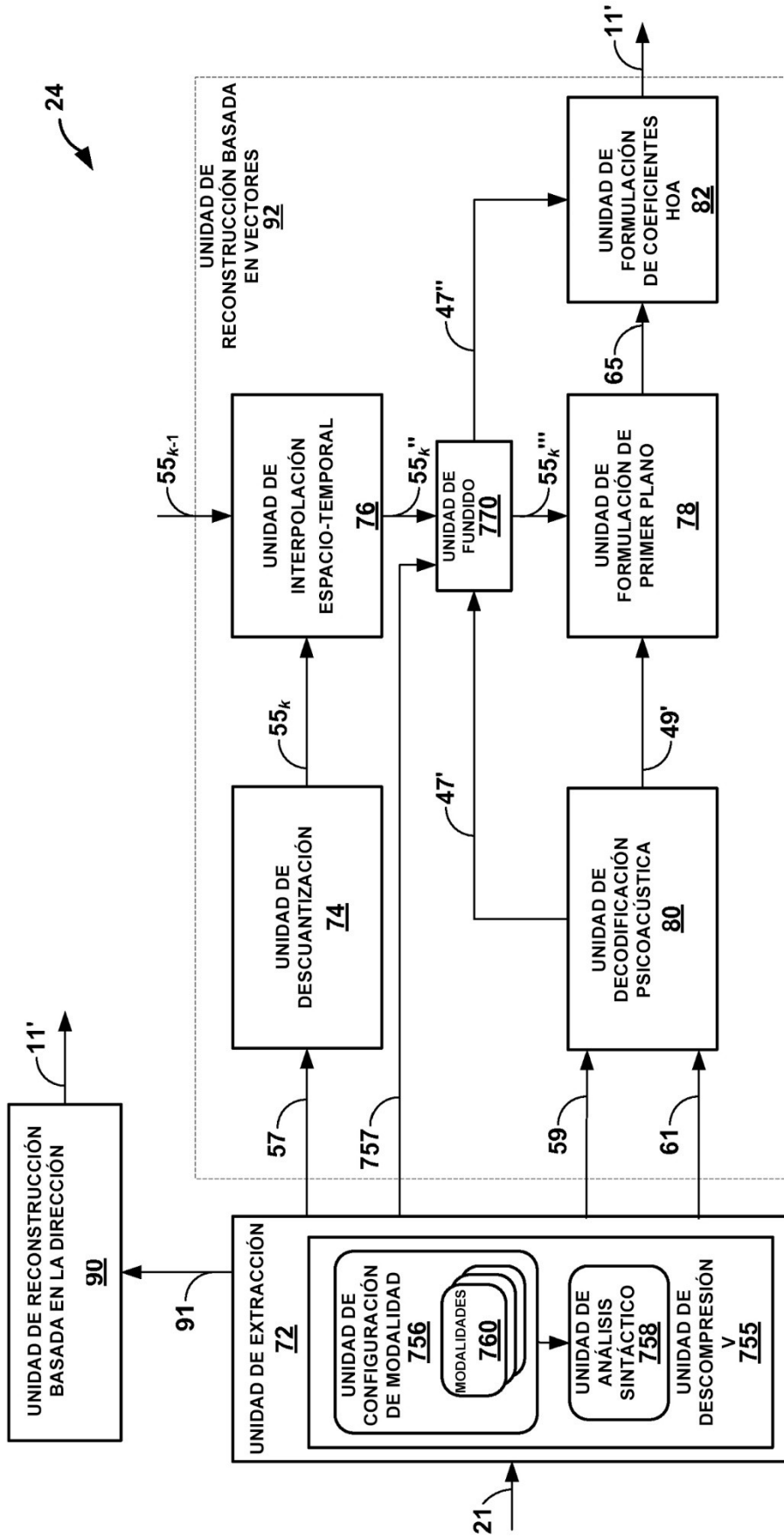


FIG. 4

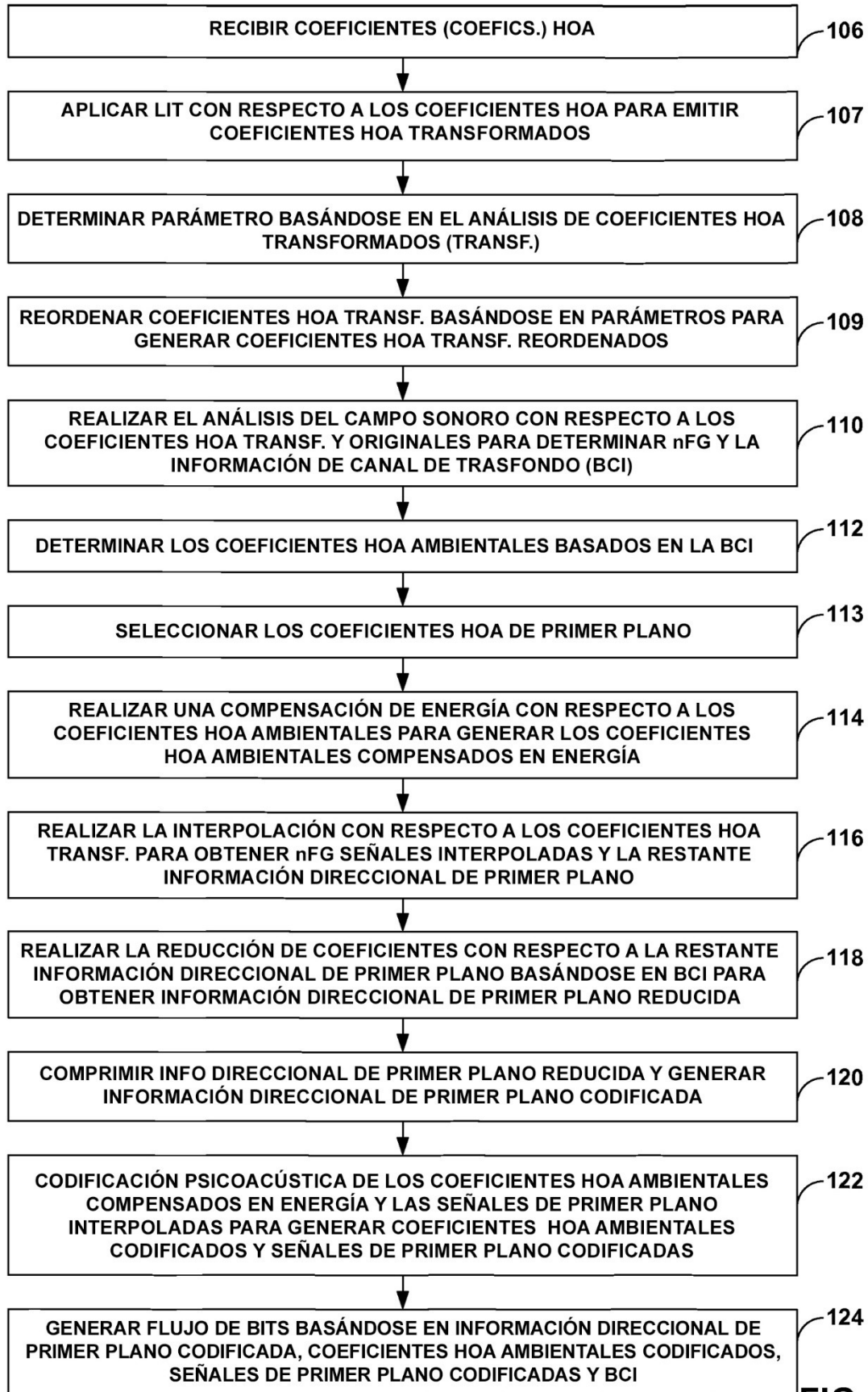


FIG. 5A

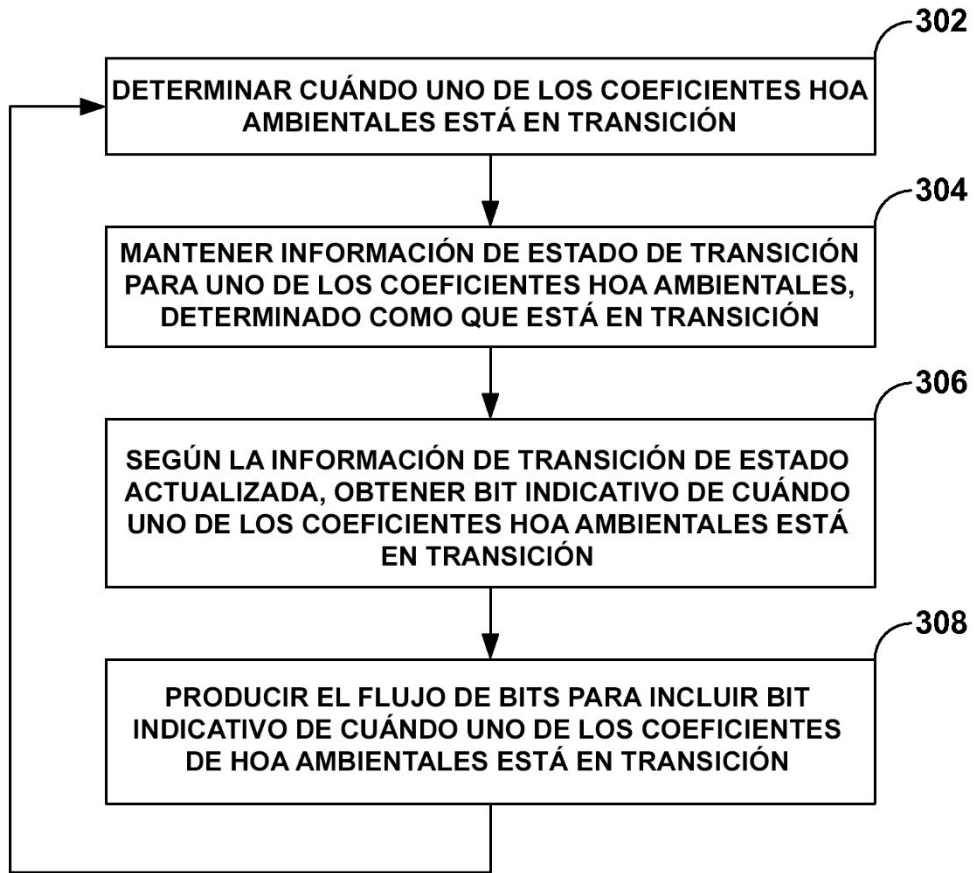


FIG. 5B

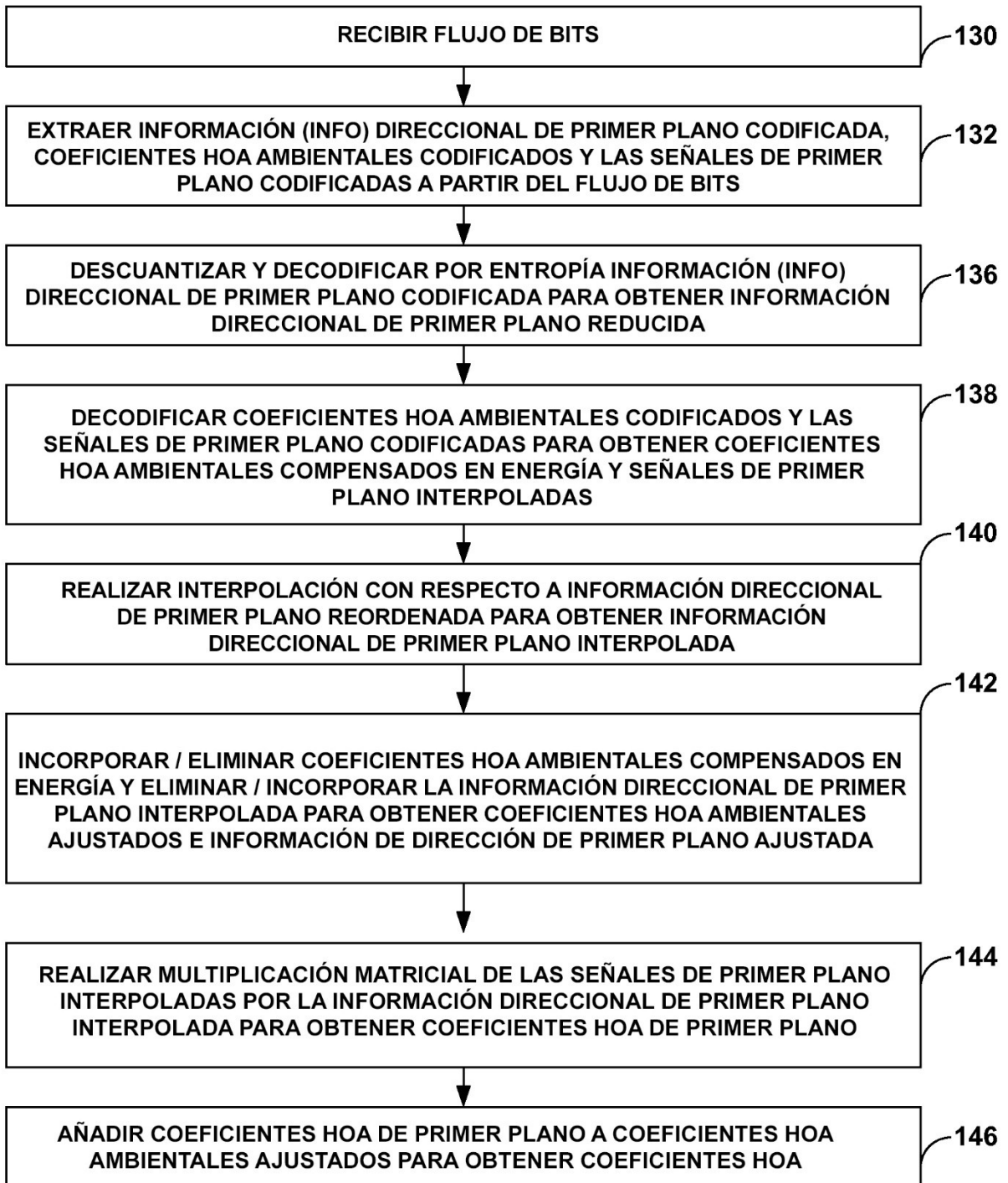


FIG. 6A

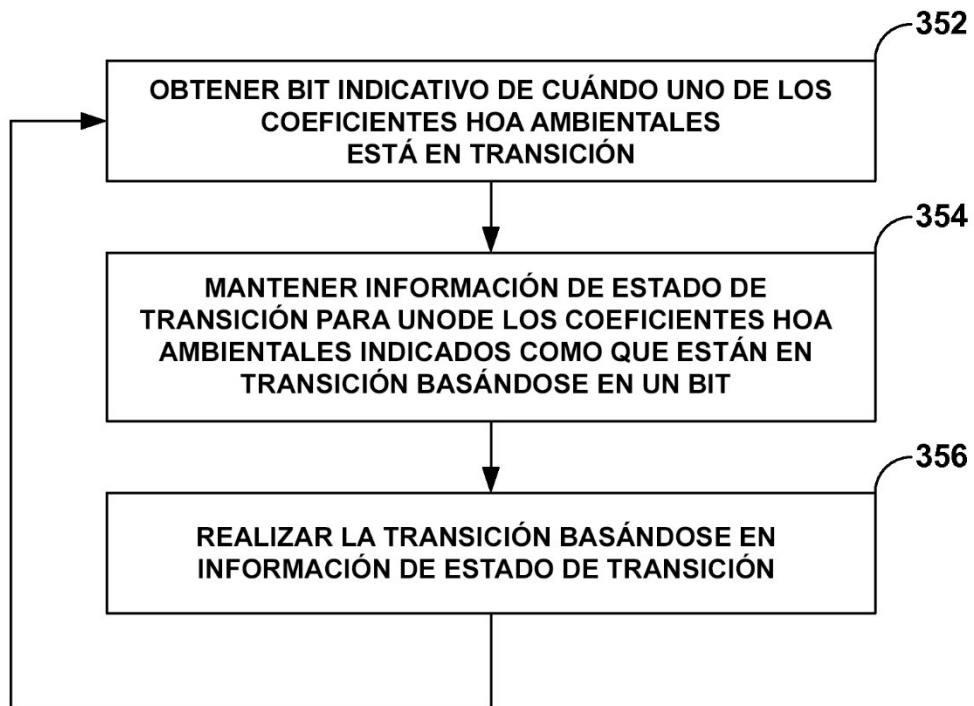


FIG. 6B

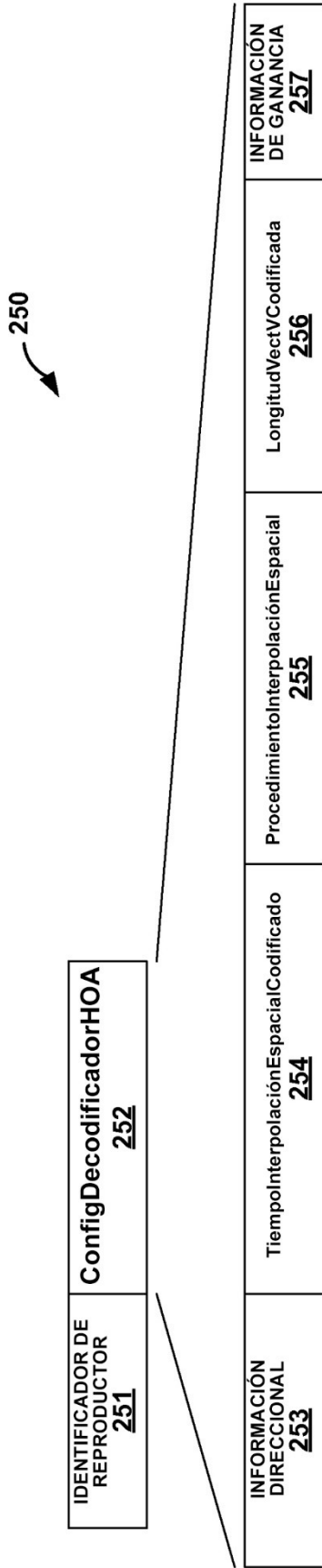


FIG. 7A

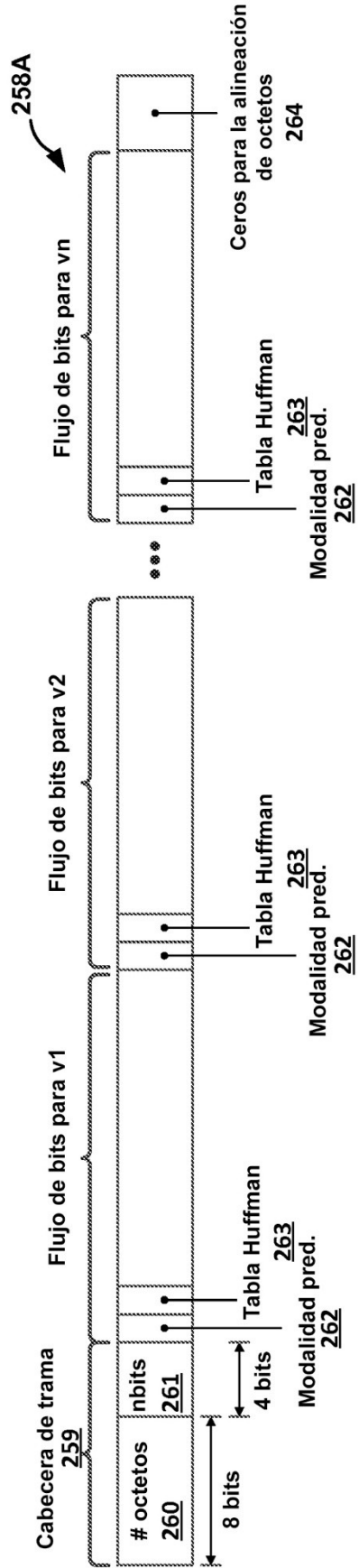


FIG. 7B

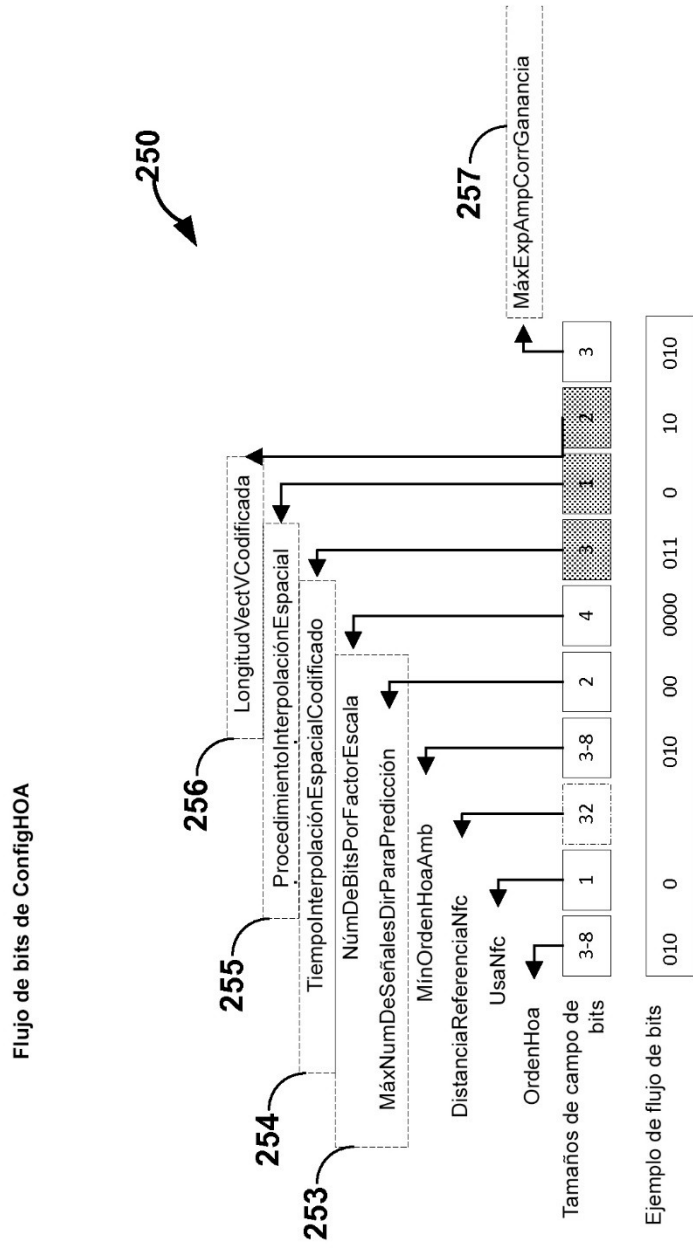
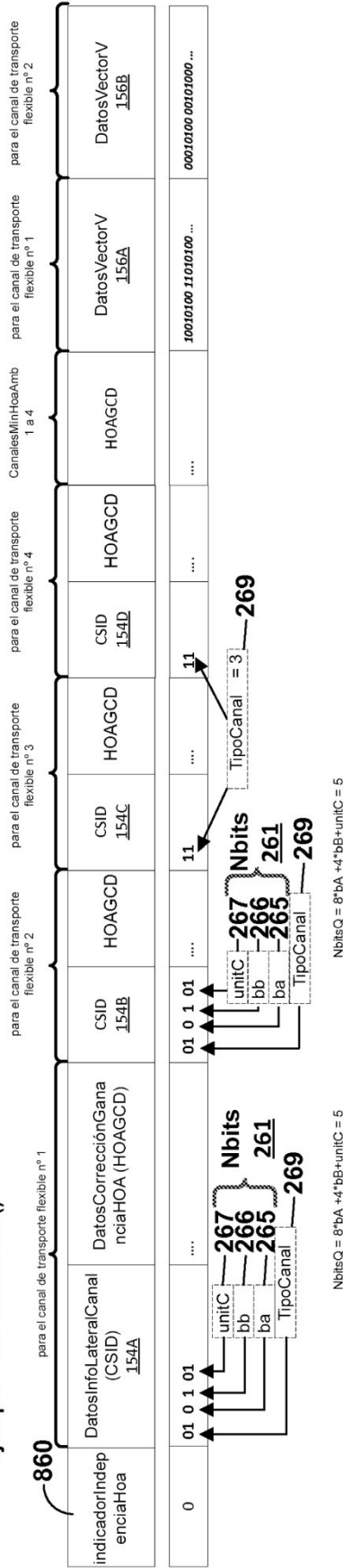


FIG. 7C

249Q

Ejemplo 1 de TramaHOA()



249R

Ejemplo 2 de TramaHOA()

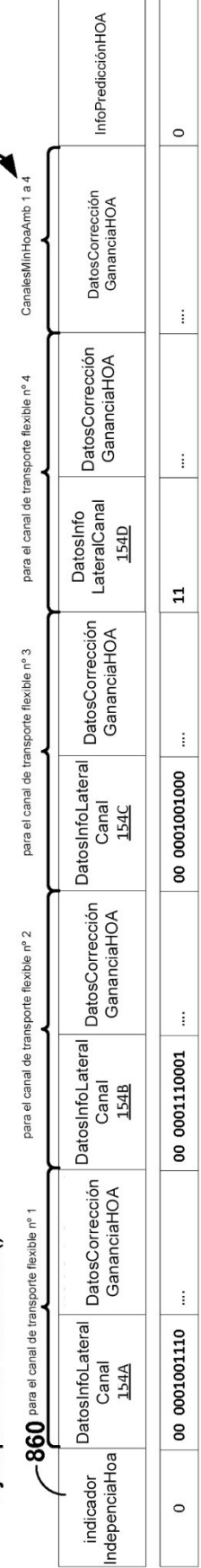


FIG. 7D

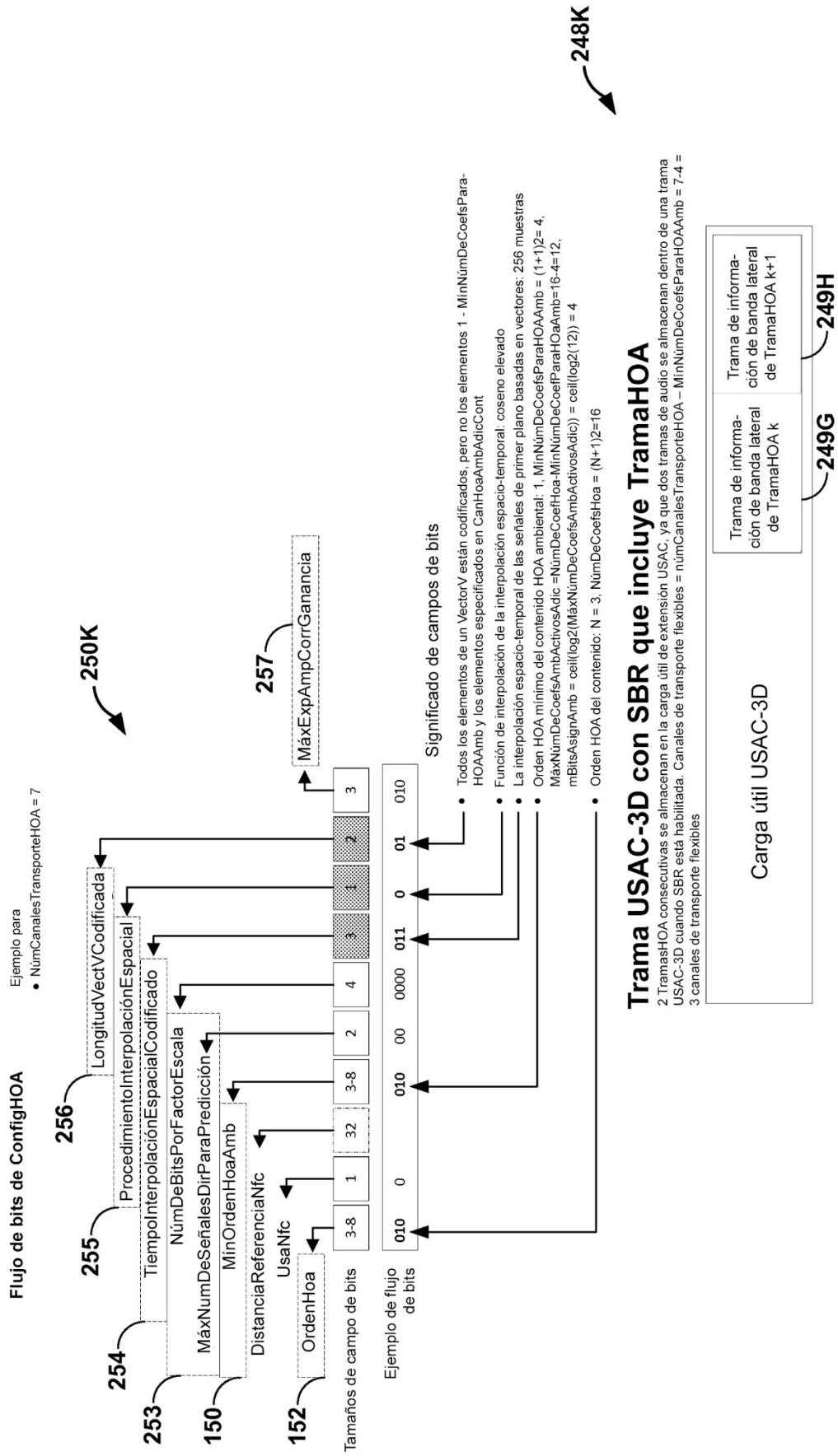
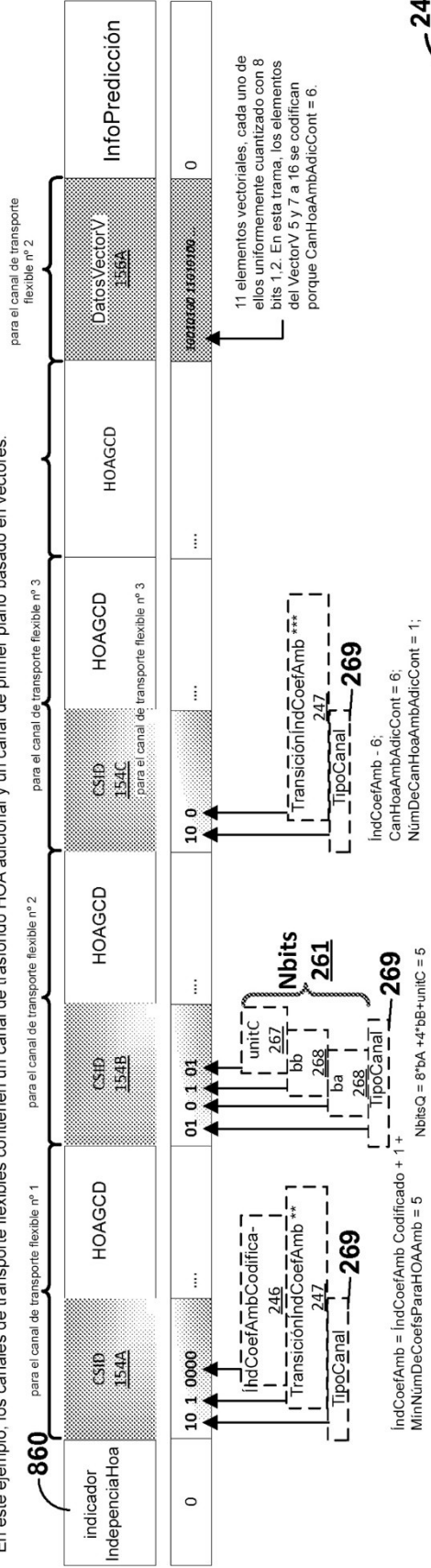


FIG. 7F

249G

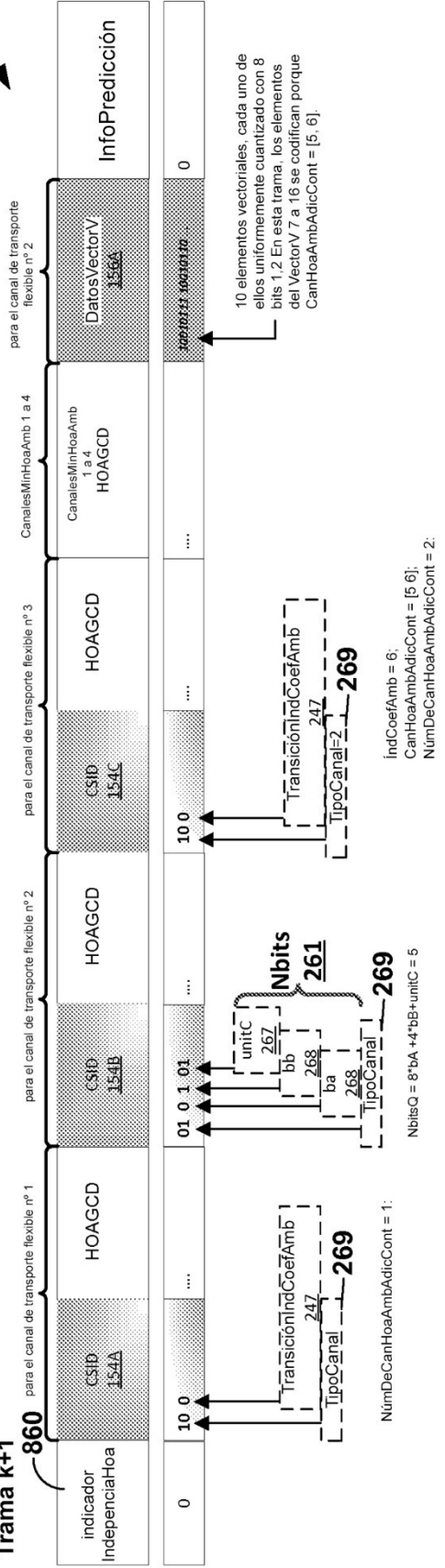
Trama k

En este ejemplo, los canales de transporte flexibles contienen un canal de trasfondo HOA adicional y un canal de primer plano basado en vectores.



249H

Trama k+1



1 Número e índices de elementos de DatosVectorV codificados especificados por el parámetro LongitudVectV Codificada, MinOrdenHoaAmb, OrdenHoa Esquema de codificación 2 señalizado por NbitsQ = 5 en la InfoLateraCanal para el canal de transporte flexible nº 2

** Para el canal de transporte flexible nº1 se supone aquí que el estado interno del decodificador es EstadoTransiciónIndCoefAmb = 2. Por lo tanto, se señala el campo de bits IndCoefAmbCodificado.

*** El canal de transporte flexible nº 3 se ha definido en una trama previa para llevar el coeficiente HOA adicional 6. Por lo tanto, el campo de bits IndCoefAmbCodificado no está señalado.

FIG. 7G

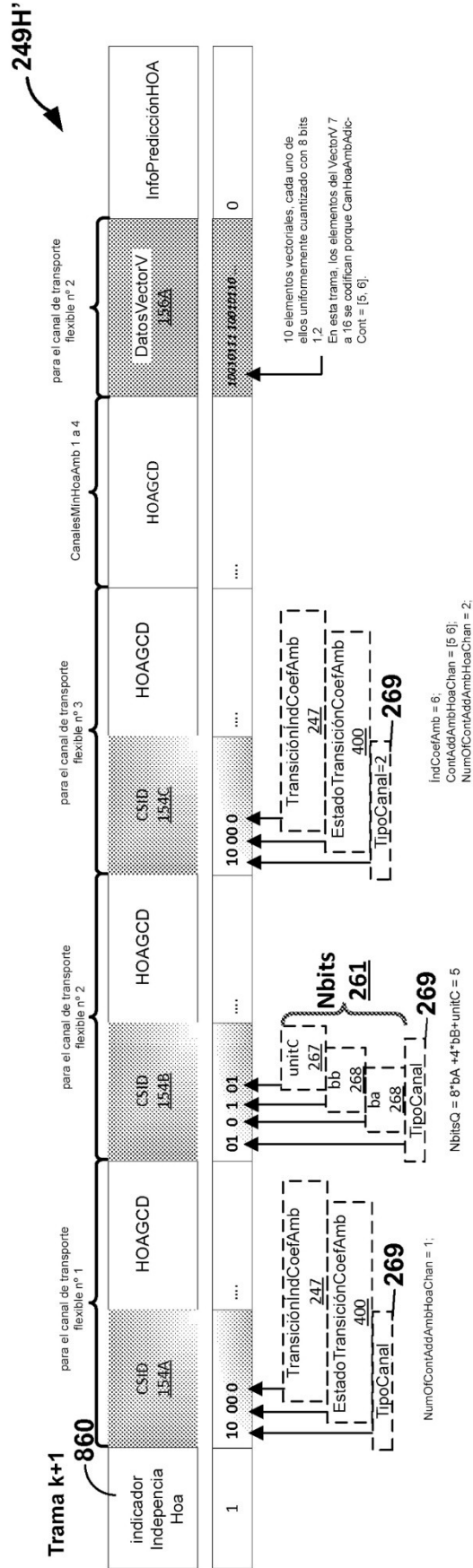


FIG. 7H

808

(i): Trama independiente

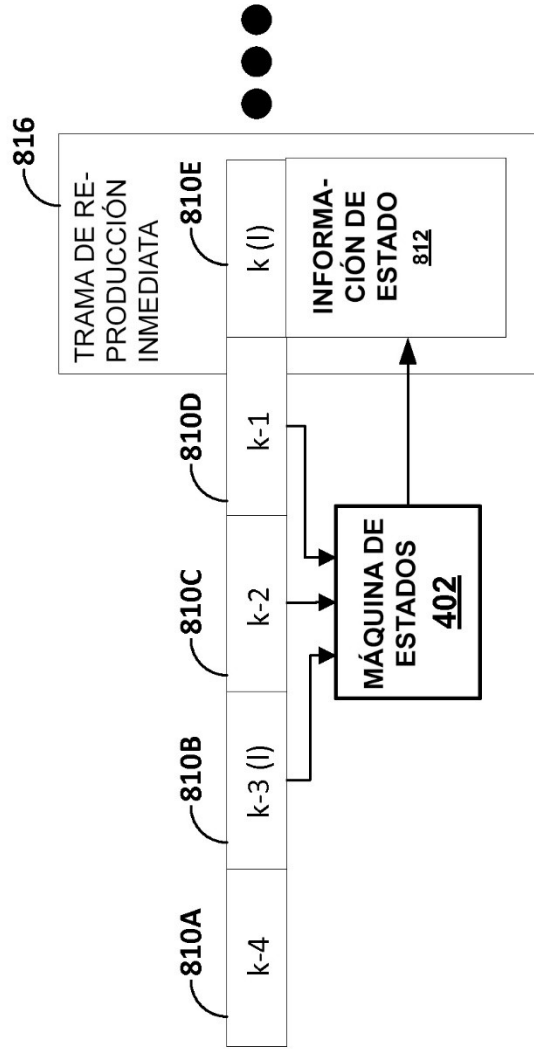


FIG. 71

450

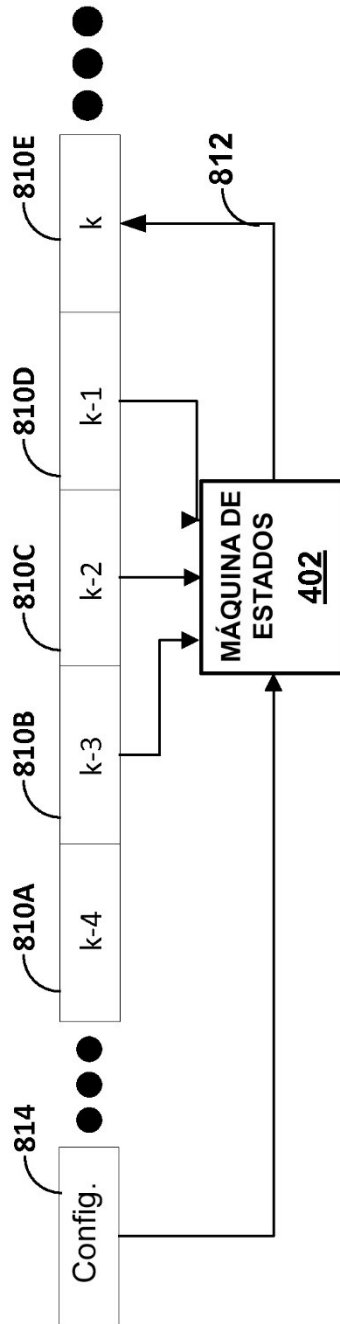


FIG. 7J

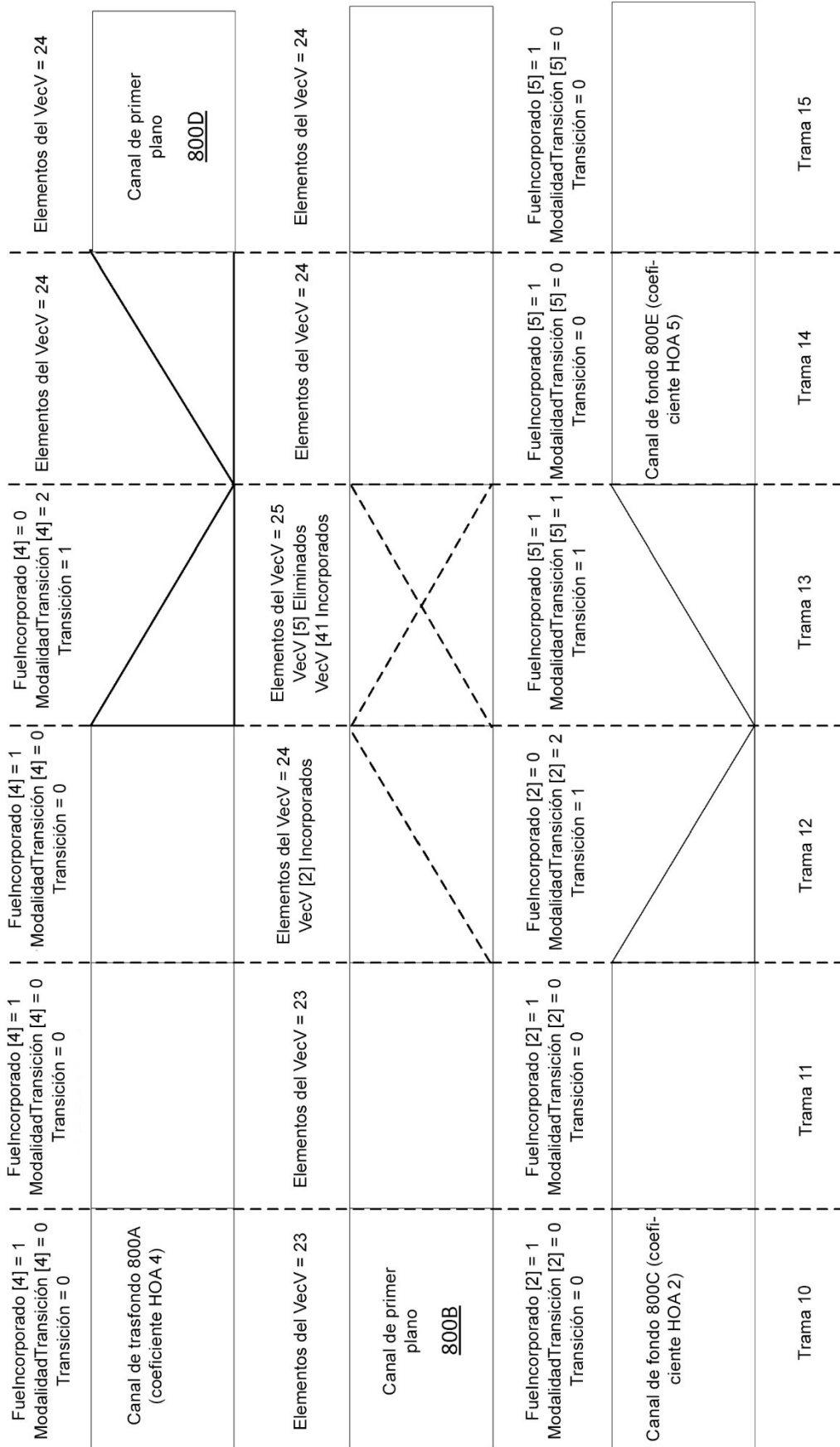


FIG. 8

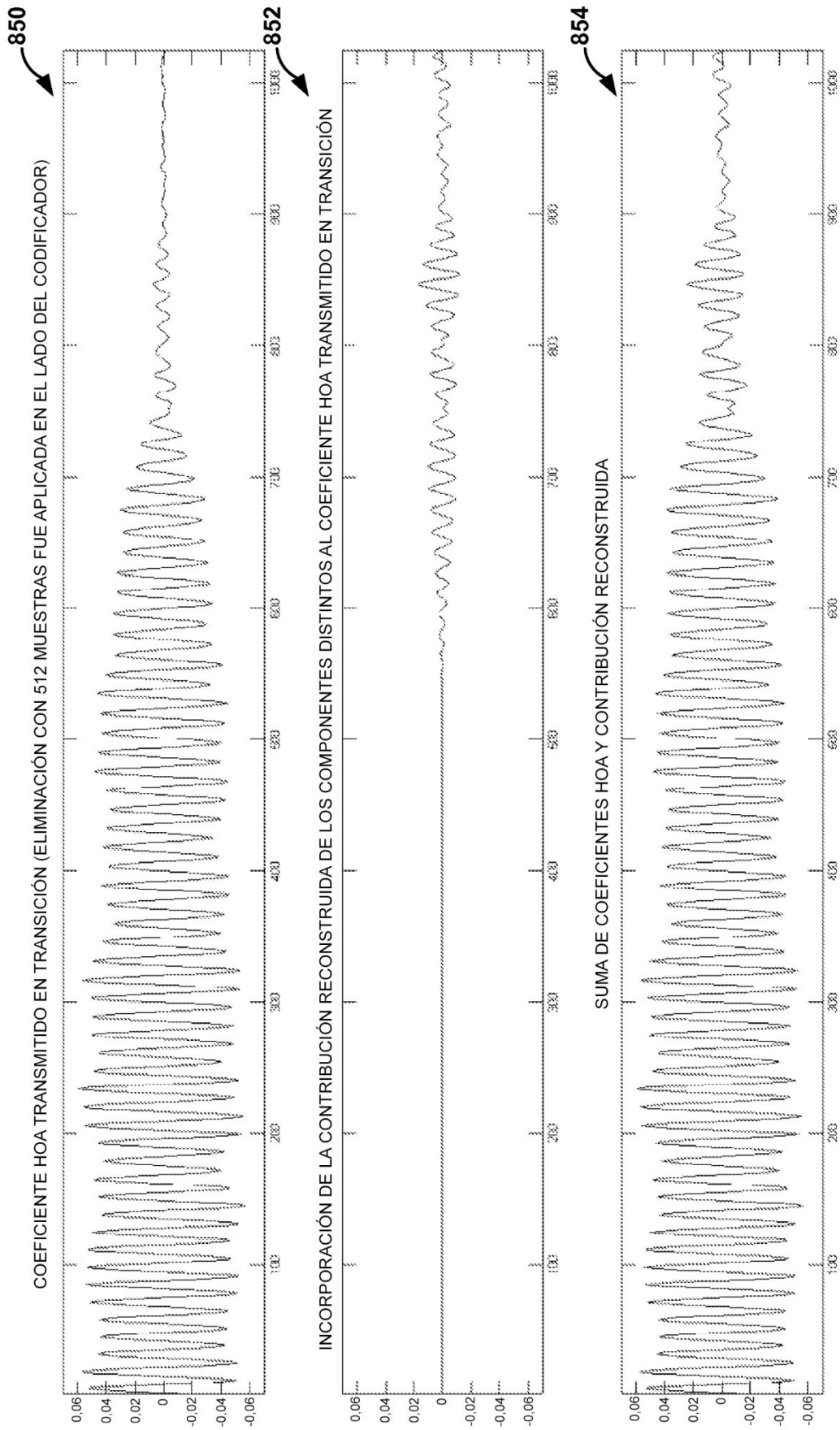


FIG. 9